Detection and Estimation of an Optical Image by Photon-Counting Techniques

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Information and Computer Science

by

Lily Lee Wang

Committee in charge:

Professor Carl W. Helstrom, Chairman
Professor Adolf W. Lohmann
Professor Elias Masry
Professor Michael J. Sharpe
Professor Harold W. Sorenson

1973
The dissertation of Lily Lee Wang is approved, and it is acceptable in quality and form for publication on microfilm:

A. Lofgren
Henri W. Grimson
Michael J. Amico
Eliot W. Haug
W. W. Kelton

Committee Chairman

University of California, San Diego

1973
Table of Contents

List of Figures ........................................ vi
List of Tables ......................................... x
Acknowledgements ......................................... xi
Vita ................................................... xii
Abstract ............................................... xiii

Introduction ........................................... 1

Chapter I. Image formation and photon-counting
statistics ............................................... 8

Image formation when diffraction theory is
applied ...................................................... 9

Intensity representation of an incoherent
object ..................................................... 14

The statistical description of a photoelectric
detector .................................................. 16

Distribution of the photoelectrons from a single
small area when the time-bandwidth product
$TW >> 1$ .................................................. 18

Distribution of the photoelectrons from two
neighborhood areas when $TW >> 1$ ................. 26

Distribution of the sum of the photoelectrons from
a large number of small areas when $TW >> 1$ .... 30

Chapter II. Binary detection of the object light from
a photosensitive surface .................................. 37

Binary hypothesis test .................................. 38

The ideal photoelectric detector ....................... 42

Threshold detector ...................................... 50

Simple detector .......................................... 53

Reliability and error probability of the ideal
detector in the absence of the background light .... 55
Reliability and error probability of the ideal detector in the presence of the background light 57

Performance of the ideal detector for a Gaussian image 65

Performance of the threshold detector for a Gaussian image in the presence of background light 82

Performance of the simple detector for a Gaussian image in the presence of background light 88

Discussion 92

Chapter III. Simultaneous detection and estimation of the mean intensity of an optical image 96

Simultaneous detection and estimation for a single parameter 97

Bayes strategy for detecting the image and estimating its intensity with a quadratic cost function 105

Maximum-likelihood strategy and the almost optimum intensity estimate 110

Performance of the almost optimum intensity estimate for a truncated Gaussian or a parabolic image 113

Discussion 134

Chapter IV. Detectors for resolving two point sources in optical communication by photo-counting techniques 136

Ideal detector for resolving two light sources 137

Ideal detector for binary detection in optical communication 141

Simple detector for detection of binary bits 149

Counting comparator for detection of binary bits 163

Image location effects on the detectors in the absence of background light 169
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discussion</td>
<td>177</td>
</tr>
<tr>
<td>List of References</td>
<td>181</td>
</tr>
<tr>
<td>Appendix A. Asymptotic expansion of the tail distribution by the method of steepest descent</td>
<td>184</td>
</tr>
<tr>
<td>Appendix B. Iteration procedure for finding the saddle point for the method of steepest descent</td>
<td>198</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Image formation system</td>
<td>9</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Block diagram of hypothesis test for binary detection.</td>
<td>41</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Photoelectrical emission surface</td>
<td>42</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Detection probability $Q_d$ as a function of $N_s$ defined in (2.38) for the ideal detector in the absence of background light; false-alarm probability $Q_0 = 0.01, 0.1, 0.2$</td>
<td>69</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Probability $Q_0$ and probability $Q_d$ as a function of decision level $g_0$ for the ideal detector calculated by the saddle point, Gaussian, and gamma approximations; $N_0 = N_s = 5$, where $N_0$ is defined in (2.57).</td>
<td>73</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Probability $Q_d$ as a function of the signal-to-noise ratio $D^2$ defined in (2.58) for the ideal detector; $N_0 = 0.5, 5, Q_0 = 10^{-3}, 10^{-5}$</td>
<td>75</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Probability $Q_d$ as a function of $N_s$ for ideal, threshold and simple detectors; $Q_0 = 10^{-3}$ at $N_0 = 0, 0.5, 5$.</td>
<td>77</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Probability $Q_d$ as a function of $N_s$ for ideal, threshold and simple detectors, $Q_0 = 10^{-5}$ at $N_0 = 0, 0.5, 5$.</td>
<td>79</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Average error probability $P_e$ in detecting the Gaussian image against the uniform background light vs. the mean number $N_s$ for ideal and simple detectors; $N_0 = 0, 0.01, 0.5, 5$.</td>
<td>81</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Probability $Q_0$ and probability $Q_d$ vs. decision level $g_0$ for threshold detector; $N_0 = 0.5$,</td>
<td></td>
</tr>
</tbody>
</table>
$D^2 = 0.1, 1, 2, 3, 4, 6$  

Figure 11 Block diagram for simultaneous detection and estimation of a single parameter $\theta$ from the signal.  

Figure 12 Image surface divided into rings.  

Figure 13 Mean intensity estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for truncated Gaussian image; $Q_0 = 10^{-2}$, total number of rings is 100. Mean background counts $\mu = 0.1$ at $R_A = \sigma = 1$ for time $T$. Curves are indexed by parameter $R_A$ defined in (3.44).  

Figure 14 Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for truncated Gaussian image; $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_A = 1$ for 100 rings for time $T$. Curves are indexed by the parameter $M$ defined in (3.41).  

Figure 15 Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for truncated Gaussian image; $Q_0 = 10^{-2}$, $M = 100$, $\mu = 0.1$ at $R_A = 1$ for time $T$. Curves are indexed by observation intervals $T$.  

Figure 16 Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for parabolic image; $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_o = 1$ for time $T$. Curves are indexed by parameter $M$.  

Figure 17 Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for parabolic image; $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_o = 1$ for time $T$. Curves are indexed by observation intervals $T$.  

Figure 18 Error vs. mean counts $\mu$ for parabolic image; $Q_0 = 10^{-2}$. Curves are indexed by parameters $I_S$ defined in (3.47) and $M$.  

vii
Figure 19  Average error probability $P_e$ in binary bits
detection vs. distance of separation
$(2 \times_0/\sigma)$ between images for ideal detector,
counting comparator, or simple detectors in the
absence of background light; $N_0 = 0.1, 1, 2,$
5 . . . . . . . . . . . . 148

Figure 20  Average error probability $P_e$ in binary bits
detection vs. distance $(2 \times_0/\sigma)$ for ideal
detection and simple detector with fixed
area or optimum area in the absence of the
background light; $N_0 = 5$ . . . . . . . . . . . . 156

Figure 21  Average error probability $P_e$ in binary
detection vs. average number $N_0$ for simple
detector, counting comparator with square
observation area $A' = 4\sigma^2$ at $R_0 = 1$ and
$x_0/\sigma = 2, N_0 = 0, 0.1, 0.5, 1$. . . . . . . . . . . . 158

Figure 22  Average error probability $P_e$ in binary
detection vs. average number $N_0$ for simple
detector and counting comparator with square
observation area $A' = 4\sigma^2$ at $R_0 = 1$ and
$N_0 = 0.1; x_0/\sigma = 1, 2$. . . . . . . . . . . . 160

Figure 23  Average error probability $P_e$ in binary
detection vs. average number difference $\Delta N$ for
simple detector and counting comparator
where $\Delta N = \bar{n}_{T1} - \bar{n}_{T0}$ with $\bar{n}_{Tk}$ as the mean number
of photoelectrons observed from a finite area of
one-half of the surface under hypothesis $H_k$ for
$k=0, 1; \bar{n}_{T0} = 0.5, 1, 2$. . . . . . . . . . . . 162
Figure 24  Error probability $P_e$ of the counting comparator vs. the length $R_0$ for the rectangular observation area; $N = 5$. The curves are indexed by the parameter $N_0$ defined in (2.57).

Figure 25  Average error probability $P_e$ in binary bits detection vs. the ratio $|\epsilon|/x_0$ due to the mislocation of the images for the ideal detector or the counting comparator at $x_0/\sigma = 1$; $N = 2, 5, 8$.

Figure 26  Average error probability $P_e$ in binary bits detection vs. the ratio $|\epsilon|/x_0$ due to the mislocation of the images for the ideal detector and the counting comparator for fixed value $N = 5$; $x_0/\sigma = 1.5, 2, 4$.

Figure 27  Saddle point $s_0$ vs. the signal-to-noise ratio $D^2$ for a Gaussian image; $N_0 = 0.5, 5$, $Q_0 = 10^{-3}, 10^{-5}$.
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Optimum radius for simple detector</td>
<td>94</td>
</tr>
<tr>
<td>Table 2</td>
<td>Optimum distance for simple detector</td>
<td>179</td>
</tr>
<tr>
<td>Table 3</td>
<td>Percentage error of exponential distribution</td>
<td>196</td>
</tr>
<tr>
<td>Table 4</td>
<td>Percentage error of Poisson distribution</td>
<td>196</td>
</tr>
</tbody>
</table>
Acknowledgements

The author wishes to express her appreciation to Professor Carl W. Helstrom who introduced her to this field of study and provided the direction and encouragement throughout its development. Thanks are due to her many professors who provided the insight into the theories and techniques that make this dissertation possible. Appreciation is also extended to both Dottie Bybee and Sue Pettijohn who typed the manuscript.

The author wishes to acknowledge the financial support provided by the National Aeronautics and Space Administration for the first three years and the National Science Foundation for the fourth year since 1969. This research was carried out under NASA Grant NGL 05-009-079 and NSF Grant GK-33811.

Finally, the author's deep appreciation is felt for her husband, Charles, for his encouragement and great consideration during these years of graduate study.
VITA

1962 - M.S.E.E., Oklahoma State University, Stillwater
1963 - 1965 Engineer, Fairchild Control Corporation, Los Angeles, California
1965 - 1967 Project Engineer, Endevco Corporation, Pasadena, California
1967 - 1969 Technical Staff Member, Bellcomm Inc., Washington, D.C.

PUBLICATIONS


FIELDS OF STUDY

Major Field: Signal Processing
Detection and Estimation Theory
Communication Theory
Electronic Engineering
ABSTRACT OF THE DISSERTATION

Detection and Estimation of an Optical Image by Photon-Counting Techniques

by

Lily Lee Wang

Doctor of Philosophy in Information and Computer Science
University of California, San Diego, 1973
Professor Carl W. Helstrom, Chairman

The statistical description of the photoelectric detector is given. The photosensitive surface of the detector is divided into many small areas, and the moment generating function of the photo-counting statistic is derived for large time-bandwidth product. The detection of a specified optical image in the presence of the background light by using the hypothesis test is discussed. The ideal detector based on the likelihood ratio from a set of numbers n of
photoelectrons ejected from many small areas of the photosensitive surface is studied and compared with the threshold detector and a simple detector which is based on the likelihood ratio by counting the total number of photoelectrons from a finite area of the surface. The intensity of the image is assumed to be Gaussian distributed spatially against the uniformly distributed background light. The numerical approximation by the method of steepest descent is used, and the calculations of the reliabilities for the detectors are carried out by the digital computer. An almost optimum intensity estimate is proposed when the intensity of the object light is unknown, and its statistical performance is studied. The application of the photon-counting techniques is further discussed where detectors are investigated for resolving two point sources with Gaussian images.
Introduction

An optical communication system, owing to its high information-carrying capacity both in the temporal and the spatial channels, is more advantageous than electromagnetic wave systems of lower frequency. In general, the retrieval of the information will rely heavily on prior knowledge about the field and the performance of the detecting system. Usually the optical field at space point \( \mathbf{r} \) and time \( t \) is not a measurable quantity, and only its average power at that point can be observed. A photoelectric detector, such as the photo-tube, the photomultiplier, the image tube, or the image orthicon, which carries the information about the optical field in the sense that the probability of the photoelectron emission is proportional to the intensity of the light at its photosensitive surface, plays an important role in the optical system for the detection and extraction of the information. For example, in observational astronomy it is often a practice to use an image tube or a photomultiplier to detect a star for the purpose of observation and navigation; in a binary communication system a phototube may be used to detect a bit "1" when a specified optical object is received.

In most applications, the background light, referred to as the "noise", may also pass through the optical system and thus corrupt the object light, referred to as the "signal". When the signal-to-noise ratio is low, the observer will not only be required to detect the optical signal, but also to seek the best strategy he can to make the decision whether there is only the noise or whether the object signal is also present. One way to observe the optical field by using a
photoelectric detector is to count the number of ejected photoelectrons. That is, as the photosensitive surface of the detector is divided like a mosaic into many small equal areas, the set of data can be obtained from observing the numbers of the photoelectrons ejected from each such area during the time interval \((0, T)\). Because of the stochastic nature of the optical field, each of the observed data will be a random variable. To make decisions, detection theory and the concepts of statistical hypothesis testing can be applied when the object signal is specified. If there are also some parameters of the object signal unknown, the estimation of such parameters as well as the detection must be carried out. In this thesis we will discuss the basic binary detection and single parameter estimation of an optical image and some of their applications in optical communication systems by using photon-counting techniques.

A point source, at a known location, emits incoherent quasimonochromatic light of given spectral density and is focused on the photosensitive surface of a photoelectric detector. Its electromagnetic field \(^1\), which is spectrally pure, is assumed to be a stationary, spatio-temporal, circular Gaussian random process. A brief description of the image formation and statistical description of the photoelectric detector will be given in Chapter I. As the observation interval \(T\) is much greater than the reciprocal of the bandwidth \(W^{-1}\), the distribution of the number of photoelectrons ejected from a single small area of the photosensitive surface, discussed by Mandel \(^2\) and Helstrom \(^3\), will be derived in this chapter, following Siegert \(^4\), and it can be approximated by a Poisson distribution
function. The derivation will be then extended to the approximation that the numbers of photoelectrons ejected from the different small areas of the surface will be Poisson distributed and statistically independent from each other. The moment-generating function (m.g.f.) of the statistic, which is the sum of the number of photoelectrons ejected from many small areas as these areas become infinitesimally small, will also be discussed.

When the optical signal is specified, the simple binary hypothesis tests can be used. Two important strategies, Bayes and Neyman-Pearson, will be briefly described in Chapter II. The optimum statistic based on the likelihood ratio and its m.g.f. for the spatially varying optical signal, discussed previously by Helstrom (5), will also be given. The performance of the ideal detector, in the absence of the background light, will be discussed, and its detection probability will be calculated for a preassigned false-alarm probability. When background light also passes through the system, only an approximate form of the distribution can be used. Because of the complexity of the m.g.f. of the statistic, a Gaussian approximation was suggested by Helstrom (5). When the signal-to-noise ratio becomes large, the Kth order cumulant of the statistic increases as K increases, and the Gaussian approximation is not valid. Farrell (6) recommended the gamma distribution approximation. We have used the method of steepest descent to approximate the distribution with a uniform asymptotic expansion series following Rice (7), as also discussed by Daniel (8). The performance of the optimum detector has been
investigated with a Gaussian image and uniform background light, where
the detection probability at preassigned values of false-alarm
probability, and the average error probability will be calculated by
using the digital computer. As the signal-to-noise ratio is small,
the optimum detector can be approximated by the threshold detector and
will be also discussed. Besides the ideal detector, which registers
the locations of the ejected photoelectrons from the surface, a
simple detector, which observes the total photoelectrons ejected from
the surface without dividing it into many small areas, will be also
discussed. The performances of both the threshold detector and the
simple detector will be investigated with the Gaussian image and
uniform background light. Comparison between these three detectors
will be made with the evaluation of the average error probabilities
and of the detection probabilities at preassigned values of the
false-alarm probabilities.

When the intensity of the point source, which is located in
the field of view, is unknown, the observer must also estimate the
intensity of the image simultaneously if he makes the decision that
the optical signal is present. Since the observed data are a
set of random variables, no two experiments will yield the same value
of the intensity estimate even though the true value of the intensity
is the same in both. The most one can hope for is that the estimate
will be close to the true value of the intensity in the sense of "on
the average". Two most important strategies, Bayes and
maximum-likelihood, will be discussed in Chapter III. The intensity
estimate, derived from the Bayes strategy, which has been discussed
also by Middleton and Esposito\(^{(9)}\), will be investigated with a quadratic cost function where the prior probability density function (p.d.f.) of the intensity will be assumed to be a gamma distribution function. When the signal-to-noise ratio is not too weak, the Bayes estimate of intensity will be approximately equivalent to the maximum-likelihood estimate. An almost optimum estimate therefore will be proposed, and its statistical performance will be investigated with a truncated Gaussian or a parabolic image. The expectations of the estimate will be calculated by the digital computer at different values of parameters such as the duration of the observation interval and the radius of observation area, all at a preassigned value of the false-alarm probability.

In Chapter IV, we will further discuss some applications to optical communication of the photon-counting techniques. For example, the ideal detector derived from the likelihood ratio given in Chapter II can be applied to resolve two point sources with equal radiant power at known locations from a single point source with twice the power located between them. The m.g.f. of the ideal detector and its cumulants will be given. Since the m.g.f. of the ideal detector bears a complicated form with double integration, we will not practice the numerical calculation though it can be carried out by the digital computer. Instead, the ideal detector for detecting a bit "1" when light from a point source located at \(x\) is received or a bit "0" when a light from a point source located at \(-x\) is received will be investigated numerically in the absence of any background light. A Gaussian image will be postulated, and the calculation of the error
probability, which is dependent on the separation distance of the point sources, will be carried out. Two more detectors, the simple detector that counts the total number of photoelectrons ejected from half of the surface of the photoelectric detector and the counting comparator that compares the numbers of photoelectrons observed from the upper and lower half-surfaces of the photoelectric detector, will be also discussed, and their error probabilities will be evaluated. As the background light also passes through the aperture, the average error probabilities of both the simple detector and the counting comparator will be calculated with a finite square observation area. Since the performances of the detector for detecting binary bits will also depend on the prior knowledge about the locations of the images, the effects of mislocation on the ideal detector and counting comparator will be discussed.
Footnotes

Introduction

1. Mandel and Wolf [1].
3. Helstrom [3].
4. Siegert [4].
5. Helstrom [5].
6. Farrell [6].
7. Rice [7].
8. Daniel [8].
9. Middleton and Esposito [9].
In this chapter, the general terms such as intensity, covariance function, and point-spread function in image forming and processing will be discussed. The optical field \( v \) will be considered as a complex circular Gaussian stochastic process which is stationary, ergodic and spectrally pure. In other words, the process can be described by the complex random functions of space \( \xi \) and time \( t \), \( V_1(\xi,t) \) and \( V_2(\xi,t) \), where \( V_1(\xi,t) \) and \( V_2(\xi,t) \) are the real and imaginary parts of the process. The statistical description of the photoelectric detector will be given. When the observation time \( T \) is much greater than the reciprocal of the bandwidth \( W^{-1} \), the numbers of the photoelectrons ejected from different small areas of the photosensitive surface will be proved to be Poisson distributed and statistically independent of each other. The moment generating function (m.g.f.) of the photon-counting statistic will be derived in a general form and discussed.
The diffraction theory of image formation has been largely used to describe imaging systems in which the process is considered both linear and spatially stationary. The basic elements for an imaging system are the luminous object \( o \), aperture \( A \) (for example, a lens) and image \( I \) located in the planes \( U \), \( A \), \( R \) respectively as shown in Fig. 1, where \( z_1 \) and \( z_2 \) are the distances of the object plane \( U \) and the image plane \( R \) from the aperture \( A \).

The optical field \( v \) at the point \((u, z_1)\) of the object plane at time \( t \) is assumed to be linearly polarized and quasimonochromatic and can be characterized by

\[
v(u, z_1, t) = \Re \, V(u, z_1, t) \, e^{i \Omega_0 t}, \tag{1.1}
\]
where

\[ V(u,z_1,t) = V_1(u,z_1,t) + iV_2(u,z_1,t) \quad (1.2) \]

is the complex amplitude of the electric field and \( V_1(u,z_1,t) \), \( V_2(u,z_1,t) \) are its real and imaginary parts, which are Gaussian random processes. \( \Omega_0 \) is the angular frequency at the center of the object spectrum. "Re" indicates the real part of the analytic signal \( V(u,z_1,t) e^{i\Omega_0 t} \). As the distances \( z_1 \) and \( z_2 \) are much greater than the wavelength of the light, the complex amplitude \( V(x,z_2,t) \) at point \((x,z_2)\) of the image plane \( R \) at time \( t \) due to the object light can be expressed as

\[ V(x,z_2,t) = \int_{-\infty}^{\infty} V(u,z_1,t) K(u,x) \, du \quad (1.3) \]

where \( K(u,x) \) is the amplitude point-spread function describing the propagation from the object plane to the image plane. \( o \) is the area of the object.

The instantaneous intensity at any point \((x,z)\) of an arbitrary plane \( X \) away from the aperture with a distance \( z \) and at time \( t \) is defined by

\[ I(x,z,t) = V(x,z,t) V^*(x,z,t) \quad (1.4) \]

where \( V^*(x,z,t) \) is the complex conjugate of the amplitude \( V(x,z,t) \).
which is not directly measurable; the observable quantity is the ensemble average intensity defined by

\[ \langle I(x,z,t) \rangle = \langle V(x,z,t) V^*(x,z,t) \rangle. \] (1.5)

When the process is also ergodic, the ensemble average is equivalent to the time average through the interval \((-T, T)\) as \(T \to \infty\).

The field is normalized so that the instantaneous power density at point \((u_1, z_1)\) of the object plane at time \(t\) is \(\frac{1}{2} |V(u_1, z_1, t)|^2\), and the complex covariance function between the complex field at point \((u_1, z_1)\) and the complex-conjugate field at point \((u_j, z_1)\) of the object plane at times \(t_1\) and \(t_j\) respectively is defined by

\[ \Gamma(u_1, u_j, t_1, t_j : z_1) = \frac{1}{2} \langle V(u_1, z_1, t_1) V^*(u_j, z_1, t_j) \rangle. \] (1.6)

(1.6) is known as the complex autocovariance function; it is proportional to the mutual coherence function with a factor \(e^{i\Delta_0(t_1 - t_j)}\). \(\Gamma(u_1, u_j, t_1, t_j : z_1)\) is the average power density at point \((u_1, z_1)\) of the object plane along the direction \(z\). For an optical system that can be approximated by only considering the paraxial rays and has small bandwidth \(W\) such that \(W \ll \Omega_0\), the property of the spectrally pure light is characterized by Mandel in such a way that the superposition of light beams will not affect the spectral distribution; the covariance function from (1.6) is then reducible to the
product of two simpler functions as

$$\Gamma(u_1, u_j, t_1, t_j; z_1) = \phi(u_1, u_j, z_1) \chi(t_1 - t_j), \quad (1.7)$$

where $\phi(u_1, u_j, z_1)$ and $\chi(t_1 - t_j)$ are the complex spatial and temporal coherence functions. In general $\chi(\tau)$ is normalized so that $\chi(0) = 1$ and $\chi(-\tau) = \chi^*(\tau)$ for real power spectrum ($\tau$ is real). The ensemble average intensity from (1.5) can also be written as

$$\langle I(x, z, t) \rangle = 2\Gamma(x, x, t, t; z), \quad (1.8)$$

which is the variance of the process $V(x, z, t)$ or the sum of the variances of the two independent processes $V_1(x, z, t)$ and $V_2(x, z, t)$ at point $(x, z)$ of the plane $X$ and at time $t$.

As the complex amplitude $V(u, z_1, t)$ is defined to be zero outside the object $o$, (1.3) can also be expressed by a convolution equation for the paraxial approximation, that is

$$V(x, z_2, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2u \; V(u, z_1, t) K(x - u). \quad (1.9)$$

The average intensity at point $(x, z_2)$ of the image plane due to the object light according to (1.5) is

$$\langle I_{im}(x, z_2, t) \rangle = \langle V(x, z_2, t) V^*(x, z_2, t) \rangle$$
\[
2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2 u_i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2 u_j \\
\Gamma(u_i, u_j, t, t; z_1) K(r-u_i) K^*(r-u_j)
\]

(1.10)

when (1.6) and (1.9) are used. \( \Gamma(u_i, u_j, t, t; z_1) \) is the covariance function between points \((u_i, z_1)\) and \((u_j, z_1)\) of the object plane at time \(t\). As the light is assumed to be spectrally pure, the average energy received within the small area \(\Delta A\) of the image plane because of the object illumination during the observation interval \((0, T)\) is given by

\[
AE = \int_{\Delta A} d^2 \mathbf{r} \int_0^T dt \Gamma(r, \mathbf{r}, t, t; z_2) \\
= T \int_{\Delta A} d^2 \mathbf{r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2 u_i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2 u_j \\
\phi(u_i, u_j, z_1) K(r-u_i) K^*(r-u_j)
\]

(1.11)

as (1.6), (1.7) and (1.9) are used.
Intensity representation of light from an incoherent object

The average intensity at point \((r, z_2)\) of the image plane and at time \(t\) due to the coherent object illumination can be expressed by (1.10) or can also be written as

\[
\langle I_{\text{im}}(r,z_2,t) \rangle = \langle |\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(r-u_1) V(u_1,z_1,t) d^2u_1 |^2 \rangle. \tag{1.12}
\]

For an incoherent object, the spatial coherence function from (1.7) can be defined as

\[
\phi(u_1,u_j,z_1) = \pi k_o^{-2} B(u_1) \delta(u_1-u_j), \tag{1.13}
\]

where \(B(u_1)\) is the radiance so defined that \(B(u_1)/4\pi\) is the power emitted per unit area per steradian in the direction \(z\) normal to the object plane. The radiance \(B(u_1)\) is of limited extent. Here \(k_o = \omega_0/C_o\) with \(C_o\) the velocity of the light, and \(\delta(u)\) is the two-dimensional Dirac delta function. The average intensity given by (1.10) can be expressed in terms of the radiance \(B(u_1)\) as

\[
\langle I_{\text{im}}(r,z_2,t) \rangle = 2\pi k_o^{-2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(r-u_1) d^2u_1 \cdot B(u_1) \delta(u_1-u_j) K^*(r-u_j) \cdot B(u_1) d^2u_1
\]

\[
= 2\pi k_o^{-2} \int_{-\infty}^{\infty} B(u_1) S(r-u_1) d^2u_1, \tag{1.14}
\]
where \( S(r-u_1) = |K(r-u_1)|^2 \) is defined as the incoherent point-spread function. Thus we can see from (1.12) that for a coherent object illumination the imaging system is linear in complex amplitude. However, for an incoherent object the imaging system is linear only in intensity as given by (1.14).
The statistical description of a photoelectric detector

When the photoelectric detector is used for image detection, the photosensitive surface—for example, the surface of a photomultiplier—will be placed at the image plane. The output of the illuminated photodetector carries information about the radiation field in the sense that the instantaneous probability of the photoelectron emission is proportional to the classical instantaneous intensity $|\mathcal{V}(x, z, t)|^2$ of the light as long as the intensity of the light is not so strong that the photoelectron emission cannot be described by the first-order perturbation theory. The output of the detector is a sequence $\{t_i\}$ of time-instants of absorption-emission phenomena. These instants $\{t_i\}$ are random and constitute a point process. If we observe this process for a fixed time interval $[t, t+T]$, a total number $n$ of photoelectrons will be ejected from the area $A$ of the photosensitive surface. The sequence $t_1, t_2, \ldots, t_i$ of instants when the photoelectrons are emitted can be written symbolically as $\{t_i\}$. When the total energy received from that area $A$ during the interval $[t, t+T]$ is given, the probability that $n$ photoelectrons are ejected is a Poisson function:

$$P(n | \lambda_T(t)) = (\lambda_T(t))^n \exp(-\lambda_T(t)) / n!$$

where

16
\[ \lambda_T(t) = \alpha \int \int_{t}^{t+T} |V(\xi, t_1)|^2 \, dt_1 \, d^2 \xi \quad (1.16) \]

is the average number of photoelectrons and can be predicted if the coefficient \( \alpha \) is given. \( \alpha \) is proportional directly to the quantum efficiency and inversely to the energy \( \hbar \omega/2\pi \) of each quantum of light, where \( \hbar \) is the Planck's constant and \( \omega \) is the angular frequency of the field. \( V(\xi, t) \) is the complex amplitude of the light field at the point \( \xi \) of the photosensitive surface at time \( t \). Since \( V(\xi, t) \) is a stochastic process, so is the number \( \lambda_T(t) \). The probability for \( n \) photoelectrons in any counting interval \([t, t+T]\) can then be expressed only by the ensemble average \(^{(10)}\) of the conditional probability \( p(n|\lambda_T(t)) \). That is,

\[ P_n(T, t) = \langle (\lambda_T(t))^n \exp(-\lambda_T(t)) / n! \rangle \quad (1.17) \]

The derivation \(^{(11,12)}\) of (1.15) is well known and will not be given here. The operation of the expectation on the probability from (1.17) makes it depart from the Poisson distribution. For a stationary process, the ensemble average is independent of the initial time \( t \). We will use the expression \( P_n(T) \) to replace \( P_n(T, t) \) in (1.17) from now on.
Distribution of the photoelectrons from a single small area when the time-bandwidth product $TW >> 1$

When linearly polarized light impinges normally on the photosensitive surface, the probability that a number $n_i$ photoelectrons are emitted from a small area $dA_i$ centered at point $r_i$ during an interval $(0,T)$ can be written according to (1.17) as

$$P_{n_i}(T) = \langle \lambda_i^{n_i} \exp(-\lambda_i) / n_i \rangle$$

(1.18)

where

$$\lambda_i = \alpha \int_0^T \int_{dA_i} |V(r,t)|^2 \, dt \, d^2r$$

(1.19)

with $f_i(r)$ as the function defined by

$$f_i(r) = 1 \quad \text{for} \quad r \in dA_i$$

$$= 0 \quad \text{for} \quad r \notin dA_i$$

(1.20)

$A$ is the entire area of the surface. $\lambda_i$ depends on the coefficient $\alpha$, the size $dA_i$ and the amplitude $|V(r,t)| \in dA_i$. The distribution of the random variable $n_i$ can also be derived from its moment generating function (m.g.f.), which is the ensemble average of the conditional m.g.f. with given $\lambda_i$. That is,
\[ h_1(s) = \langle \mathbb{E}[\exp(sn_1) \mid \lambda_1] \rangle \]
\[ = \langle \sum_{n_1=0}^{\infty} \exp(sn_1) \lambda_1^{n_1} e^{-\lambda_1 / n_1} \rangle \]
\[ = \langle \exp[\lambda_1 (e^s - 1)] \rangle \] \hspace{1cm} (1.21)

Now

\[ \lambda_1 (e^s - 1) = \alpha \int_0^T \int_A f \left( x \right) (e^s - 1) \ |V(x, t)|^2 \ dt \ d^2x \]
\[ = \int_0^T \int_0^T \int_0^T \int_A K_1(x_1, x_2; t_1, t_2) V(x_1, t_1) V^*(x_2, t_2) \ dt_1 dt_2 d^2x_1 d^2x_2, \] \hspace{1cm} (1.22)

where the kernel

\[ K_1(x_1, x_2; t_1, t_2) = F_1(s; x_1) \delta(x_1 - x_2) \delta(t_1 - t_2) \] \hspace{1cm} (1.23)

and

\[ F_1(s; x_1) = \alpha (e^s - 1) f_1(x_1) \] \hspace{1cm} (1.24)

The m.g.f. from (1.21) can be worked out as in the problem of finding the m.g.f. of a quadratic functional of a circular Gaussian random process \(^{(4,13)}\). One can show that \( h_1(s) \) can be determined from the expression
\[ h_1(s) = \exp \left[ \int_0^1 du \int_0^T \int_A L(\xi, t, \xi, t; -u) \, dt \, d^2 \xi \right], \quad (1.25) \]

where \( L(x_1, t_1, x_2, t_2; u) \) can be solved \(^{(4,14)}\) from the following integral equation

\[
L(x_1, t_1, x_2, t_2; u) + 2u \int_0^T \int_0^T \int_A d^2 y_1 \, d^2 y_2 \, dh_1 \, dh_2 \\
\cdot K_1(x_1, y_1; t_1, h_1) \, \Gamma(y_1, y_2, h_1, h_2) \, L(y_2, h_2, x_2, t_2; u)
\]

\[ = 2 \int_0^T \int_A K_1(x_1, y_1; t_1, h_1) \, \Gamma(y_1, y_2, h_1, t_2) \, d^2 y_1 \, dh_1 \]  

(1.26)

For spectrally pure light, the autocovariance function can be written according to (1.7) as

\[
\Gamma(x_1, x_2; t_1, t_2) = \phi(x_1, x_2) \, \chi(t_1 - t_2). \quad (1.27)
\]

Thus, substituting (1.23), (1.24) and (1.27) into (1.26) and solving for \( L(x_1, t_1, x_2, t_2; u) \) recursively we have

\[
L(x_1, t_1, x_2, t_2; u) = 2F_1(s; x_1) \, \phi(x_1, x_2) \, \chi(t_1 - t_2) \\
- u^2 \int_0^T \int_A d^2 y_2 \, dh_2 \, F_1(s; x_1) \, \phi(x_1, y_2) \, \chi(t_1 - h_2) \\
F_1(s; y_2) \, \phi(y_2, x_2) \, \chi(h_2 - t_2)
\]
\[ + u^2 \int_0^T \int_0^T \int_A d\gamma_2 d\gamma_3 dh_2 dh_3 F_1(s; X_1) \phi(X_1, Y_2) \]
\[ \cdot X(t_1 - h_2) F_1(s; Y_2) \phi(Y_2, Y_3) \chi(h_2 - h_3) F_1(s; Y_3) \phi(Y_3, Y_2) \chi(h_3 - t_2) \]
\[ + \ldots \ldots \quad (1.28) \]

By substituting (1.28) into (1.25) we can write the logarithm of the m.g.f. as

\[ \ln h_1(s) = 2\chi(0) T \int_A F_1(s; x) \phi(x, x) \, d^2x \]
\[ + \frac{1}{2} (2T)^2 \int_0^T \int_0^T F_1(s; x) F_1(s; y_1) \phi(x, y_1) \phi(y_1, y_2) \]
\[ \cdot d^2x \, d^2y_1 \, d^2y_2 \, T^{-3} \int_0^T \int_0^T \int_0^T \chi(t - h_1) \chi(h_1 - h_2) \chi(h_2 - t) \, dt \, dh_1 \, dh_2 \]
\[ + \ldots \frac{1}{12} (2T)^3 \int_0^T \int_0^T F_1(s; x) F_1(s; y_1) \ldots F_1(s; y_{k-1}) \phi(x, y_1) \phi(y_1, y_2) \]
\[ \ldots \phi(y_{k-1}, x) \, \frac{d^2x \, d^2y_1 \ldots d^2y_{k-1}}{(2T)^{T-k}} \int_0^T \int_0^T \chi(t - h_1) \ldots \chi(h_{k-1} - t) \, dt \, dh_1 \]
\[ \ldots \frac{d \, h_{k-1} + \ldots}{} \]

\[ = 2 \chi(0) T \int_A F_1(s; x) \phi(x, x) \, d^2x \]
\[
\sum_{k \geq 2} \frac{1}{k} (2T)^k \int \cdots \int F_1(s ; x) F_1(s ; y_1) \cdots F_1(s ; y_{k-1}) \phi(y_1 y_2) \cdots \phi(y_{k-1} y_k) d^2 x d^2 y_1 \cdots d^2 y_{k-1} \cdot P_k ,
\]

where

\[
P_k = T^{-k} \int_0^T \int_0^T \cdots \int_0^T x(h_1 - h_2) x(h_2 - h_3) \cdots x(h_k - h_1) dh_1 dh_2 \cdots dh_k ,
\]

\( k \geq 2 \) \hspace{1cm} (1.30)

Now the bandwidth \( W \) is defined \(15\) by

\[
W = |x(0)|^2 \left[ \int_{-\infty}^{\infty} |x(t)|^2 dt \right]^{-1} = \left[ \int_{-\infty}^{\infty} |\chi(\Omega)|^2 d\Omega / 2\pi \right]^{-1} ,
\]

as \( \chi(0) = 1 \) \hspace{1cm} (1.31)

where

\[
\chi(\Omega) = \int_{-\infty}^{\infty} x(t) e^{-i\Omega t} dt
\]

(1.32)

is the temporal spectral density of the complex field at the image plane and is real so that \( \chi(\tau) = \chi^*(-\tau) \). We now let

\[
\begin{align*}
h_1 - h_2 &= t_1 , \\
h_2 - h_3 &= t_2 , \\
\cdots & \cdots \cdots
\end{align*}
\]
\[ h_{k-1} - h_k = t_{k-1}, \]
\[ h_k - h_1 = -(t_1 + t_2 \ldots + t_{k-1}) . \] \hspace{1cm} (1.33)

When \( T >> \frac{1}{W} \) we can extend the range of integration to \((-\infty, \infty)\) and by changing variables we can write (1.30) as

\[ P_k = T^{-k+1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \chi(t_1) \chi(t_2) \ldots \chi(-t_1 - t_2 \ldots t_{k-1}) dt_1 dt_2 \ldots dt_{k-1}. \] \hspace{1cm} (1.34)

According to the convolution theorem (16) we have

\[ \int_{-\infty}^{\infty} \chi(t_1) \chi(-t_1 - t_2 \ldots - t_{k-1}) dt_1 \]
\[ = \int_{-\infty}^{\infty} x^2(\Omega) \exp[-i\Omega(t_2 + t_3 + \ldots + t_{k-1})] d\Omega / 2\pi . \] \hspace{1cm} (1.35)

By substituting (1.31), (1.32) and (1.35) into (1.34) we have

\[ P_k = T^{-k+1} \int_{-\infty}^{\infty} x^2(\Omega) \frac{d\Omega}{2\pi} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \chi(t_2) \ldots \chi(t_{k-1}) \]
\[ \cdot \exp[-i\Omega(t_2 + t_3 \ldots t_{k-1})] dt_2 \ldots dt_{k-1} \]
\[ = T^{-k+1} \int_{-\infty}^{\infty} x^2(\Omega) \frac{d\Omega}{2\pi} \]
\[ = (TW)^{-k+1} R_1 . \] \hspace{1cm} (1.36)
where

\[ R_I = \int_{-\infty}^{\infty} x(\Omega) \frac{d\Omega}{2\pi} \left/ \int_{-\infty}^{\infty} x^2(\Omega) \frac{d\Omega}{2\pi} \right. \] \quad (1.37)

As \( WT >> 1 \), we can approximately write \( x(\Omega) \sim \frac{1}{W} \) for \( \Omega \in (-\pi W, \pi W) \) and \( x(\Omega) = 0 \) otherwise. Thus \( R_I \sim l \) is finite and \( P_2 \rightarrow 0 \) for \( l \geq 2 \). For spatially slowly varying light and small area \( dA_1 \) we define

\[ \bar{n}_I = 2aT \int_{dA_1} \phi(x_i, x) \frac{d^2x}{dA_1} \]

\[ \propto 2aT \phi(x_i, x) \frac{dA_1}{dA_i} \quad (1.38) \]

as the average number of photoelectrons ejected from the small area \( dA_1 \) centered as \( x_i \). Since for the spatial coherence function

\[ |\phi(x_i, x_j)| \leq \phi(x_i, x_1) \quad \text{for all } i \neq j, \]

\[ |\int_A \int_A \cdots F_i(s; x) \phi(x_i, x_1) F_i(s; x_1) \phi(x_1, x_2) \cdots \]

\[ F_i(s; x_{l-1}) \phi(x_{l-1}; x) \frac{d^2x}{dA_1} \frac{d^2x_1}{dA_1} \cdots \frac{d^2x_{l-1}}{dA_1} | \]

\[ \leq |\int_A F_i(s; x) \phi(x, x)|^l. \quad (1.39) \]

Thus as long as \( \bar{n}_I \) defined by (1.38) is not too large, all the terms
in (1.29) for \( k \geq 2 \) are negligible by comparing with the first term as \( W^T \gg 1 \). That is

\[
\ln h_i(s) \approx \bar{n}_i (e^s - 1),
\]

(1.40)

which is the logarithm of the m.g.f. of a Poisson distribution \(^{(17)}\) with mean \( \bar{n}_i \).
Distribution of the photoelectrons from two small areas for TW $\gg 1$

When the energy received by each of the two small areas $dA_1$ and $dA_2$ centered at $x_1$ and $x_2$ is given, the numbers of photoelectrons $n_1$ and $n_2$ ejected from these areas during the time interval $[t, t+T]$ are Poisson distributed with the given mean values $\lambda_i$ for $i = 1, 2$ defined in (1.19) and are statistically independent. The joint m.g.f. of the random variables $n_1$ and $n_2$ can be derived from the expected value of the conditional m.g.f.:

$$h_{12}(s_1, s_2) = \langle e^{s_1 n_1 + s_2 n_2} | \lambda_1, i=1,2 \rangle$$

$$= \langle \exp[\lambda_1 (e - 1) + \lambda_2 (e - 1)] \rangle, \quad (1.41)$$

where now

$$\lambda_1 (e - 1) + \lambda_2 (e - 1)$$

$$= \int \int_A dt \int d^2x [f_1(x)(e - 1) + f_2(x)(e - 1)] |V(x, t)|^2$$

$$= \int \int_A \int \int_0^T dt_1 dt_2 d^2x_1 d^2x_2 K_{12}(x_1, x_2, t_1, t_2) V(x_1, t_1) V^*(x_2, t_2), \quad (1.42)$$

where $f_i(x)$ for $i = 1$ or $2$ is given by (1.20) and
\[ K_{12}(x_1, x_2; t_1, t_2) = F_{12}(s_1, s_2; x_1) \delta(x_1 - s_1) \delta(t_1 - t_2), \]  

(1.43)

with

\[ F_{12}(s_1, s_2; x_1) = \alpha f_1(x_1) e^{s_1-1} + \alpha f_2(x_1) e^{s_2-1}. \]  

(1.44)

The m.g.f. can also be expressed in the form of (1.25) and the following integral equation must be satisfied.

\[ L_{12}(x_1, t_1, x_2, t_2; u) + 2u \int_A \int_0^T F_{12}(s_1, s_2; x_1) \phi(x_1, t_1, x_2, h_2) \]

\[ L_{12}(x_2, h_2, x_2, t_2; u) d^2 x_2 d h_2 \]

\[ = F_{12}(s_1, s_2; x_1) \phi(x_1, t_1, x_2, t_2). \]  

(1.45)

The kernel \( L_{12}(x_1, t_1, x_2, t_2; u) \) is solved for recursively and substituted into the joint m.g.f. as described before where \( K_{l}(x_1, x_2; t_1, t_2) \) and \( F_{l}(s; x_1) \) in (1.23) and (1.24) are now replaced by \( K_{12}(x_1, x_2; t_1, t_2) \) and \( F_{12}(s_1, s_2; x_1) \). We have

\[ \ln h_{12}(s_1, s_2) = 2 \chi(0) T \int_A F_{12}(s_1, s_2; x) \phi(x, x) d^2 x \]

\[ + \frac{1}{2} (2T)^2 \int_A \int_A d^2 x d^2 y_1 F_{12}(s_1, s_2; x) F_{12}(s_1, s_2; y_1) \]

\[ \phi(x, y_1) \phi(y_1, x) P_2 \]
\[ + \frac{1}{3} (2T)^3 \int_\mathcal{A} \int_\mathcal{A} \int_\mathcal{A} d^2x d^2x_1 d^2x_2 \int_{s_1, s_2, x} P_{12}(s_1, s_2, x) P_{12}(s_1, s_2, y_1) \]

\[
\cdot P_{12}(s_1, s_2, x) \phi(x, y_1) \phi(x_1, y_2) \phi(y_2, x) \text{ P}_3 \\
+ \ldots + \frac{1}{4} (2T)^2 \int_\mathcal{A} \int_\mathcal{A} d^2x d^2x_1 \int_{s_1, s_2, x, y_{\ell-1}} \int_{s_1, s_2, x} d^2x_\ell \phi(x, y_1) \phi(x_1, y_2) \phi(y_2, x) \phi(y_{\ell-1}, x) \text{ P}_\ell \\
+ \ldots \\
(1.46) \]

where \( P_\ell \) for \( \ell \geq 2 \) is given by (1.30). By the same argument, that is, when the average numbers \( n_i \) defined in (1.38) for \( i = 1, 2 \) are not too large and \( WT \gg 1 \), (1.46) can be approximated by the first term or

\[ \ln h_{12}(s_1, s_2) \approx 2T \int_\mathcal{A} \int_{s_1, s_2, x} P_{12}(s_1, s_2, x) \phi(x, x) d^2x \]

\[ \approx 2aT d_{A_1} \phi(x_1, x_{\ell-1})(e^{-1}) + 2aT d_{A_2} \phi(x_2, x_\ell)(e^{-1}) \]

\[ = -n_1(e^{-1}) + n_2(e^{-1}) \] (1.47)

Since \( \ln h_i(s_i) = -n_i(e^{-1}) \) for \( i = 1, 2 \) as given by (1.40), this implies that the joint m.g.f. \( h_{12}(s_1, s_2) \) is the product of the m.g.f.'s \( h_1(s_1) \) and \( h_2(s_2) \) as \( TW \gg 1 \). In other words, the number of
photoelectrons ejected from areas $dA_1$ and $dA_2$ during any observation interval are Poisson distributed and statistically independent of each other when the time-bandwidth product is large.
Distribution of the sum of the photoelectrons from a large number of small areas when $WT >> 1$

When there are a large number $N$ of small areas $dA_i$ for $i = 1, 2 \ldots N$, the distribution of the sum of the photoelectrons from these areas can be derived in the same way just described as for $N = 2$. For generalization, we discuss the distribution of a statistic given by

$$ Y = \sum_{i=1}^{N} n_i \beta(x_i) ,$$

(1.48)

where $x_i$ is the center of each small area $dA_i$ where $n_i$ photoelectrons are ejected. $\beta(x_i)$ is a weighting function to characterize the location of $n_i$. For example, when $\beta(x_i) = 1$, $Y$ is the sum of the photoelectrons ejected from the $N$ small areas. As $N$ is very large we will then pass to the limit as $dA_i \to 0$.

The m.g.f. of $Y$ can be written as

$$ h_Y(s) = \langle \exp \left[ s \sum_{i=1}^{N} n_i \beta(x_i) \right] \rangle $$

$$ = \langle \exp \left[ \sum_{i=1}^{N} \lambda_i (e^{s\beta(x_i)} - 1) \right] \rangle ,$$

(1.49)

where
\[
\sum_{i=1}^{N} \lambda_i (e^{\beta(x_i)} - 1)
\]
\[
= \sum_{i=1}^{N} (e^{\beta(x_i)} - 1) \int_0^T \int_{\mathcal{A}_i} |V(x,t)|^2 \, dt \, d^2x
\]
\[
= \sum_{i=1}^{N} (e^{\beta(x_i)} - 1) \int_0^T \int_{\mathcal{A}_i} f_1(x) |V(x,t)|^2 \, dt \, d^2x
\]
\[
= \int_0^T \int_0^T \int_{\mathcal{A}_i} \int_{\mathcal{A}_i} K_N(x_i,x_2,t_1,t_2) V(x_1,t_1) V^*(x_2,t_2)
\]
\[
\cdot dt_1 dt_2 d^2x_1 d^2x_2
\]

(1.50)

Now
\[
K_N(x_1,x_2,t_1,t_2) = F_N(s \omega x_1) \delta(x_1-x_2) \delta(t_1-t_2),
\]

(1.51)

with
\[
F_N(s \omega x_1) = \sum_{i=1}^{N} (e^{\beta(x_i)} - 1) f_1(x_i),
\]

(1.52)

and \(f_1(x_i)\) is defined in (1.20).

We must first solve the integral equation
\[
L_0(x_1,t_1,x_2,t_2,u) + 2u \int_0^T \int_{\mathcal{A}_i} F_N(s \omega x_1) \phi(x_1,x_2) \chi(t_2-t_1)
\]

(1.53)
$$L_0(Y_2, h_2, x_2, t_2; u) \ d^2h_2 \ d^2x_2$$

$$= 2 \ F_N(s, x_1) \phi(x_1, x_2) \chi(t_1-t_2). \quad (1.53)$$

Now

$$\int_A F_N(s, x) \phi(x, y) \ d^2x$$

$$= \alpha \sum_{i=1}^{N} \left( e^{s \beta(x_i)} - 1 \right) \int \phi(x, y) \ d^2x$$

$$\simeq \alpha \sum_{i=1}^{N} \left( e^{s \beta(x_i)} - 1 \right) \phi(x_i, y) \ dA_i, \quad (1.54)$$

as $\phi(x, y)$ is slowly varying spatially. When $N$ is very large we can pass the limit as $dA_i \to 0$. \ (1.54) can be written as

$$\int_A F_N(s, x) \phi(x, y) \ d^2x$$

$$\simeq \alpha \int_A \left( e^{s \beta(x)} - 1 \right) \phi(x, y) \ d^2x. \quad (1.55)$$

The logarithm of the m.g.f. can be obtained when the kernel $L_0(x_1, t_1, x_2, t_1; u)$ is solved first recursively from (1.53) where (1.54) is substituted. We have

$$L_0(x_1, t_1, x_2, t_2; u) \simeq 2 \ F_N(s, x_1) \phi(x_1, x_2) \chi(t_1-t_2)$$
\[-2^2 \Phi_{N}(s, x_1) \int_0^T \int_A \, dh_2 \, d^2 y_2 \]

\[= \beta(\gamma_1) \phi(x_1, x_2) \phi(x_2, x_3) \chi(t_1 - h_2) \chi(h_2 - t_2) \]

\[+ 2^3 (\Phi_{N}(s, x_1) \int_0^T \int_T \int_A \int_A \, dh_2 \, dh_3 \, d^2 y_2 \, d^2 y_3 \]

\[= \beta(\gamma_2) \beta(\gamma_3) \phi(x_1, y_2) \phi(y_2, y_3) \phi(y_3, x_2) \chi(t_1 - h_2) \chi(h_2 - h_3) \chi(h_3 - t_2) \]

\[+ \ldots \ldots \quad (1.56) \]

The logarithm of the m.g.f. can be derived as we pass to the limit

\[dA_1 \to 0 \text{ where (1.55) and (1.56) are used.} \]

\[= n \, h_{Y(s)} = \int_0^1 du \int_0^T \int_A \int_0^T \int_A \int_0^T \int_A \, L_0(x, t, t; u) \, dt \, d^2 x \]

\[= 2\alpha x(0) T \int_A \phi(x, x) \, d^2 x \]

\[+ \frac{1}{2} (2\alpha T)^2 \int_A \int_A \, d^2 y_1 \, d^2 y_2 \]

\[= \beta(\gamma_1) \beta(\gamma_2) \phi(x_1, y_2) \phi(y_2, y_1) \phi(y_1, y_2) \phi(y_2, y_1) \cdot P_2 + \ldots \ldots \]

\[+ \frac{1}{2} (2\alpha T)^2 \int_A \int_A \ldots \int_A \, d^2 y_1 \, d^2 y_2 \ldots \, d^2 y_2 \]
\[
\left[ \prod_{i=1}^{\ell} \left( e^{\beta(x_i^2)} - 1 \right) \right] \phi(x_1, x_2) \ldots \phi(x_{\ell}, x_1) P_{\ell} + \ldots \ldots
\]  

(1.57)

where \( P_{\ell} \) for \( \ell \geq 2 \) is given by (1.30). As \( \ell T >> 1 \), \( P_{\ell} \to 0 \) for all \( \ell \geq 2 \), and by the same argument used in (1.39), all the terms in (1.57) except the first term become very small. Then we can approximate

\[
\ln h_{\lambda}(s) = 2aT \int_A (e^{\beta(x^2)} - 1) \phi(x, x) d^2 x,
\]

(1.58)

or the m.g.f. of the statistic \( Y \) can be written approximately

\[
h_{\lambda}(s) = \exp \left[ 2aT \int_A (e^{\beta(x^2)} - 1) \phi(x, x) d^2 x \right],
\]

(1.59)

where the mean and variance in general are

\[
E[Y] = 2aT \int_A \beta(x) \phi(x, x) d^2 x,
\]

\[
\text{Var}[Y] = 2aT \int_A \beta^2(x) \phi(x, x) d^2 x.
\]

(1.60)

For \( \beta(x) = 1 \), (1.59) becomes

\[
h_{\lambda}(s) = \exp \left[ 2aT \int_A \phi(x, x) d^2 x (e^s - 1) \right],
\]

(1.61)
which is the m.g.f. of a Poisson distribution with equal cumulants

\[ 2\alpha T \int_A \phi(x_i, x) \, d^2x \] for all orders.

On the other hand, as we expand

\[ (e^{s \beta(x)} - 1) = \sum_{m=1}^{\infty} \frac{(s \beta(x))^m}{m!}, \quad (1.62) \]

if

\[ \frac{1}{m!} \int_A \beta^m(x) \phi(x_i, x) \, d^2x << \frac{1}{2} \int_A \beta^2(x) \phi(x_i, x) \, d^2x \]

\[ \text{for } m \geq 3, \quad (1.63) \]

we can further approximate (1.59) as

\[ h_Y(s) = \exp \left[ 2\alpha T \int_A s \beta(x) \phi(x_i, x) \, d^2x \right. \]

\[ + \left. 2\alpha T \int_A s^2 \beta^2(x) \phi(x_i, x) \, d^2x \right], \quad (1.64) \]

which is the m.g.f. of a Gaussian distribution.
Footnotes

Chapter 1

4. Helstrom [10], Chapter XI, p. 375.
5. Wolf [12].
6. Mandel and Wolf [1].
7. Mandel [13].
8. Helstrom [14], pp. 165-166.
9. Helstrom [3], eq. (2).
10. Mandel [15], eq. (69).
11. Mandel [15], Appendix B.
12. Bar-David [16].
13. Siegert [4].
14. Helstrom [17], Appendix.
15. Helstrom [18], eq. (1.6).
16. Papoulis [19], p. 27.
17. Mandel [2].
Chapter II  Binary detection of an optical signal at a photosensitive surface

Two important criteria, Bayes and Neyman–Pearson, used in the simple hypothesis test for binary detection will be briefly described. The optical signal is detected by using the photoelectric detector. The photosensitive surface of the device is divided like a mosaic into many small equal areas so that the observed data \( n \) are the set of numbers of photoelectrons ejected from all the small areas. The optimum statistic \(^{(1)}\) obtained from the likelihood ratio in the presence of the uniform background light will be discussed and its moment-generating function (m.g.f.) will be derived. Because of the complexity of the m.g.f., the distribution of the optimum statistic will be approximated by the method of steepest descent \(^{(2)}\) and compared with two other approximate forms, the Gaussian and gamma distributions. The detection of the optical signal in the absence of the background light will also be discussed. Two other detectors, the threshold and the simple detector, will be studied. In order to investigate the performances of these detectors, a Gaussian image will be postulated and the false-alarm probability, the detection probability, and the average error probability will be calculated.
Binary hypothesis test

In a communication system when a set of data \( b = \{b_1, b_2 \ldots b_N\} \) is observed, the decision between binary hypotheses \( H_0 \) and \( H_1 \), where \( H_0 \) is the hypothesis of background noise alone and \( H_1 \) is the hypothesis that a signal is also present, can be viewed as a rule for dividing the total N-dimensional observation space \( \Sigma \) of the data \( b \) into two regions, \( \Sigma_0 \) and \( \Sigma_1 \). Hypothesis \( H_0 \) will be chosen if the observed data \( b \) fall in the region \( \Sigma_0 \); \( H_1 \) will be chosen otherwise.

The joint conditional probability density function (p.d.f.) \( P(b|H_k) \) under \( H_k \) for \( k = 0 \) or \( 1 \) is to be given, and the likelihood ratio defined by

\[
\Lambda(b) = \frac{P(b|H_1)}{P(b|H_0)} \tag{2.1}
\]

will be calculated. If the Bayes criterion is used, a decision level defined by

\[
\Lambda_0 = \frac{\xi[C_{10} - C_{00}]}{(1-\xi)[C_{01} - C_{11}]} \tag{2.2}
\]

can be evaluated, where the prior probability \( \xi \) of hypothesis \( H_0 \) and the costs \( C_{ij} \) of choosing hypothesis \( H_i \) when \( H_j \) happens to be true \((i,j = 0,1)\), are given. The regions \( \Sigma_0 \) and \( \Sigma_1 \) are divided in such a way that \( \Sigma_0 \) contains all the data for \( \Lambda(b) \leq \Lambda_0 \) and \( \Sigma_1 \) for \( \Lambda(b) > \Lambda_0 \). The false-alarm probability and the detection probability are defined as
When there is no information about the prior probability \( \xi \) and the cost functions, the appropriate way to make the decision according to the Neyman-Pearson criterion is that the false-alarm probability \( Q_0 \) will not exceed the preassigned value. The decision level \( \Lambda_0 \) must be solved from the expression of \( Q_0 \) by (2.3), and the hypothesis \( H_1 \) or \( H_0 \) will be selected according as \( \Lambda(b) > \Lambda_0 \) or \( \Lambda(b) \leq \Lambda_0 \).

If \( b \) is a set of discrete random variables, the likelihood ratio \( \Lambda(b) \) defined in (2.1) will be the ratio of the probabilities \( P(b|H_1) \) and \( P(b|H_0) \). When the Neyman-Pearson criterion is used, randomization must be applied. That is, a probability \( f \) of choosing \( H_1 \) must be assigned at \( \Lambda(b) = \Lambda_0 \) so that the false-alarm probability

\[
Q_0 = \Pr[\Lambda(b) > \Lambda_0 | H_0] + f \Pr[\Lambda(b) = \Lambda_0 | H_0] = \sum_{\Lambda(b) > \Lambda_0} P(b|H_0) + f P(b: \Lambda(b) = \Lambda_0 | H_0)
\]  

(2.4)
equals exactly the preassigned value. For given \( Q_0 \), both \( \Lambda_0 \) and \( f \) can be solved from (2.4) and will be discussed later for Poisson distribution. Hypothesis \( H_1 \) or \( H_0 \) will be chosen for \( \Lambda(b) > \Lambda_0 \) or \( \Lambda(b) \leq \Lambda_0 \).

The detection probability for the correct decision can be evaluated in the same way after the decision level \( \Lambda_0 \) and probability \( f \) are obtained. That is,
\[
Q_d = \sum_{\Lambda(b) > \Lambda_0} P(b|H_1) + fP(h|\Lambda(b) = \Lambda_0|H_1).
\] (2.5)

The average error probability for each decision is defined by

\[
P_e = \xi Q_0 + (1-\xi)(1-Q_d).
\] (2.6)

If the relative costs \(C_{10} - C_{00}\) and \(C_{01} - C_{11}\) given in (2.2) are equal, the hypothesis \(H_k\) with greater posterior probability \(P(H_k|b)\) for \(k = 0\) or 1 is always selected when the Bayes criterion is used and the error probability \(P_e\) given by (2.6) is minimized \(3\).

The reliability and error probability are used to measure the performance of the detectors. One detector is said to be better than the other detector if for a fixed value \(Q_0\), the former has higher detection probability and hence smaller error probability.

The operation of the hypothesis test for binary detection can best be described by the block diagram on the next page. The selection of the Bayes or the Neyman-Pearson criterion depends on the information about the system we have and its application. The block labeled "computer" implies that certain numerical calculations will be involved so that the decision level \(\Lambda_0\), the reliability \(\{Q_0, Q_d\}\) or the error probability \(P_e\) can be calculated.
Figure 2  Block diagram of hypothesis test for binary detection.
The ideal photoelectric detector

When a luminous object is located in the field of view of an optical detector and focused on the photosensitive surface, the photoelectrons emitted from the surface will be proportional to the intensity of the object light, which is a fluctuating function of time as discussed in Chapter I. If there is also background light incident through the aperture, the information carried by the object light will be corrupted. As the intensity of the object light is not much stronger than the background light, in order to decide whether a certain luminous object is present or not, we can use strategies according to the hypothesis tests which we have just discussed. The basic elements of the optical detector have been shown in Fig. 1, where the photoelectrical emission surface of the photo-tube is placed in the image plane and is divided like a mosaic with a large number $N$ of small equal areas $dA_i$ as shown in Fig. 3.

![Photoelectrical emission surface](image)

Figure 3    Photoelectrical emission surface
The received data are the values \( \mathbf{n} = \{n_1, n_2, \ldots, n_N\} \) obtained by measuring the numbers \( n_i \) of photoelectrons ejected from each area \( dA_i = dA \) for \( i = 1, 2, \ldots, N \) during the interval \((0, T)\). Since the background light in general also enters the aperture, the observer must choose between the two hypotheses: \((H_0)\) only background light is present, or \((H_1)\) object light is also incident through the aperture in addition to the background light. When the observation time \( T \) is much greater than \( W^{-1} \), the numbers \( n_i \) of photoelectrons emitted from each area \( dA_i \) under hypothesis \( H_k \) \((k = 0, 1)\) are Poisson distributed and statistically independent of each other as discussed in Chapter I.

The probability is given by

\[
P_{n_i}(T | H_k) = \frac{n_i^{n_i} e^{-n_{ik}}}{n_i!}, \quad (2.7)
\]

where \( \overline{n_{ik}} \) is the average number of photoelectrons emitted during the interval \((0, T)\) for \( k = 0, 1 \) from the \( i \)th area \( dA_i \) and

\[
\overline{n_{ik}} = \alpha T dA I_k(x_i), \quad k = 0, 1 \quad (2.8)
\]

with

\[
I_k(x_i) = 2 \phi_k(x_i, x_j), \quad k = 0, 1 \quad (2.9)
\]

when (1.38) is used. \( \phi_k(x_i, x_j) \) is the spatial coherence function under hypothesis \( H_k \) for \( k = 0, 1 \) at point \( x_i \) and point \( x_j \). \( I_k(x_i) \) is the
mean intensity at point \( x_1 \). The joint conditional probability for observing the data \( n \) under hypothesis \( H_k \) for \( k = 0 \) or \( 1 \) is

\[
P_{\mathcal{R}}(T|H_k) = \prod_{i=1}^{N} p_{n_i}(T|H_k)
\]

\[
= \prod_{i=1}^{N} \frac{n_i}{n_{ik}} e^{-\frac{n_{ik}}{n_i}} \cdot n_{ik} \cdot n_{i1} \cdot e^{-\frac{n_{i1}}{n_{i0}}} .
\] (2.10)

The likelihood ratio defined by (2.1) becomes

\[
\Lambda(n) = \frac{p_{\mathcal{R}}(T|H_1)}{p_{\mathcal{R}}(T|H_0)}
\]

\[
= \prod_{i=1}^{N} \left( \frac{n_{il}}{n_{i0}} \right)^{n_i} e^{-\frac{n_{il}}{n_{i0}} + \frac{n_{i0}}{n_{i0}}} \cdot n_{il} + n_{i0} .
\] (2.11)

and will be compared with the decision level \( \Lambda_0 \) according to the Bayes strategy or Neyman-Pearson strategy. Hypothesis \( H_1 \) will be chosen if \( \Lambda(n) > \Lambda_0 \), and \( H_0 \) will be chosen otherwise. It is also equivalent to use the optimum statistic

\[
g = \sum_{i=1}^{N} n_i \ln \left( \frac{n_{il}}{n_{i0}} \right) .
\] (2.12)

since \( \Lambda(n) \) is a monotone function of \( g \) which is obtained from the logarithm of the likelihood ratio \( \Lambda(n) \). The decision level of this statistic is given by
\[ e_0 = \ln A_0 + \Delta N, \quad (2.13) \]

where

\[ \Delta N = \sum_{i=1}^{N} (\bar{n}_{11} - \bar{n}_{10}) \quad (2.14) \]

is the total average number of photoelectrons emitted over the entire area of the surface when the signal field alone is present. Thus we can compare \( g \) with \( e_0 \) instead and select \( H_1 \) if \( g > e_0 \) and \( H_0 \) when \( g \leq e_0 \). When Neyman-Pearson criterion is used, the false-alarm probability \( Q_0 \) is preassigned. The distribution of the statistic \( g \) under \( H_0 \) must be known so that the decision level \( e_0 \) can be determined from either of the expressions in (2.3) or (2.4) as discussed previously. To examine the reliability or the error probability of making each decision we must also know the distribution of the statistic \( g \).

The mean intensity \( I_k(x_i) \) given in (2.9) can be written as

\[ I_k(x_i) = I_0(x_i) + k I_s(x_i), \quad k = 0, 1 \quad (2.15) \]

where \( I_0(x_i) \) is the mean intensity due to the background light. In most cases, such as thermal background light, \( I_0(x_i) \) can be assumed to be spatially invariant for all areas \( dA_i \) over the surface. \( I_s(x_i) \) is the mean intensity due to the object light in the area \( dA_i \). Therefore
we can write

$$I_0(x_1) = I_b, \text{ for all } dA_1$$

$$I_s(x_1) = I_s \gamma(x_1), \quad i = 1, 2 \ldots N.$$  \hspace{1cm} (2.16)

where $I_s = \int_A I_s(x) \, d^2x$ is the total power transmitted to the surface from the object and $\gamma(x_1)$ is the weighting function to characterize the distribution of the power at point $x_1$. $A$ is the entire area which is divided into $N$ equal areas $dA$.

We also define a function

$$H(x_1) = \frac{n_{11}}{n_{10}}$$

$$= \frac{[I_0(x_1) + I_s(x_1)]}{I_0(x_1)}$$

$$= 1 + I_s \gamma(x_1)/I_b$$

$$= 1 + D^2 u(x_1),$$  \hspace{1cm} (2.17)

where $D^2$ is called the signal-to-noise and is the ratio of the total average number $\alpha T I_s$ of photoelectrons ejected from the surface by the object light illumination to the average number $\alpha T I_b A_0$ of photoelectrons ejected from a finite area $A_0$ of the surface by the background light illumination. The function $u(x_1)$ is only different from $\gamma(x_1)$ by the
factor $A_0$, that is $u(x_i) = A_0 \gamma(x_i)$. $A_0$ is arbitrarily defined. The statistic $g$ from (2.12) can be written as

$$g = \sum_{i=1}^{N} n_i \ln H(x_i), \quad (2.18)$$

which is similar to the statistic $Y$ of (1.48) with $\beta(x_i) = \ln H(x_i)$ as discussed in Chapter I. As $N$ is very large, by passing to the limit as $dA \to 0$, the m.g.f. of the statistic $g$ under hypothesis $H_k$ as $TW >> 1$ can be written according to (1.59) as

$$h_k(s) = E\left[ e^{-sg} \mid H_k \right]$$

$$= \exp \left[ aT \int_{A} \left( e^{-s \ln H(x)} - 1 \right) I_k(x) \, d^2x \right]$$

$$= \exp \left[ aTI_b \int_{A} H^k(x) \left( H^{-s}(x) - 1 \right) d^2x \right],$$

$$k = 0, 1 \quad \text{(2.19)}$$

The logarithm of $h_k(s)$ is

$$\ln h_k(s) = aTI_b \int_{A} H^k(x) \left( H^{-s}(x) - 1 \right) d^2x,$$

for $k = 0, 1 \quad \text{(2.20)}$

The distribution of the statistic $g$ cannot in general be obtained in
exact form. Only an approximation can be used and will be discussed later.

The logarithm of \( h_k(s) \) from (2.20) is differentiable for any order with respect to the variable \( s \), and the \( n \)th cumulant under hypothesis \( H_k \) for \( k = 0, 1 \) can be expressed as

\[
G_k^n = (-1)^n \frac{d^n}{ds^n} \ln h_k(s) \bigg|_{s=0}
\]

\[
= \alpha T \int_a^b \frac{h_k(x)}{H(x)} \ln H(x) \, d^2x .
\] (2.21)

In particular, the mean and variance of the statistic \( g \) can be calculated by

\[
E[g|H_k] = \alpha T \int_a^b \frac{h_k(x)}{H(x)} \ln H(x) \, d^2x
\]

\[
\text{Var}[g|H_k] = \alpha T \int_a^b \frac{h_k(x)}{H(x)} \ln^2 H(x) \, d^2x ,
\]

\[
k = 0, 1 .
\] (2.22)

If the average intensity \( I_s(x_1) \) given in (2.16) is zero outside a finite area \( A_1 \) of the photosensitive surface, that is, if

\[
I_1(x_1) = I_s(x_1) + I_0(x_1) \quad \text{for} \quad x_1 \in A_1 \quad \text{and} \quad I_1(x_1) = I_0(x_1) \quad \text{otherwise},
\]

then \( H(x_1) = 1 \) when \( x_1 \notin A_1 \) and the optimum statistic \( g \) involves only the emission within the area \( A_1 \), so that there will be a finite probability of getting no electrons under hypothesis \( H_k \).
\[
Fr \left[ g=0 \left| H_k \right. \right] = \exp \left[ -\alpha T \int_{A_1} I_k(x) \, d^2 x \right],
\]

\[k = 0, 1\]

which is the result of (2.19) as \( s \to \infty \).

When \( I_k(x) \) is also spatially invariant, so is the function \( H(x) \). The optimum statistic is therefore equivalent to \( g/2n \, C \), where \( C = H(x) \) is constant for all \( i \), and the statistic can be written as

\[
g' = \sum_{i=1}^{N} n_i
\]

which is the sum of the photoelectrons from the total number \( N \) of small areas \( dA \). As \( TW \gg 1 \), the m.g.f. of the statistic \( g' \) is given by (1.59) with \( \beta(x) = 1 \) as discussed in Chapter I. \( g' \) is therefore a random variable with a Poisson distribution.

The detector just described represents a particular way of processing the light field behind the aperture of the optical system in order to choose between the hypotheses. It has been compared with the optimum means of processing the field from the point of view of quantum detection theory\(^{(5)}\). When diffraction is negligible and the surface has unit quantum efficiency, this detector is equivalent to the optimum detector of the image-forming light\(^{(6)}\).
Threshold detector

The structure of the optimum statistic is based on the function $H(x)$ which is a function of $D^2 u(x_1)$ as given by (2.17). When the signal-to-noise ratio $D^2$ is not known, the detector must be designed for some reasonable standard value. If $D^2$ is very small, the appropriate thing to do in detection theory is to expand the logarithm of the conditional likelihood ratio $H[R|D^2]$ from (2.11) into a power series in $D^2$, and the optimum statistic is equivalent to the threshold statistic, which is based on the lowest order of $D^2$ in that expansion (7). For detecting the image on the photoelectric surface we have just discussed, the optimum statistic $g$ from (2.18) can be expanded into a power series in $D^2$ or

$$g = \sum_{i=1}^{N} n_i \sum_{m=1}^{\infty} (-1)^{m-1} \frac{[D^2 u(x_1)]^m}{m!}.$$  \hspace{1cm} (2.25)

Since $D^2$ is very small and is independent of the location, we can base the decision on the threshold statistic, which is the coefficient of the lowest order of $D^2$ in (2.25). We have

$$g_0 = \sum_{i=1}^{N} n_i u(x_1).$$  \hspace{1cm} (2.26)

The m.g.f. of this statistic can be derived for $TW >> 1$ and is given by (1.59) where $\beta(x_1) = u(x_1)$. That is when $N$ is very large. By passing to the limit as $dA \to 0$ we have

50
\[
\begin{align*}
\eta_{0k}(s) &= E \left[ e^{-s \theta} \mid \eta_k \right] \\
&= \exp \left\{ \alpha T \int_A H^k(\chi) \left[ e^{-s u(\chi)} - 1 \right] d^2 \chi \right\}, \\
&\quad k = 0, 1 , \quad (2.27)
\end{align*}
\]

where \( H(\chi) \) and \( u(\chi) \) are given in (2.17), and the logarithm of (2.27) is

\[
\ln \eta_{0k}(s) = \alpha T \int_A H^k(\chi) \left[ e^{-s u(\chi)} - 1 \right] d^2 \chi ,
\]

for \( k = 0, 1 \) . \quad (2.28)

The nth cumulant of the threshold statistic \( g_\theta \) under \( \eta_k \)
can be written as

\[
G_{\theta k}^n = (-1)^n \left. \frac{d^n}{ds^n} \ln \eta_{0k}(s) \right|_{s=0}
\]

\[
= \alpha T \int_A H^k(\chi) u^n(\chi) d^2 \chi
\]

for \( k = 0, 1 \) , \quad (2.29)

and the mean and variance are

\[
E[\theta \mid \eta_k] = \alpha T \int_A H^k(\chi) u(\chi) d^2 \chi
\]
\[ \text{Var}[g_\theta | H_k] = aT I_b \int_A H^k(\chi) u^2(\chi) \, d^2\chi. \] (2.30)

Since the m.g.f. from (2.27) is dependent on the signal-to-noise ratio \( D^2 \) only through the function \( H(\chi) \), the distribution of the threshold statistic \( g_\theta \) will not depend on \( D^2 \) when \( H_0 \) is true.
**Simple detector**

Instead of weighting the number \( n_1 \) of photoelectrons from each area \( dA_1 \) with a factor \( \ln H(x_1) \), a simple way is to collect all the photoelectrons ejected from a certain area \( A_1 \) of the surface without any weighting factor about the location. As \( TW >> 1 \), the probability of \( n \) photoelectrons ejected from the area \( A_1 \), under hypothesis \( H_k \) has been proved in Chapter I to be Poisson distributed and

\[
P_n(T|H_k) = \frac{\overline{n}_k^n e^{-\overline{n}_k}}{n!},
\]

where

\[
\overline{n}_k = aT \int_{A_1} I_k(x) d^2x, \quad k = 0 \text{ or } 1,
\]

is the average number of photoelectrons ejected from the area \( A_1 \), under \( H_k \), and \( I_k(x) \) is given by (2.9). The likelihood ratio of the simple detector according to (2.1) is

\[
\Lambda_s(n) = \frac{P_n(T|H_1)}{P_n(T|H_0)} = \left( \frac{\overline{n}_{T1}}{\overline{n}_{T0}} \right)^n e^{-\overline{n}_{T1} + \overline{n}_{T0}}.
\]

Hypothesis \( H_0 \) will be chosen if \( \Lambda_s(n) \leq \Lambda_{s0} \) and \( H_1 \) will be chosen otherwise, where \( \Lambda_{s0} \) is the decision level, which can be obtained according to the criterion we use. Because the likelihood ratio \( \Lambda_s(n) \)
of the simple detector is a monotone function of $n$, it is therefore equivalent to use the optimum statistic from the logarithm of $A(n)$

$$g_s = n$$ \hspace{1cm} (2.34)

and hypothesis $H_1$ will be chosen if $g_s > g_{s0}$ and $H_0$ otherwise except that at $g_s = g_{s0}$ a probability $f$ will be assigned for choosing $H_1$. When the Bayes criterion is used the decision level $g_{s0}$ of the statistic can be determined from the likelihood-ratio decision level $A_{s0}$ and

$$g_{s0} = (\ln A_{s0} + \Delta n_T) \left[ \ln \left( 1 + \frac{\Delta n_T}{\overline{n}_{T0}} \right) \right]^{-1}$$ \hspace{1cm} (2.35)

where $\Delta n_T = \overline{n}_{T1} - \overline{n}_{T0}$ is the average number of photoelectrons ejected from the area $A_1$ by the object light illumination alone and can be calculated from the expression

$$\Delta n_T = \alpha T I_S \int_{A_1} \gamma(x) \, d^2x$$ \hspace{1cm} (2.36)

The optimum area $A_1$ for the observer to adopt is the area where the detection probability is maximum for a fixed false-alarm probability or the error probability is minimum.
Reliability and error probability of the ideal detector in the absence of the background light

The performance of a detector can be characterized by the reliability, $Q_0$ and $Q_d$, and the error probability $P_e$ as defined by (2.3) and (2.6). When there is no background light, there will be no photoelectrons ejected from the photosensitive surface under the hypothesis $H_0$. For both Bayes and Neyman-Pearson criteria the strategy is to choose $H_1$ whenever one or more than one photoelectron have been observed. As $TW >> 1$, the joint probability for observing the data $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ from the N small areas under hypothesis $H_1$ is given by (2.10). Since the probability of zero photoelectron emission under $H_0$ is 1, the likelihood ratio at $n = 0$ is

$$
A(n=0) = e^{-N_A} \quad (2.37)
$$

where

$$
N_A = \alpha_T I_S \quad (2.38)
$$

is the total average number of photoelectrons ejected from the photosensitive surface due to the object light. For the detection of the image when Bayes criterion is used, hypothesis $H_1$ will be chosen at $n = 0$ if $e^{-N_A} > \Lambda_0$ and $H_0$ otherwise, with $\Lambda_0$ given by (2.2). If the relative costs $C_{10} - C_{00}$ and $C_{01} - C_{11}$ are equal, the hypothesis with greater posterior probability $P(H_k | n=0)$ for $k = 0, 1$ is always
selected where

\[ P(H_k \mid n=0) = P(H_k) \frac{P(n=0 \mid H_k)}{P(n=0)}, \]

\( k = 0 \) or \( 1 \). \hspace{1cm} (2.39)

The minimum average error probability will be

\[ p_e = \min(\xi, (1-\xi) e^{-N^s}). \]

(2.40)

In particular, when the hypotheses are equally likely

\[ p_e = \frac{1}{2} e^{-N^s}, \quad \xi = \frac{1}{2}. \]

(2.41)

For the Neyman-Pearson criterion, the randomized strategy must be used. As the false-alarm probability \( Q_0 \) is preassigned, hypothesis \( H_1 \) will be chosen with a probability \( Q_0 \) when no photoelectron is emitted from the surface during the interval \((0,T)\). The detection probability at this case becomes

\[ Q_d = 1 - (1-Q_0) e^{-N^s}. \]

(2.42)

Thus the error probability \( p_e \) can be calculated again according to (2.6).
Reliability and error probability of the ideal detector in the presence of the background light

When the background light is present and \( I_0(x) = I_b \), the optimum statistic \( g \) is expressed by (2.18), where its m.g.f. is given by (2.19). It is not in general possible to evaluate the distribution, which is known as an infinitely divisible distribution \((8,9)\).

Helstrom \((10)\) treated the distribution by a Gaussian approximation, where the p.d.f. of the statistic \( g \) is approximated by

\[
N_k(g) = \frac{1}{\sqrt{2\pi} \sigma_k} \exp\left[-\frac{(g-\bar{g}_k)^2}{2\sigma_k^2}\right] \tag{2.43}
\]

with

\[
\bar{g}_k = \mathbb{E}[g|H_k], \quad \text{for } k = 0, 1,
\]

\[
\sigma_k^2 = \text{Var}[g|H_k],
\]

given in (2.22).

The reliability of the detector can then be expressed by the error-function integral,

\[
Q_0 = \Pr[g > g_0 | H_0] = \text{erfc} \left( y_0 \right),
\]

\[
Q_d = \Pr[g > g_0 | H_1] = \text{erfc} \left( y_1 \right), \tag{2.44}
\]
As we have discussed in Chapter I, with the m.g.f. given by (2.19) the distribution of the statistic \( g \) can be approximated by the Gaussian function if only the condition from (1.63) is true. In other words, with \( \beta(x) = \ln \mu(x_j) \) the condition from (1.63) will be satisfied if \( D^2 << 1 \). For large values of the signal-to-noise ratio \( D^2 \), the Gaussian approximation will not be accurate because the \( m \)th order cumulant of the statistic increases as \( m \) increases. Farrell (11) recommended approximating the distribution of the statistic \( g \) by a gamma distribution function. The p.d.f. of the statistic is approximated by

\[
G_k(g) = \left[ \frac{a_k}{\Gamma(b_k)} \right] (a_k g)^{b_k-1} e^{-a_k g} \quad (2.45)
\]

with

\[
a_k = \frac{E[g|H_k]}{\text{Var}[g|H_k]} \quad , \quad k = 1, 0
\]

\[
b_k = \left( \frac{E[g|H_k]}{\text{Var}[g|H_k]} \right)^2 \quad \text{Var}[g|H_k]
\]
so that the mean and variance of the gamma random variable are matched with the mean and variance given by (2.22). Here \( \Gamma(b_k) \) is the gamma function.

The reliability of the optimum detector is then approximated by

\[
Q_0 = P[g > g_0 | H_0] = \int_{z_0}^{\infty} \beta_0(z) \, dz,
\]
\[
Q_d = P[g > g_0 | H_1] = \int_{z_1}^{\infty} \beta_1(z) \, dz, \tag{2.46}
\]

where

\[
f_k(z) = \frac{1}{\Gamma(b_k)} b_k^{-\frac{1}{2}} z^{b_k-\frac{1}{2}} e^{-z},
\]

\[
z_k = a_k g_0, \quad k = 1, 0.
\]

On the other hand, it is reasonable to use the method of steepest descent as illustrated in Appendix A, where well-known distributions, such as Gaussian, gamma and Poisson, have been used as examples and good numerical agreement has been obtained, as listed in Table 3 and Table 4.

The optimum statistic \( g \) given by (2.18) is a non-negative random variable since the function \( H(x) \) exceeds 1 for all \( x \). The m.g.f. is given by (2.19). The probability that \( g \) exceeds the decision level under \( H_k \) can be written as
\[ \Pr[g > g_0 | H_k] = \int_{g_0}^{\infty} dF_k(g), \]
\[ k = 0, 1 \quad (2.47) \]

and can be approximated by the asymptotic expansion series given by (A12) in Appendix A, where \( F_k(g) \) is the distribution function of the random variable \( g \). The complex phase of the integral defined by (A7) under hypothesis \( H_k \) for \( k = 0 \) or \( 1 \) is

\[ \phi_k(s) = g_0 \ln h_k(s) + s \]
\[ = g_0^{-1} \alpha_t I_b \int_A H(x)^{k-s} (H(x))^{-s-1} d^2 x + s, \]
\[ \text{for } k = 0, 1 \quad (2.48) \]

when (2.20) is used.

There is only one saddle point for this complex phase as discussed in Appendix B, and it can be determined from the equation

\[ \frac{d}{ds} \phi_k(s) = 0 \text{ or} \]
\[ \varepsilon_0 = \alpha_t I_b \int_A (H(x))^{k-s} \ln H(x) d^2 x \quad (2.49) \]

If we use the statistic

\[ g' = \sum_{i=1}^{N} n_i \ln \frac{n_i!}{n_{10}} + N_s \]
instead of the statistic from (2.12) we can derive the m.g.f. of \( g' \) under hypothesis \( H_k \) for \( k = 0, 1 \) as

\[
h_k'(s) = \exp\left\{ \alpha T \int B \left[ \left( H(x) \right)^{k-s} - 1 \right] d^2x + (s-k) N_s \right\},
\]

which gives us the fundamental relation for all likelihood-ratio detectors as

\[
h_1'(s) = h_0'(s-1).
\]  

(2.50)

The saddle point \( s_1 \) under \( H_1 \) can be directly obtained through the saddle point \( s_0 \) under \( H_0 \) for real values by

\[
s_1 = s_0 + 1.
\]  

(2.51)

The relation from (2.51) also satisfies (2.49) for a fixed value of \( g_0 \) for both \( H_1 \) and \( H_0 \). Whether the saddle point \( s_k \) for \( k = 0 \) or \( 1 \) is positive or negative will depend on the following conditions

\[
s_k < 0 \quad \text{if} \quad g_0 > E[g|H_k],
\]

\[
> 0 \quad \text{if} \quad g_0 < E[g|H_k],
\]

\[
= 0 \quad \text{if} \quad g_0 = E[g|H_k],
\]

\[k = 0, 1.\]  

(2.52)
The $n$th derivative of the complex phase $\phi_k(s)$ for $n \geq 2$ under hypothesis $H_k$ is

\[
\phi_k^n(s) = \frac{d^n}{ds^n} \phi_k(s)
\]

\[
= (-1)^n \epsilon_0^{-1} aT_i b \int_A (\mathbb{H}(\chi))^{k-s} \ln^n \mathbb{H}(\chi) \, d^2x ,
\]

for $k = 0, 1$. \hspace{1cm} (2.53)

From the relation (2.51) we can have

\[
\phi_0^n(s_0) = \phi_1^n(s_1) , \quad \text{for } n \geq 1 . \hspace{1cm} (2.54)
\]

The false-alarm probability and the detection probability can be approximated\(^\text{12,13}\) by using (A12) from Appendix A

\[
Q_0 = 1 - E_0(g_0) - I_0(g_0) ,
\]

\[
Q_d = 1 - E_1(g_0) - I_1(g_0) , \hspace{1cm} (2.55)
\]

where

\[
E_k(g_0) = \begin{cases} 
1 - \text{erfc} \left[ (-2 g_0 \phi_k(s_k))^\frac{1}{2} \right] , & \text{for } s_k < 0 \\
\text{erfc} \left[ (-2 g_0 \phi_k(s_k))^\frac{1}{2} \right] , & \text{for } s_k > 0
\end{cases}
\]
and

\[ I_k(g_0) = \frac{\exp[g_0 \phi_k(s_k)]}{[2\pi g_0 \phi_k(2)(s_k)]^{1/2}} \sum_{j=0}^{\infty} \left( \frac{-2}{g_0 \phi_k(2)(s_k)} \right)^j s_k^{-1} \]

\[ \sum_{n=0}^{2j} (-s_k)^{-2j+n} \sum_{l=0}^{n} A_{k,n} \left( \frac{1}{2} \right)_{l+j} \]

\[ -\text{sign}(s_k) \left( \frac{1}{2} \right)_k \left( \frac{-\phi_k(2)(s_k)}{2\phi_k(s_k)} \right)^{1/2} (g_0 \phi_k(s_k))^{-j} \]

where \( A_{k,n} \left( \frac{1}{2} \right)_{l+j} \), \( \text{sign}(x) \) are defined in (A12) and (A10) for integers \( l, n, j \). If the Bayes criterion is used, the decision level \( g_0 \) is calculated by (2.13). The saddle point can be obtained by solving (2.49), and \( \phi_k(s_k) \) and \( \phi^n_k(s_k) \) for \( k = 0, 1 \) can be evaluated from (2.48), (2.53) and (2.54) where numerical integration will be required. The probability \( Q_0 \) or \( Q_d \) will then be approximated by summing the series given in (2.55) up to the term after which the absolute values of the terms either begin to increase or become insignificantly small. If the Neyman-Pearson criterion is used, the value of \( Q_0 \) is preassigned and the decision level \( g_0 \) must be hunted by iteration. Since \( g_0 \) is a monotone function of the real value \( s_0 \), both \( g_0 \) and \( s_0 \) can be found simultaneously by iteration until the false-alarm probability calculated by the asymptotic expansion approaches the preassigned value \( Q_0 \) within a tolerable error such that (2.49) is satisfied. The saddle point \( s_1 \) under \( H_1 \) will be obtained according to (2.51), and the detection probability \( Q_d \) will
be approximated by the expansion (2.55). The numerical iteration procedure for searching for the saddle points \( s_0, s_1 \), and the decision level \( g_0 \) is described in Appendix B.

When the relative costs \( C_{10} - C_{00} \) and \( C_{01} - C_{11} \) are equal, the decision level \( g_0 \) for Bayes criterion is given by (2.13) with \( \Lambda_0 = \xi / (1 - \xi) \). The saddle points \( s_0 \) and \( s_1 \) are calculated from (2.49) and (2.51). The average error probability is minimized and can be obtained after the probabilities \( Q_0 \) and \( Q_d \) are approximated by the asymptotic expansion (2.55). The performance of the ideal detector will be studied with a Gaussian image and compared with the threshold detector and simple detector in the following paragraphs.
Performance of the ideal detector for a Gaussian image

In order to investigate the performance of the ideal detector and compare it with the other two detectors, the threshold and the simple detector, a Gaussian image will be postulated. The average intensity function on the photosensitive surface due to the object light given in (2.16) is described by the weighting function

\[ \gamma(x) = \frac{1}{2\pi\sigma^2} \exp \left[ -\frac{x_1^2 + x_2^2}{2\sigma^2} \right] \]  

(2.56)

where \( \sigma \) is the width of the image, this might be the image of a circular nebula or, more important, of a point source whose light has passed through a turbulent medium. If we define

\[ N_0 = 2\pi \sigma^2 I_b \alpha T \]  

(2.57)

as the average number of photoelectrons emitted from an area \( A_0 = 2\pi \sigma^2 \) of the photosensitive surface due to the background light and with \( N_s \) given by (2.38), the signal-to-noise ratio \( D^2 \) and function \( u(x) \) from (2.17) can be written as

\[ D^2 = \frac{N_s}{N_0}, \quad \text{for } N_0 \neq 0 \]

\[ u(x) = \exp \left[ -\frac{x_1^2 + x_2^2}{2\sigma^2} \right] \]  

(2.58)
In the absence of the background light, that is \( N_0 = 0 \), and for a false-alarm probability \( Q_0 \) preassigned at 0.2, 0.1 or 0.01, the detection probability \( Q_d \) is plotted as a function of \( N_b \) in Fig. 4; (2.42) is used to calculate \( Q_d \).

If the background light is present, the m.g.f. of the ideal statistic \( g \) at \( \text{WT} \gg 1 \) for the Gaussian image is given by

\[
h_k(s) = \exp \left\{ \frac{N_0}{2\pi \sigma^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1+D^2u(x)) \frac{1}{(1+D^2u(x))^s - 1} \, d^2x \right\},
\]

\[ k = 0, 1 \]  

(2.59)

when (2.19), (2.56), (2.57), and (2.58) are used.

Now we can change variables by letting

\[ x_1 = R \sigma \cos \theta, \quad x_2 = R \sigma \sin \theta \]

so that

\[ u(x) = \exp \left[-\frac{R^2}{2}\right] \]

The m.g.f. from (2.59) becomes

\[
h_k(s) = \exp \left\{ \frac{N_0}{2\pi \sigma^2} \int_0^{2\pi} \int_0^{\infty} (1+D^2 \exp(-R^2/2)) \frac{1}{(1+D^2 \exp(-R^2/2))^s - 1} \, \sigma^2 \, dR \, d\theta \right\}
\]
\[
= \exp \left\{ N_0 \int_0^1 (1+D^2y)^k \left[ (1+D^2y)^{-s} - 1 \right] y^{-1} dy \right\},
\]
\[k = 0 \text{ or } 1, \quad (2.60)\]

where (2.60) is obtained by changing the variable \( y = e^{-R^2/2} \). The complex phase required in the asymptotic expansion series under \( H_k \) for \( k = 0, 1 \) as given by (2.48) can be written as

\[
\phi_k(s) = g_0^{-1} N_0 \int_0^1 (1+D^2y)^k \left[ (1+D^2y)^{-s} - 1 \right] y^{-1} dy + s,
\]
\[k = 0, 1 \quad (2.61)\]

where the saddle point is obtained by solving the equation

\[
g_0 = N_0 \int_0^1 (1+D^2y)^{k-s} \ln(1+D^2y) y^{-1} dy,
\]
\[k = 0, 1 \quad (2.62)\]

and \( s_1 \) and \( s_0 \) satisfy (2.51).

The \( n \)th derivative of the complex phase \( \phi_k(s) \) for \( n \geq 2 \) under hypothesis \( H_k \) is

\[
\phi_k^n(s) = (-1)^n g_0^{-1} N_0 \int_0^1 (1+D^2y)^{k-s} \ln^n(1+D^2y) y^{-1} dy,
\]
\[k = 0 \text{ or } 1 \quad (2.63)\]
Figure 4  Detection probability $Q_d$ as a function of $N_s$ defined in (2.38) for the ideal detector in the absence of background light; false-alarm probability $Q_0 = 0.01, 0.1, 0.2$. 
The mean and variance of the optimum statistic \( g \) from (2.22) for the Gaussian image under \( H_k \) for \( k = 0, 1 \) becomes

\[
E[g | H_k] = N_0 \int_0^1 (1 + D^2 y)^k \ln(1 + D^2 y) y^{-1} dy,
\]

\[
\text{Var}[g | H_k] = N_0 \int_0^1 (1 + D^2 y)^k \ln^2(1 + D^2 y) y^{-1} dy.
\]  

(2.64)

In order to compare the three approximation methods, steepest descent, Gaussian, and gamma, the reliability of the ideal detector has been calculated by the three approximate forms as given in (2.55), (2.44), and (2.46) respectively. At typical values of \( N_0 = N_s = 5 \) \((D^2 = 1)\), the false-alarm and detection probabilities of the statistic are plotted as functions of the decision level \( g_0 \) in Fig. 5. It can be noticed that both Gaussian and gamma approximations with the mean and variance matched by (2.64) are least accurate in the tails of the distributions. For small false-alarm probability \( Q_0 \), there may be a serious error in the decision level \( g_0 \) if either one of the approximate forms is used under the Neyman-Pearson criterion where \( Q_0 \) is preassigned. To apply the asymptotic expansion series given in (2.55), an iterative search method was used to determine the decision level \( g_0 \) and the associated saddle point \( s_0 \). Detailed description of the iteration is given in Appendix B.

To investigate the performance of the optimum detector and later the threshold detector, the detection probability \( Q_d \) of the optimum detector is calculated by using the asymptotic expansion series...
from (2.55) and is plotted as a function of the signal-to-noise ratio $D^2$ for typical values of $N_0$ at 0.5 and 5 in Fig. 6 where the false-alarm probability $Q_0$ is preassigned at $10^{-3}$ and $10^{-5}$. The detection probability is also plotted as a function of the mean number $N_g$ by the solid curves in Fig. 7 at $Q_0 = 10^{-3}$ and in Fig. 8 at $Q_0 = 10^{-5}$.

The detection probability $Q_d$ at $N_0 = 0$, given by (2.42), is also plotted in Fig. 7 and Fig. 8 for the purpose of comparison. When the relative costs $C_{10} - C_{00}$ and $C_{01} - C_{11}$ are equal, the minimum error probability of the optimum detector at $\xi = \frac{1}{2}$ can be calculated by using (2.6) and the asymptotic expansion series (2.55) and is plotted as a function of the mean number $N_s$ by the solid curves in Fig. 9 where the decision level $g_0 = N_s$ is obtained according to (2.13) and the saddle point $s_0$ is searched for by iteration to satisfy (2.62).
Figure 5  Probability $Q_0$ and probability $Q_d$ as a function of decision level $g_0$ for the ideal detector calculated by the saddle point, Gaussian and gamma approximations; $N_0 = N_s = 5$, where $N_0$ is defined in (2.57).
$N_0 = N_s = 5$

- Saddle point
- Gamma
- Gaussian

PROBABILITY (%) vs. $g_0$

$Q_d$

$Q_o$
Figure 6  Probability $Q_d$ as a function of the signal-to-noise ratio $D^2$ defined in (2.58) for the ideal detector; $N_0 = 0.5, 5, Q_0 = 10^{-3}, 10^{-5}$. 
Figure 7  Probability $Q_d$ as a function of $N_s$ for ideal, threshold and simple detectors; $Q_0 = 10^{-3}$ at $N_0 = 0, 0.5, 5$. 
Graph showing the relationship between $Q_d(\%)$ and $N_s$ for different values of $N_0$ (0, 0.5, 5). The graph includes three types of detectors: Ideal, Threshold, and Simple, each represented by a different line style. The legend indicates $Q_o = 10^{-3}$.
Figure 8  Probability $Q_d$ as a function of $N_s$ for ideal, threshold and simple detectors, $Q_0 = 10^{-5}$ at $N_0 = 0, 0.5, 5$. 
$Q_d(\%)$ vs $N_s$

- $N_0 = 0$
- $N_0 = 0.5$
- $N_0 = 5$

$Q_0 = 10^{-5}$

- Ideal Detector
- Threshold
- Simple
Figure 9  Average error probability $P_e$ in detecting the Gaussian image against the uniform background light vs. the mean number $N_s$ for ideal and simple detectors; $N_0 = 0, 0.01, 0.5, 5.$
Performance of the threshold detector for a Gaussian image in the
presence of the background light

When the signal-to-noise ratio $D^2$ is small and the threshold
statistic $g_0$ from (2.26) is used, the m.g.f. given by (2.27) for a
Gaussian image becomes

$$h_{0k}(s) = \exp \left\{ aT \int \int (1+D^2u(x))^k \right. $$

$$\left[ \exp(-su(x))-1 \right] d^2x \left. \right\}, $$

$$k = 0 \text{ or } 1 . \quad (2.65)$$

Again by changing variables, we have

$$h_{0k}(s) = \exp \left\{ \frac{N_0}{2\pi\sigma^2} \int \int \left( 1+D^2 \exp(-R^2/2) \right)^k \right. $$

$$\left( \exp(-s \exp(-R^2/2))-1 \right) \sigma^2 R dR d\theta \left. \right\} $$

$$= \exp \left\{ N_0 \int \left( 1+D^2y \right)^k \left( \exp(-sy)-1 \right) y^{-1} dy \left. \right\}, $$

$$k = 0 \text{ or } 1 . \quad (2.66)$$

The mean and variance of the threshold statistic $g_0$ under $H_k$ for
$k = 0, 1$ are
\[ E[g_0 | H_k] = N_0 \int_0^1 (1 + D^2 y)^k dy = N_0 (1 + kD^2/2), \]
\[ \text{Var}[g_0 | H_k] = N_0 \int_0^1 (1 + D^2 y)^k y dy = N_0 (k + kD^2/3), \]
\[ k = 0, 1. \] (2.67)

The complex phase defined by (A7) under \( H_k \) for \( k = 0, 1 \) is

\[ \phi_0(s) = g_{60}^{-1} \ln h_0(s) + s \]
\[ = g_{60}^{-1} N_0 \int_0^1 (1 + D^2 y)^k [\exp(-sy) - 1] y^{-1} dy + s \] (2.68)

from which the saddle point \( s \) is determined by the equation

\[ \frac{d}{ds} \phi_0(s) = 0 \quad \text{or} \]
\[ g_{60} = N_0 \int_0^1 (1 + D^2 y)^k \exp(-sy) dy \]
\[ = s^{-1} N_0 \left\{ 1 - \exp(-s) + kD^2 [s^{-1} - (1 + s^{-1}) \exp(-s)] \right\} , \]
\[ k = 0, 1. \] (2.69)

Since the relation of the m.g.f.'s by (2.51) does not hold for the
the statistic \( g_0 \), the saddle points \( s \) under hypothesis \( H_k \) for both \( k = 0, 1 \) must be calculated from (2.69), which has only one root of real value as discussed in Appendix B. The \( n \)th derivatives of \( \phi_k(s) \) for \( n \geq 2 \) are given by

\[
\phi_k^n(s) = (-1)^n g_0^{-1} N_0 \int_0^1 (1 + D^2 y)^k y^{n-1} \exp(-sy) \, dy
\]

\[
= (-1)^n g_0^{-1} N_0 \left\{ f_{n-1}(s) + kD^2 f_n(s) \right\},
\]

\[ k = 0, 1 \quad (2.70) \]

where

\[
f_n(s) = \int_0^1 y^n \exp(-sy) \, dy
\]

\[
= n! s^{-n-1} - s^{-1} \exp(-s) \sum_{j=0}^{n} \left\{ n! (n-j)! s^{-j} \right\}
\]

\[ (2.71) \]

Again, with the mean number \( N_0 \) and the signal-to-noise ratio \( D^2 \) given, the reliability of the statistic can be approximated by the asymptotic expansion as given in (2.55). For each decision level \( g_0 \) the saddle points \( s_0, s_1 \) can be solved for from (2.69). The false-alarm probability and the detector probability are plotted as a function of \( g_0 \) in Fig. 10 at \( N_0 = 5 \) and \( D^2 = 0.1, 1, 2, 3, 4, \) and 6. Since the
distribution of the threshold statistic \( g_0 \) is independent of the signal strength under hypothesis \( H_0 \), the advantage of the threshold detector is that for each preassigned \( Q_0 \) the decision level \( g_{60} \) is fixed for all values of \( D^2 \) as shown by Fig. 10. The detection probability \( Q_d \) is also plotted as a function of the mean number \( N_s \) by the dashed curves in Fig. 7 and Fig. 8 for typical values of \( N_0 \) at \( Q_0 = 10^{-3} \) and \( Q_0 = 10^{-5} \). The detection probability \( Q_d \) of the threshold detector approaches the curves of the optimum detector when \( D^2 \) is very small; however, the optimum detector always has higher detection probability. To calculate the minimum error probability \( P_e \), one has to solve the decision level \( g_{60} \) such that the ratio of the probabilities for \( g_0 = g_{60} \) under \( H_1 \) and \( H_0 \) is equal to \( \xi/(1-\xi) \). We have not carried out the calculation.
Figure 10  Probability $Q_0$ and probability $Q_d$ vs. decision level $g_0$ for threshold detector; $N_0 = 0.5$, $d^2 = 0.1, 1, 2, 3, 4, 6$. 
Performance of the simple detector for a Gaussian image in the presence of the background light

For the simple detector, the optimum statistic from the likelihood ratio is the total number $n$ of the photoelectrons emitted from the optimum area $A_1'$ as given by (2.34). If the radius of the area $A_1'$ is given by $\sigma_R$, the average number $\bar{n}_{Tk}$ of the photoelectrons emitted from $A_1'$ under hypothesis $H_k$ for the Gaussian image according to (2.32) becomes

$$\bar{n}_{Tk} = a_T \int_{A_1'} \left[ I_b + kI_S \gamma(x) \right] \, d^2x$$

$$= a_T \int_0^{2\pi} \int_0^{\sigma_R^2} \left( I_b + kI_S \frac{1}{2\pi \sigma^2} \exp(-R^2/2\sigma^2) \right) \cdot R \, d\theta$$

$$= \alpha \pi \sigma^2 \sigma_R^2 \left[ I_b T + k \alpha I_S T \left( 1 - \exp(-\sigma_R^2/2) \right) \right]$$

$$= N_0 \left[ \rho + kD^2(1 - e^{-\rho}) \right],$$

where $\rho = \frac{1}{2} \sigma_R^2$ and $A_1' = \pi \sigma^2 \sigma_R^2$.

Since the statistic $g_s$ is a discrete random variable, when the Neyman-Pearson criterion is used, randomization must be applied.
That is, the reliability of the simple detector according to (2.4) and (2.5) can be written as

\[ Q_0 = \sum_{n>g_{s0}} \frac{-n_{T0}}{n!} e^{-n_{T0}} / n! + \int_{n_{T0}}^{g_{s0}} e^{-n_{T0}} / g_{s0} \]

\[ Q_d = \sum_{n>g_{s0}} \frac{-n_{T1}}{n!} e^{-n_{T1}} / n! + \int_{n_{T1}}^{g_{s0}} e^{-n_{T1}} / g_{s0} \]

(2.73)

where for preassigned \( Q_0 \) the decision level \( g_{s0} \) is the smallest integer such that

\[ \sum_{n=0}^{g_{s0}} \frac{-n_{T0}}{(n_{T0})!} e^{-n_{T0}} / n! > 1 - Q_0 \]

(2.74)

\( f \) is the probability to choose \( H_1 \) when \( n = g_{s0} \) and can be determined from

\[ f = \left[ \left( \sum_{n=0}^{g_{s0}} \frac{-n_{T0}}{(n_{T0})!} e^{-n_{T0}} / n! \right) + Q_0 - 1 \right] \]

\[ \cdot \frac{-g_{s0}}{(n_{T0})!} e^{-n_{T0}} / g_{s0} \]

(2.75)

where \( n_{T0}, n_{T1} \) are given by (2.72) for \( k = 0, 1 \). Hypothesis \( H_1 \) will be then chosen if \( n > g_{s0} \) and \( H_0 \) will be chosen if \( n < g_{s0} \). As the
false-alarm probabilities $Q_0$ are preassigned at $10^{-3}$ and $10^{-5}$, the detection probabilities $Q_d$ of the simple detector are calculated from (2.73) at $N_0 = 0.5$ and 5 and are plotted as a function of the mean value $N_e$ in Fig. 7 and Fig. 8. For each point on the curves, the radius $R_o$ was varied to yield maximum detection probability.

When the Bayes criterion is used, with $A_{s0}$ given by (2.2), the decision level from (2.35) becomes

$$g_{s0} = \left[ \ln A_{s0} + N_0 D_2 (1-e^{-\rho}) \right] \left[ \ln (1+D_2 e^{-1}) (1-e^{-\rho})\right]^{-1}. \quad (2.76)$$

Hypothesis $H_1$ will be chosen if the number $n > g_{s0}$ or $H_0$ will be chosen otherwise. The reliability of the detector can be evaluated according to (2.73), except that we let $f = 0$ because the probability for $g_{s0}$ to be an integer is negligible. When the relative costs are equal, the average error probability given by (2.5) is minimum and can be written as

$$p_e = \xi \sum_{n=g_{s0}}^{\infty} \frac{n!}{n!} \left( -nT_0 \right) e^{-nT_0} / n!$$

$$+ (1-\xi) \sum_{n=0}^{<g_{s0}} \frac{n!}{n!} \left( -nT_1 \right) e^{-nT_1} / n! \quad (2.77)$$

At $\xi = \frac{1}{2}$, $g_{s0}$ is calculated from (2.76) with $A_{s0} = 1$, and the error probabilities $p_e$ are calculated and plotted as a function of mean
number difference $\Delta N = N_8$ at $N_0 = 0, 0.01, 0.5$ and 5 by the dashed curves in Fig. 9. For each point on these curves the value $R_o$ was varied to yield maximum $P_e$.

When the false-alarm probability $Q_0$ is preassigned, the decision level $g_{s0}$ determined from (2.74) is a monotonically increasing function of $R_o$. The optimum radius $R_o$ at which the detection probability $Q_d$ yields a maximum can easily be searched by the digital computer. The values of the optimum radius $R_o$ at different values of $N_8$ are listed in Table 1 at typical values of $N_0 = 0.5, 5$. The maxima are quite flat and the radius $R_o$ of the observation area is not critical. For the calculation of the average error probability when the decision level $g_{s0}$ is determined from (2.76), $g_{s0}$ is also a monotonically increasing function of $R_o$. However, both $Q_0$ and $Q_d$ are varied by changing $g_{s0}$, and their sensitivities toward the change are different. We found that there are several minima in a certain small region as we vary the radius $R_o$, and we have picked the smallest one for the plots in Fig. 9.
Discussion

It is important in detection theory to calculate the distribution of the statistic as accurately as possible so that the statistical performances of the detectors can be investigated. The statistic given by (2.18) or (2.26) is the sum of $N$ independent random variables and is very common in detection and estimation theory. Since the exact distribution of the statistic cannot be obtained, the Gaussian approximation has been largely used for the statistical evaluations in the past. On the other hand, the method of steepest descent has been proved mathematically to provide approximations of an accuracy that is often high; usually the error of the approximation is of the order of the first term neglected in the asymptotic expansion series. Furthermore, the asymptotic expansion series expressed by (2.55) gives the exact expression of a Gaussian distribution and a fairly good numerical agreement with the other two well-known distributions, the exponential and Poisson distributions, as discussed in Appendix A. It is therefore appropriate for us to use the method of steepest descent to approximate the tail distribution of the ideal and threshold detectors in order to investigate their performances, although more numerical calculations may be required such as searching for saddle point. With the availability of the digital computer at the present time, the asymptotic-expansion approximation should be very useful for solving many problems in detection and estimation theory.
Each of the detection statistics we have discussed is the sum of every ejected number $n_i$ of photoelectrons weighted by a factor as given by (2.18), (2.26), and (2.34). In order to design a better detector, we must make the false-alarm probability of the detector small and its detection probability as large as possible, or, equivalently, make the magnitude of the statistic as small under hypothesis $H_0$ and as large under hypothesis $H_1$ as we can. From Figs. 7, 8 and 9 we can see that the ideal and threshold detectors, which utilize the information about the shape of image, are not much better than the simple detector when the optimum observation area of the simple detector is used to yield maximum detection probability or minimum error probability. In other words, to register the locations of photoelectrons does not help much to improve the detection of the image. This is because when the background light is assumed to be uniform, the effectiveness of reducing the magnitude of the statistic on the average under $H_0$ or increasing the magnitude of the statistic on the average under $H_1$ is limited, especially when the signal-to-noise ratio $D^2$ is low. On the other hand, by varying the observation area, we can optimize the mean numbers of photoelectrons under $H_0$ and $H_1$ effectively and reduce the error probability or improve the reliability of the simple detector.
<table>
<thead>
<tr>
<th>$N_0 = 0.5$</th>
<th>$N_s$</th>
<th>$R_o(Q_0 = 10^{-3})$</th>
<th>$R_o(Q_0 = 10^{-5})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.31</td>
<td>1.37</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.72</td>
<td>1.68</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.72</td>
<td>1.68</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.10</td>
<td>1.68</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2.10</td>
<td>1.98</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.10</td>
<td>1.98</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2.46</td>
<td>2.28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N_0 = 5$</th>
<th>$N_s$</th>
<th>$R_o(Q_0 = 10^{-3})$</th>
<th>$R_o(Q_0 = 10^{-5})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.36</td>
<td>1.37</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.44</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.60</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.60</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.75</td>
<td>1.65</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.82</td>
<td>1.78</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1.82</td>
<td>1.78</td>
<td></td>
</tr>
</tbody>
</table>
Footnotes

Chapter II

1. Helstrom [5].

2. Papoulis [19]. Appendix II-4 gives a brief description of the saddle point method of integration. For a more thorough explanation, see Erdélyi [33] and Jeffreys [34].

3. Helstrom [10], Chapter III.

4. Van Trees [20], Chapter II.

5. Helstrom [21].

6. Helstrom and Wang [22].

7. Helstrom [10], Chapter V, §1d.


9. Feller [24], Chapter VI, §3.

10. Helstrom [5].

11. Farrell [6].

12. Rice [7].


14. Helstrom [14].
Chapter III  Simultaneous detection and estimation of the mean intensity of an optical image

The detection of the image on a photosensitive surface described in Chapter II is based on the knowledge about the intensity and location of the light source. In practice, however, the estimation of the parameters such as the intensity $I_S$ and the center location $z_0$ of the image may be needed as well as the detection. The observer must estimate these parameters as best he can on the basis of the observed set of data $n$. In this chapter we will assume that if the object image is detected, the location of the image is known, but its intensity must be simultaneously estimated. One seeks a strategy to estimate the intensity, $\hat{I}_S = I_S(n)$, which is a function of the data $n = (n_1, n_2, \ldots, n_N)$. Two most important strategies, Bayes and maximum likelihood $(1,2)$, will be discussed. A quadratic cost function will be used to derive the Bayes estimate. When the signal is not too weak, the Bayes estimate is approximately equivalent to the maximum likelihood estimate. An almost optimum estimate is therefore proposed and its statistical properties will be studied. The expectation of the biased estimate will be evaluated and plotted as a function of relative intensity at different observation times or total numbers $N$ of small areas with a truncated Gaussian or parabolic image.
Simultaneous detection and estimation for a single parameter

As we have discussed in connection with the binary hypotheses tests in Chapter II, when a set of data \( \mathbf{b} = \{b_1, b_2, \ldots, b_N\} \) is obtained, the observer will choose the hypothesis between the binary hypotheses \( H_0 \) and \( H_1 \). Since the signal received under hypothesis \( H_1 \) now depends on a certain unknown parameter \( \theta \), the observer must seek a strategy based on the \( N \) measurements of \( \mathbf{b} \) so that when hypothesis \( H_1 \) is chosen, the parameter \( \theta \) will also be estimated\(^{(3)}\). If, however, \( H_0 \) is chosen, it implies that there is only background noise; therefore the parameter \( \theta \) does not exist or is equal to zero. The conditional p.d.f. \( P(\mathbf{b} | \theta) \) is assumed to be given in general. If the cost functions \( C(\hat{\theta}, \theta) \), which are the costs for choosing hypothesis \( H_0 \) or \( H_1 \) and making estimate \( \hat{\theta} = 0 \) or \( \hat{\theta}(\mathbf{b}) \) for given true parameter \( \theta \), are given, and if the prior p.d.f. \( z(\theta) \) of the parameter \( \theta \) and the conditional p.d.f. \( \gamma(\hat{\theta} | \mathbf{b}) \) of the estimate \( \hat{\theta} \) for given data \( \mathbf{b} \) are known, then the Bayes strategy can be used. The average risk\(^{(4)}\) per experiment for simultaneous detection and estimation of the parameter is given by

\[
\bar{C} = \int_{\Omega} d\theta \int_{\Delta} d\hat{\theta} \int_{\Sigma} d^N \mathbf{b} \sigma(\theta) \gamma(\hat{\theta} | \mathbf{b}) P(\mathbf{b} | \theta) C(\hat{\theta}, \theta),
\]

(3.1)

where \( \Omega, \Delta \) and \( \Sigma \) are the spaces for parameters \( \theta \), \( \hat{\theta} \) and \( \mathbf{b} \) respectively.

The Bayes strategy is to properly choose the function \( \gamma(\hat{\theta} | \mathbf{b}) \)
so that the average risk $\overline{C}$ can be minimized. Let us define

$$\gamma(\hat{\theta}|b) = \psi(H_0|b) \delta(\hat{\theta}) + \Pi(\hat{\theta}|b), \quad (3.2)$$

where $\delta(\hat{\theta})$ is the Dirac delta function. $\psi(H_0|b)$ is the conditional probability for the system to choose $H_0$. $\Pi(\hat{\theta}|b)$ is the conditional p.d.f. for estimate $\hat{\theta}$ when $H_1$ is chosen with observed data $b$. Furthermore, we have

$$\int_{\Delta} \Pi(\hat{\theta}|b) \, d\hat{\theta} = \psi(H_1|b), \quad (3.3)$$

where $\psi(H_1|b)$ is the conditional probability for the system to choose $H_1$ with given data $b$ and $\psi(H_1|b) + \psi(H_0|b) = 1$. We also let

$$\sigma(\theta) = \xi \delta(\theta) + (1-\xi) z(\theta), \quad (3.4)$$

where $\xi, 1-\xi$ are the prior probabilities for hypotheses $H_0$ and $H_1$ as discussed in Chapter II. $z(\theta)$ is the prior p.d.f. of parameter $\theta$ under $H_1$.

By substituting (3.2), (3.3) and (3.4) into (3.1) we can write the average risk as

$$\overline{C} = \int_{\Sigma} d^N \{ \xi [\psi(H_0|b)c_{00} + \int_{\Delta} d\hat{\theta} \Pi(\hat{\theta}|b) c_{10}(\hat{\theta}) ] P_b(\xi) \}$$
\[ + (1-\varepsilon) \int_\Omega d\theta \, z(\theta) \left[ \psi(H_0 | b) \right] C_{01}(\theta) + \int_\Delta d\theta \, \Pi(\hat{\theta} | b) \]
\[ c_{11}(\hat{\theta}, \theta) \bigg] P(b | \theta) \bigg) \]

(3.5)

where \( P_0(b) = P(b | H_0) \),
\( C_{00} \) is the cost of choosing \( H_0 \) when \( H_0 \) is true,
\( C_{01}(\theta) \) is the cost of choosing \( H_0 \) when \( H_1 \) is true with parameter \( \theta \),
\( C_{10}(\hat{\theta}) \) is the cost of choosing \( H_1 \) with estimate \( \hat{\theta} \) when \( H_0 \) is true,
\( C_{11}(\hat{\theta}, \theta) \) is the cost of choosing \( H_1 \) with estimate \( \hat{\theta} \) when \( H_1 \) is true with parameter \( \theta \).

Now since
\[
P(b | \theta) \, z(\theta) = P(\theta | b) \, P(b) , \]

(3.6)

where \( P(\theta | b) \) is the posterior p.d.f. and
\[
P(b) = \int_\Omega P(b | \theta) \, z(\theta) \, d\theta , \]

(3.7)

we can also write the average risk \( \overline{C} \) in terms of the conditional risks as
\[
\bar{C} = \int_{\Sigma} \text{d}N_{\theta} \, P(\theta) \left[ C_{H_{0}|b} + \int_{\Delta} \text{d}\theta \, \Pi(\hat{\theta}|b) \, C(\hat{\theta}|b) \right], \tag{3.8}
\]

where

\[
C_{H_{0}|b} = \psi(H_{0}|b) \left[ \xi \frac{C_{00} P_{0}(b)}{P(b)} + (1-\xi) \int_{\Omega} \text{d}\theta \, C_{01}(\theta) \frac{P_{1}(b)}{P(b)} \right], \tag{3.9}
\]

is the conditional risk when \(H_{0}\) is chosen for given \(b\) and

\[
C(\hat{\theta}|b) = \xi \frac{C_{10}(\hat{\theta}) P_{0}(b)}{P(b)} + (1-\xi) \int_{\Omega} \text{d}\theta \, C_{11}(\hat{\theta},\theta) \frac{P_{1}(b)}{P(b)} \tag{3.10}
\]

is the conditional risk when \(H_{1}\) is chosen with \(\hat{\theta}\) for given \(b\). Now the p.d.f. \(P(b)\) and conditional p.d.f. \(\Pi(\hat{\theta}|b)\) are both positive over the space \(\Sigma\) of the outcome \(b\), and the conditional risk \(C_{H_{0}|b}\) is not a function of estimate \(\hat{\theta}\), the average risk \(\bar{C}\) is minimized with respect to \(\hat{\theta}\) by making the conditional risk \(C(\hat{\theta}|b)\) as small as possible for every set of data \(b\). Now we choose the p.d.f. \(\Pi(\hat{\theta}|b)\) as

\[
\Pi(\hat{\theta}|b) = \psi(H_{1}|b) \delta(\hat{\theta} - \hat{\theta}(b)), \tag{3.11}
\]

where \(\hat{\theta}(b)\) is the optimum estimate \(\hat{\theta}\) obtained from minimizing the conditional risk \(C(\hat{\theta}|b)\) and can be determined from the following
Thus the minimum average risk from (3.8) at $\hat{\theta} = \hat{\theta}(b)$ becomes

$$
\overline{c}_{\text{min}} = \int_\mathcal{E} d^N \mathcal{E} \left\{ \left[ 1 - \psi(H_1 | b) \right] \left[ \xi C_{00} P_0(b) + (1 - \xi) P(b) \int_{\Omega} d\theta \ C_{01}(\theta) P(\theta | b) \right] + \psi(H_1 | b) \left[ \xi C_{10}(\hat{\theta}(b)) P_0(b) + (1 - \xi) P(b) \int_{\Omega} d\theta \ C_{11}(\hat{\theta}(b), \theta) P(\theta | b) \right] \right\} \tag{3.13}
$$

If we now define

$$
A(b) = \xi C_{00} P_0(b) + (1 - \xi) P(b) \int_{\Omega} d\theta \ C_{01}(\theta) P(\theta | b)
$$

$$
B(b, \hat{\theta}(b)) = \xi C_{10}(\hat{\theta}(b)) P_0(b) + (1 - \xi) P(b) \int_{\Omega} d\theta \ C_{11}(\hat{\theta}(b), \theta) P(\theta | b), \tag{3.14}
$$
then (3.13) can be written as

\[ C_{\text{min}} = \int_{\Sigma} d^{n} \{ A(b) + \psi(H_1 b) [B(b, \hat{\theta}(b)) - A(b)] \}. \]

(3.15)

As the cost functions can be chosen so that both functions \( A(b) \) and \( B(b, \hat{\theta}(b)) \) are positive. Then we can further minimize the average risk from (3.15) so that

\[
\begin{align*}
\psi(H_1 | b) &= 1 \quad \text{if } B(b, \hat{\theta}(b)) < A(b), \\
\psi(H_1 | b) &= 0 \quad \text{or} \\
\psi(H_0 | b) &= 1 \quad \text{otherwise}.
\end{align*}
\]

(3.16)

This is equivalent to defining a cost likelihood ratio \( \Lambda_c \) where

\[
\Lambda_c = \frac{(1-\xi) P(b) \int_{\Omega} d\theta [C_{11}(\theta) - C_{10}(\hat{\theta}(b), \theta)] P(\theta | b)}{\xi P_0(b) [C_{10}(\hat{\theta}(b)) - C_{00}]},
\]

(3.17)

and a decision level \( \Lambda_{c0} = 1 \) so that hypothesis \( H_1 \) will be chosen if \( \Lambda_c > \Lambda_{c0} \) and estimate \( \hat{\theta}(b) \) will be issued; otherwise, \( H_0 \) will be
chosen.

When there is no information about the cost functions, the prior probabilities of the hypotheses, or the prior p.d.f. \( z(\theta) \) of the parameter, the maximum-likelihood strategy can be used. The estimate \( \hat{\theta} \) will be determined from maximizing the posterior p.d.f. \( P(\theta | b) \), which is expressed in (3.6). Since \( P(b) \) is independent of \( \theta \), it is equivalent to maximizing the product of \( P(b | \theta) \) and \( z(\theta) \). With very little knowledge about \( z(\theta) \), we must assume that the prior p.d.f. \( z(\theta) \) is very broad, so that \( z(\theta) \) will not affect the decision and estimate we have made. This implies that the estimate \( \hat{\theta} \) is simply obtained by maximizing the joint conditional p.d.f. \( P(b | \theta) \) or determined from the following equation

\[
\frac{\partial}{\partial \theta} P(b | \theta) = 0 .
\] (3.18)

The estimate \( \hat{\theta} \) determined from (3.18) will be a function of the data \( b \).

When the data \( b \) are a set of discrete random variables, the function \( P(b | \theta) \) will be the conditional probability function instead of conditional p.d.f. All the equations we have discussed in this chapter are still valid except that we have to change the integrals \( \int_{\mathbb{D}}^{N} dN \) into summations \( \sum_{b} \). For the detection of the image and estimation of its intensity of a light source, both Bayes strategy and maximum likelihood strategy will be discussed in the following paragraphs.
The operation of simultaneous detection and estimation of the signal and its parameter $\theta$ can be described by the following block diagram.

![Block diagram for simultaneous detection and estimation of a single parameter $\theta$ from the signal.](image)

Figure 11  Block diagram for simultaneous detection and estimation of a single parameter $\theta$ from the signal.
Bayes strategy for detecting the image and estimating its intensity

of a point source with a quadratic cost function

When the photosensitive surface is divided into N small areas dA, the data n are the numbers of photoelectrons ejected from the N small areas dA, i = 1, 2 . . . N; n is a set of N discrete random variables. As TW >> 1, they are Poisson distributed and statistically independent. The joint conditional probability under hypothesis H_0 can be written as

\[ P(n|I_S=0) = P_0(n) = \prod_{i=1}^{N} \mu_i^{n_i} e^{-\mu_i}/n_i! , \]  

(3.19)

where

\[ \mu_i = \alpha T I_b dA \]  

(3.20)

is the average number of photoelectrons ejected from each area dA due to the background light illumination. The joint conditional probability under hypothesis H_1 can be written as

\[ P(n|I_S) = \prod_{i=1}^{N} (\mu + I_S q_i)^{n_i} e^{-(\mu + I_S q_i)}/n_i! , \]  

(3.21)

where

\[ q_i = \alpha T Y_i dA , \quad Y_i = \gamma(\chi_i) \]  

(3.22)
with \( \gamma(x) \) as the weighting function given by (2.16). We let the cost functions described in (3.5) be

\[
\begin{align*}
C_{10}(I_S) &= C_{10}, \\
C_{11}(I_S, I_S) &= C_0(I_S - I_S)^2,
\end{align*}
\tag{3.23}
\]

where \( C_{00}, C_{01}, C_{10}, C_0 \) are all given constants.

The Bayes estimate \( \hat{I}_S(n) \) can be determined from (3.12)

\[
\hat{I}_S(n) = \int_0^\infty I_S P(I_S|n) \, dI_S,
\tag{3.24}
\]

which is the conditional expectation of the intensity \( I_S \). The posterior p.d.f. \( P(I_S|n) \) can be expressed by using (3.6) and (3.7) as

\[
P(I_S|n) = \frac{z(I_S) P(n|I_S)}{\int_0^\infty z(I_S) P(n|I_S) \, dI_S}
\tag{3.25}
\]

for positive values \( I_S \). The denominator of (3.25) can be written as

\[
\int_0^\infty z(I_S) P(n|I_S) \, dI_S
= \int_0^\infty dI_S \, z(I_S) \prod_{i=1}^N \left( \frac{\mu + I_S q_i}{1 + I_S} \right)^{n_i} e^{-\left(\mu + I_S q_i\right)} / n_i!.
\tag{3.26}
\]
When the p.d.f. \( z(I_S) \) is given, the Bayes estimate from (3.24) can be evaluated.

We now assume that the p.d.f. \( z(I_S) \) is a gamma distributed probability function

\[
z(I_S) = \frac{C^B}{\Gamma(B)} I_S^{B-1} e^{-CI_S} U(I_S) \quad (3.27)
\]

with \( B, C \) real and positive constants. \( U(x) \) is the step function.

If the object light is very much weaker than the background light, \( I_S \ll I_b, \quad (3.26) \) can be calculated approximately as

\[
\int_0^\infty z(I_S) P(n|I_S) \, dI_S \\
= \int_0^\infty dI_S \frac{C^B}{\Gamma(B)} I_S^{B-1} e^{-CI_S} \\
\quad \cdot \prod_{i=1}^N \\mu \left( 1 + n_i \frac{I_S}{I_b} \gamma_i \right) e^{-\mu} / n_i! \\
= \prod_{i=1}^N e^{-\mu} \mu \frac{n_i}{\mu} \left( 1 + n_i \gamma_i \frac{B}{CI_b} \right) / n_i! \quad (3.28)
\]

The Bayes estimate from (3.24) becomes

\[
\hat{I}_S(n) \approx \prod_{i=1}^N e^{-\mu} \frac{n_i}{\mu} \frac{B}{\gamma_i} \left[ 1 + n_i \gamma_i (B+1)/CI_b \right] / n_i! \\
= \frac{e^{-\mu} \mu n_i \left( 1 + n_i \gamma_i B/CI_b \right) / n_i!}{e^{-\mu} \mu n_i \left( 1 + n_i \gamma_i B/CI_b \right) / n_i!}
\]
\[ = \frac{B}{C} \prod_{i=1}^{N} \left[ 1 + \frac{1}{B \left( 1 + \frac{C}{B} \frac{I_b}{n_i Y_i} \right)} \right] \]  

(3.29)

where all the values \( n_i \), \( Y_i \), \( I_b \), \( B \), \( C \) are known. If the object light is so strong that we can assume that there are \( m_1 \) areas \( dA \) which are so close to the image center and such that \( \gamma_i I_s \gg I_b \), and \( m_0 \) areas which are far away from the image center and \( \gamma_i I_s \ll I_b \) \((m_0 + m_1 = N\) is the total number of areas into which we have divided the image), then (3.26) can be calculated approximately

\[
\int_0^{\infty} z(I_s) P(n|I_s) \, dI_s \approx 2 \frac{C^B}{F(B)} \mu^M_0 e^{-m_0 \nu} \left( \prod_{i=1}^{m_1} q_i n_i \right) F_N^{-1} \int_0^{\infty} B^{-1 + M_1} e^{-(C+M q) I_s} \, dI_s ,

(3.30)

where

\[
F_N = \prod_{i=1}^{N} n_i! , \quad M_q = \sum_{i=1}^{m_1} q_i , \quad M_0 = \sum_{i=1}^{m_0} n_i ,
\]

is the total number of photoelectrons ejected from the \( m_0 \) areas far from the image center.
\[ M_1 = \sum_{i=1}^{m_1} n_i \]

is the total number of photoelectrons ejected from the \( m_1 \) areas close to the image center, where

\[
\int_0^{\infty} I_S B^{-1} + M_1 e^{-(C+M_q)I_S} dI_S = \frac{\Gamma(B+M_1)}{(C+M_q)^{B+M_1}} .
\]

(3.31)

The Bayes estimate \( \hat{I}_S(n) \) according to (3.24) can be approximated as

\[
\hat{I}_S(n) \approx \frac{\Gamma(B+1+M_1)}{(C+M_q)^B} = \frac{B+M_1}{C+M_q} .
\]

(3.32)

When \( M_1 >> B \) we have

\[
\hat{I}_S(n) \approx K_b \sum_{i=1}^{m_1} n_i
\]

(3.33)

with \( K_b^{-1} = C + M_q \).
Maximum-likelihood strategy and the almost optimum intensity estimate

When there is no information about the cost function, the prior probabilities of the hypotheses, or the prior p.d.f. \( z(I_S) \) of the intensity, the maximum-likelihood strategy will be used and the estimate \( \hat{I}_S(n) \) will be determined from (3.18); it satisfies the following equation

\[
\sum_{i=1}^{N} \frac{n_i q_i}{q_i \hat{I}_S(n) + \mu} = \sum_{i=1}^{N} q_i. \tag{3.34}
\]

When the object light is so much weaker than the background light that \( I_S \ll I_b \), the estimate \( \hat{I}_S(n) \) can be determined approximately from

\[
\sum_{i=1}^{N} q_i \approx \sum_{i=1}^{N} n_i q_i \mu^{-1} [1 - q_i \hat{I}_S(n) \mu^{-1}]
\]

or

\[
\hat{I}_S(n) = \sum_{i=1}^{N} \mu^2 (n_i \mu^{-1} - 1) q_i / \sum_{i=1}^{N} n_i q_i^2. \tag{3.35}
\]

If the object light is so strong that there are \( m_1 \) areas close to the image center and \( q_i \hat{I}_S(n) \gg \mu \), and \( m_0 \) areas away from the image center and \( q_i \hat{I}_S(n) \ll \mu \) so that they contribute little to the sum, then we approximate the estimate \( \hat{I}_S(n) \) from (3.34) by

\[
\hat{I}_S(n) \approx \kappa \sum_{i=1}^{m_1} n_i, \tag{3.36}
\]
The estimate \( \hat{I}_S(n) \) we have just obtained resembles the approximate form of the Bayes estimate given by (3.33) when the object light is not too weak. In general we may not have any information about the cost functions and the prior p.d.f. \( z(I_S) \). We therefore propose an almost optimum estimate \( \hat{I}_{Sa}(n) \) as

\[
\hat{I}_{Sa}(n) = K \sum_{i=1}^{N} (n_i - \Theta) U(n_i - \Theta),
\]

where \( U(x) = 1 \) for \( x > 0 \) and \( 0 \) for \( x \leq 0 \), \( K \) is defined in (3.36). \( \Theta \) is the threshold level which can be determined from a preassigned false-alarm probability \( Q_0 \) such that \( \Theta \) is the smallest integer that satisfies the following equation

\[
N \sum_{n=0}^{\Theta} \frac{\mu^n}{n!} e^{-\mu} / n! > N - Q_0.
\]

In other words, the number \( n_i \) of photoelectrons ejected from each area \( dA_i, i = 1, 2 \ldots N, \) will be compared with the decision level \( \Theta \); if \( n_i > \Theta \), we will consider that some of the photoelectrons ejected from the \( i \)th area of the surface are due to the object illumination. We therefore will keep the term. If \( n_i < \Theta \) we
assume that the number \( n_i \) of photoelectrons ejected from the surface are due only to the background illumination; thus we will discard the term. When all the numbers \( n_i, i = 1, 2 \ldots N \) are less than \( \theta \), the hypothesis \( H_0 \) will be chosen. Otherwise, the hypothesis \( H_1 \) will be chosen, and the intensity estimate \( \hat{I}_S(u) \) is calculated according to (3.37).
Performance of the almost optimum intensity estimate for a truncated Gaussian or a parabolic image

The estimate \( \hat{I}_S(n) \) we have discussed is a function of the \( N \) measurements \( n \). Since the data \( n \) are a set of random variables, no two experiments will yield the same value of the estimate \( \hat{I}_S(n) \) even though the true value of the parameter \( I_S \) is the same in both. The most one can hope for is that the estimate \( \hat{I}_S(n) \) will be close to the true value \( I_S \) in the sense of "on the average".

The mean and variance of the estimate \( \hat{I}_{S_a} \) we have just discussed can be derived as follows

\[
E[\hat{I}_{S_a}] = E \left[ \sum_{i=1}^{N} \frac{n_i - \theta}{n_i - \theta} U(n_i - \theta) \right]
\]

\[
= K \sum_{i=1}^{N} E[(n_i - \theta) U(n_i - \theta)]
\]

\[
= K \sum_{i=1}^{N} \left[ \frac{n_i - \theta}{n_i - \theta} - \sum_{n_i = 0}^{\theta} (n_i - \theta) P(n_i | I_S) \right]
\]

\[
\text{Var}[\hat{I}_{S_a}] = \text{Var} \left[ \sum_{i=1}^{N} \frac{n_i - \theta}{n_i - \theta} U(n_i - \theta) \right]
\]

\[
= K^2 \sum_{i=1}^{N} \text{Var}[(n_i - \theta) U(n_i - \theta)]
\]

\[
= K^2 \sum_{i=1}^{N} \left[ \frac{n_i}{n_i^2} + \frac{1}{n_i^2} - 2n_i \theta + \theta^2 \right]
\]
\[- \sum_{n_i=0}^{\theta} (n_i-\theta)^2 P(n_i | I_S) \]
\[- \left[ \bar{n}_i - \theta - \sum_{n_i=0}^{\theta} (n_i-\theta) P(n_i | I_S) \right]^2 \]  

(3.39)

where

\[ P(n_i | I_S) = \frac{n_i^{n_i} e^{-n_i}}{n_i!} \]
\[ \bar{n}_i = \mu + q_i I_S \]

To investigate the statistical performance of the estimate \( \hat{I}_{Sa} \) as discussed in (3.37), a truncated Gaussian image will be postulated. The image surface with radius \( R_o \) is divided into \( M \) small rings as shown in Fig. 12.

The average number of photoelectrons ejected from each ring by the background light illumination is assumed to be evenly distributed and symmetrical with respect to the image center. The area \( dR_i \) of the \( i \)th ring can be calculated as

\[ dR_i = \int_{(i-1)\Delta \rho}^{i\Delta \rho} \int_{0}^{2\pi} \rho d\rho d\phi \]
\[ = \pi(2i - 1)\Delta \rho^2 = (2i - 1) A_o \]  

(3.40)
where $A_0 = \pi \Delta \rho^2$, $\Delta \rho = \rho_i - \rho_{i-1}$ for $i = 1, 2 \ldots M$. Here $\rho^2 = x_1^2 + x_2^2$ with $x_1 = \rho \cos \phi$, $x_2 = \rho \sin \phi$ as the rectangular coordinates used in Fig. 3.

The total number of rings is

$$M = \frac{R_o}{\Delta \rho} \quad (3.41)$$

If we further divide each $i$th ring into a number $(2i-1)$ of small equal areas $dA_i$ such that $dA_i = A_0$, the total number of small equal areas $dA_i$ from the whole image is

$$N = \sum_{i=1}^{M} (2i-1) = M^2. \quad (3.42)$$
The distribution of the mean number of photoelectrons ejected from the image surface by the object light is described by the weighting function \( \gamma(x_i) \) as given in (2.16) and (2.56) for a Gaussian image, or in polar coordinates we can write

\[
x(\rho) = B_0 e^{-\rho^2/2\sigma^2},
\]

(3.43)

where \( B_0 \) is a constant and \( \sigma \) is the width of the Gaussian function.

The average number of photoelectrons emitted from each small area \( dA_i \) is assumed to be evenly distributed and is proportional to the intensity \( I_S \). We calculate the coefficient

\[
q_i' = \alpha_T \int_0^{2\pi} \int_{(i-1)\Delta\rho}^{i\Delta\rho} \gamma(\rho) \rho d\rho d\phi
\]

\[
= \alpha_T B_0 2\pi \sigma^2 \left\{ e^{-(i-1)^2 R_A^2/2N} - e^{-i^2 R_A^2/2N} \right\},
\]

(3.44)

where

\[
q_i' = q_i(2i-1)
\]

(3.45)

with \( q_i \) defined in (3.22). \( R_A = R_o/\sigma \). \( R_o \) is the radius of the truncated image under observation. \( i \) is the index \((i = 1, 2 \ldots M)\).

Because of the presence of the background light, one would
like to investigate the estimate $\hat{I}_{Sa}$ with respect to the true intensity $I_S$ in terms of the signal-to-noise ratio. In other words one could assume that the energy of the background light falling on the image surface within the radius $R_0$ of the truncated image is equivalent to that from the object light with intensity

$$ I_{SO} = \frac{\mu M^2}{\sum_{i=1}^{M} q_i^2} \tag{3.46} $$

Since the total average number of photoelectrons ejected from the surface of the image is proportional to the incident energy, the intensity of the point source can be described by

$$ I_S = d^2 I_{SO}, \tag{3.47} $$

where $d^2$ is called the signal-to-noise ratio for the truncated Gaussian image.

The mean and variance of the estimate $\hat{I}_{Sa}$ according to (3.37) can be written as

$$ E[\hat{I}_{Sa}] = \kappa \sum_{i=1}^{M} (2i-1) \left[ \bar{n}_i - \bar{n}_i - \sum_{\theta} (n_i - \theta) P_{\lambda}(n_i) \right] $$

$$ \text{Var}[\hat{I}_{Sa}] = \kappa^2 \sum_{i=1}^{M} (2i-1) \left( \bar{n}_i^2 + \bar{n}_i^2 - 2\bar{n}_i \theta + \theta^2 \right) $$
\[- \sum_{n_1=0}^{\theta} (n_1-\theta)^2 p_1(n_1) - \left[ \frac{\bar{n}_1 - \theta - \sum_{n_1=0}^{\theta} (n_1-\theta) p_1(n_1)}{\bar{n}_1} \right]^2 \]  

(3.48)

where \( p_1(n_1) = \frac{n_1!}{n_1!} e^{-n_1} \) with \( \bar{n}_1 = \mu + q_{1}'I_S/(2i-1) \) and \( \mu \) and \( q_{1}' \) are given in (3.20) and (3.45).

We also define the error of the estimate \( \hat{I}_{Sa} \) by

\[
\text{Err} = \frac{(\text{Var} \hat{I}_{Sa})^{1/2}}{E[\hat{I}_{Sa}]} 
\]

(3.49)

as the ratio of the standard deviation to the mean value of the estimate \( \hat{I}_{Sa} \) given by (3.48).

To investigate the estimate \( \hat{I}_{Sa} \) in terms of various parameters we assume the background light intensity corresponding to an object light with unit intensity \( I_{S0} = 1 \). The average number of photoelectrons ejected from each area \( dA \) due to the background light is assumed to be 0.1 for an observation time \( T \) and coefficient \( a \) (i.e., \( \mu = aT I_{b} A_0 = 0.1 \)). The false-alarm probability is preassigned at 0.01. Where the parameters \( B_0 \) and \( \sigma \) are set at 1, the expected value of the estimate \( \hat{I}_{Sa} \) is calculated at various ratios \( R_A = 0.5, 1, 2 \) and 3 and is plotted as a function of the relative intensity \( I_S \) defined in (3.47) in Fig. 13. As the size of the observation area changes, the estimate \( \hat{I}_{Sa} \) will also be affected. In Fig. 14, the estimate \( \hat{I}_{Sa} \) is plotted as a function of the relative intensity \( I_S \).
Figure 13  Mean intensity estimate $\hat{I}_{sa}$ vs. relative intensity $I_s$ for truncated Gaussian image; $Q_0 = 10^{-2}$, the number of rings is 100. Mean background counts $\mu = 0.1$ at $R_A = \sigma = 1$ for time $T$. Curves are indexed by parameter $R_A$ defined in (3.44).
Figure 14  Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for truncated Gaussian image; $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_A = 1$ for 100 rings for time $T$. Curves are indexed by the parameter $M$ defined in (3.41).
Figure 15  Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_s$ for truncated Gaussian image; $Q_0 = 10^{-2}$, $M = 100$, $\mu = 0.1$ at $R_A = 1$ for time $T$. Curves are indexed by observation intervals $T$. 
when the observation area is fixed but divided into three different numbers of rings, 25, 50 and 100. In Fig. 15 the estimate \( \hat{I}_{Sa} \) is also plotted as a function of \( I_S \) with three different observation times \( T, 10T \) and \( 100T \).

Observations also have been made by assuming the image as a parabolic function. That is (3.43) replaced by

\[
\gamma'(\rho) = B'(R_o^2 - \rho^2), \quad |\rho| < R_o, \\
= 0, \quad |\rho| > R_o,
\]

where \( R_o \) is the radius of the image; \( B' \) is a constant. The coefficient \( q_{1''} \) can be calculated from

\[
q_{1''} = aT \int_0^{2\pi} d\phi \int_{(i-1)\Delta\rho}^{i\Delta\rho} \gamma'(\rho) \rho d\rho.
\]

\[
= aT B' \pi R_o^4 M^{-2} [(2i-1)M^2 - (2i^3-3i^2+2i-1)]^{-1}.
\]

The mean, variance, and error can be calculated by using (3.48) and (3.49). At \( B' = 1, R_o = 1 \) we have plotted the expected value of the estimate \( \hat{I}_{Sa} \) as a function of \( I_S \) at three different values of parameters \( M \) and \( T \) in Fig. 16 and Fig. 17 with \( Q_0 = 0.01 \). The error of the estimate \( \hat{I}_{Sa} \) for the parabolic image is also plotted as a
function of the average background photocelectrons $\mu$ at fixed $I_S = 10$ and 20 in Fig. 18, where the error for $\mu = 0.1$ is 9.24% at $I_S = 10$ or 3.15% at $I_S = 20$ with the image surface divided into 100 rings. The average number $\mu$ is then varied either by changing the size of each $dA$ or the observation time $T$. 
Figure 16 Mean estimate $I_{sa}$ vs. relative intensity $I_S$ for parabolic image; $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_0 = 1$ for time $T$. Curves are indexed by parameter $M$. 
Figure 17  Mean estimate $\hat{I}_{Sa}$ vs. relative intensity $I_S$ for parabolic image: $Q_0 = 10^{-2}$, $\mu = 0.1$ at $R_0 = 1$ for time $T$. Curves are indexed by observation intervals $T$. 
Figure 18  Error vs. mean counts $\mu$
for parabolic image; $Q_0 = 10^{-2}$. Curves
are indexed by parameters $I_S$ defined in
(3.47) and M.
Increase Observation Time $T$

Reduce Ring number $M$

$I_5 = 10$

$I_5 = 20$

$\mu$ - Average number of charges from each area $dA$
Discussion

A statistical model for simultaneous detection and estimation of a stochastic image is proposed. When the prior p.d.f. of the source intensity and the prior probability $\xi$ are known and the cost functions are available, the Bayes strategy leads to the optimum operation. If there is no information about the prior probability and the cost functions, the maximum-likelihood strategy can be used. Upon examining the estimate $\hat{I}_S$ for both strategies, it is interesting to note that the Bayes estimate $\hat{I}_S$ from (3.33) for a quadratic cost function is approximately the same as the maximum-likelihood estimate from (3.36) when the signal is not very weak.

For a Gaussian image, the bias and linearity will be improved when $R_A$ is increased; however, after $R_A$ is larger than 2, very little improvement will be gained by increasing $R_A$ farther. In order to reduce the bias of the estimate, it is always wise to use longer observation time, to divide the observed image into a smaller number $M$ of rings, or to use larger observation image area. In addition, a suitable constant should be added into the estimate $\hat{I}_{Sa}$ to compensate for the bias. The observed image has an effective range of $R_A$ between 1 and 2 for the Gaussian truncated image. For the image of parabolic shape, one should always use the whole image area. As long as the observation area on the image surface is properly defined, the shape of the function $\gamma(p)$ has no significant effect on the quality of the estimate. To reduce the estimation error, increasing
the observation time is more effective than decreasing the total number $M$ of rings as we can see from Fig. 18.

When the false-alarm probability $Q_0$ is preassigned, randomization can be used and the probability $f$ for choosing $H_1$ when $n_1 = 0$ can be calculated from the difference in (3.38) or

$$f = \left(1 + Q_0/N - \sum_{n=0}^{\theta} \frac{\mu^n}{n!} \right) e^{\mu} \mu^0 \theta!.$$

The estimate $\hat{I}_{Sa}$ from (3.37) can be modified as

$$\hat{I}'_{Sa} = K \left[ \sum_{i=1}^{N} n_{1} U(n_{1} - \theta - 0.5) + f m_{0} - C \right] \quad (3.52)$$

$m_0$ is the total number of the areas where $n_1 = 0$. $C$ is some constant. Since the bias of the estimate approaches a constant as the intensity becomes larger, as we can see from Fig. 15 or Fig. 17, a proper constant $C$ can be found to compensate for the bias.
Footnotes

Chapter III

1. Helstrom [10], Chapter VIII.
2. Van Trees [20], Chapter II.
3. Middleton [9].
4. Middleton [26], Chapter I.
5. Helstrom [10], Chapter VIII, eq. (1.11).
Chapter IV  Detectors for resolving two point sources in optical communication by photon-counting techniques

The techniques we have discussed of counting the photoelectrons from the photosensitive surface for the detection and estimation of the object light intensity have great usefulness in binary optical communication. Since those devices, such as photomultipliers, image tubes, and image orthicons are well developed in commercial industry, photon-counting techniques, combined with adequate digital computer facilities can be used for many purposes. In this chapter we will further discuss some detectors used for resolving the point sources in optical communication by photon-counting techniques. The ideal detector described in Chapter II can be used for distinguishing two light sources of equal radiant power from a single source with twice the radiant power or for deciding which one of the two light sources is present. In the absence of background light, the performance of the ideal detector for deciding which one of the two light sources is present will be compared with two other receivers, the simple detector and the counting comparator. The intensities of the optical images due to the point sources will be postulated to be Gaussian distributed spatially. As uniform background light also passes through the aperture, the error probabilities of the simple detector and the counting comparator will be calculated with a finite observation area. When the point sources are mislocated, the effects on the ideal detector and the counting comparator will be discussed, and their average error probabilities will be calculated.
Ideal detector for resolving two light sources from the one light source

When two point sources with equal radiant power are very close to each other, they may often appear to the observer as one source with twice the radiant power, located midway between them. The criterion given by Rayleigh for two point sources to be resolved is that the peak illuminance of the diffraction pattern of one source will not fall closer than the first minimum of the diffraction pattern of the other. In this section, we will construct the optimum statistic for the resolution of the two point sources from the standpoint of a classical hypothesis test\(^{(1)}\). The observer will choose between two hypotheses, (\(H_1\)) two point sources whose images have equal intensities are present at \(x\) and \(-x\) in the image plane, and (\(H_0\)) one source with twice the power is located at the origin. The optimum statistic \(g\) described by (2.18) in Chapter II, based on observing the set of \(n\) photoelectrons, can still be applied here. The intensity \(I_k(x)\) given by (2.15) can now be written as

\[
I_k(x) = I_b + I_S \gamma_k(x), \quad \text{for } k = 0, 1, \tag{4.1}
\]

where the weighting functions are given by

\[
\gamma_k(x_{11}, x_{21}) = \frac{1}{2} [\gamma(x_{11} - kx_{0}, x_{21}) + \gamma(x_{11} + kx_{0}, x_{21})], \tag{4.2}
\]

and \(\gamma(x_{11}, x_{21}) = \gamma(x_1)\) is described by (2.16). \(I_b\) is the uniform intensity of the background light. \(I_S\) is the intensity of the point source.
The optimum statistic for binary detection of the point sources can be derived from the likelihood ratio in a way similar to (2.18), and we write

\[ e_R = \sum_{i=0}^{N} n_i \ln H_R(x_i), \]  

(4.3)

where the function \( H_R(x_i) \) defined in (2.17) becomes

\[ H_R(x_i) = \frac{I_1(x_i)}{I_0(x_i)} \]

\[ = 1 + \frac{I_S[y_1(x_i) - y_0(x_i)]}{I_b[1 + \frac{I_S}{I_b} y_0(x_i)]} \]

\[ = 1 + D^2[u_1(x_i) - u_0(x_i)][1 + D^2u_0(x_i)]^{-1}, \]  

(4.4)

and \( u_k(x_i) = A_0 \gamma_k(x_i), k = 0, 1, D^2 = I_S/I_b A_0 \). Here \( A_0 \) is an arbitrarily defined finite area, as discussed in Chapter II.

Since the weighting function lies between 0 and 1,

\[ 0 < \gamma_k(x) < 1 \]  

for \( k = 0 \) or 1. We will have \( 0 < H_R(x_i) < 1 \) for some areas where \( \gamma_0(x) > \gamma_1(x) \). This implies that the statistic \( g_R \) is no longer a sum of non-negative random variables. As \( TW >> 1 \), the m.g.f. of the optimum statistic \( g_R \) according to (2.19) from Chapter II becomes

source, whose image is centered at \( (x_0,0) \) or \( (-x_0,0) \) on the photosensitive surface under hypothesis \( H_1 \) as shown by Figure 3.
where \( A \) is the area of the receptor, \( T \) is the observation interval, and \( \alpha \) is the coefficient defined in Chapter II. Since the average number difference \( \Delta N = \sum_{i=1}^{N} (\bar{n}_{i1} - \bar{n}_{i0}) \) vanishes in this case, the statistic \( g_R \) from (4.3) is the logarithm of the likelihood ratio defined by (2.11). It is easy to show \( h_{RL}(s) = h_{R0}(s-1) \). Also when Bayes strategy is used, the decision level \( g_{R0} \) is zero when the prior probability is \( \xi = \frac{1}{2} \) as we can see from (2.13).

In order to use the numerical method of steepest descent, the complex phase, defined by (A7) in Appendix A when (4.5) is used, can be written as

\[
\phi_{Rk}(s) = \varepsilon_{R0}^{-1} a_{TI} b \int_{X} d^2y [1 + b^2 u_k(y)] (H_R(y))^{-s} - 1 + s, \quad k=0,1. \quad (4.6)
\]

The saddle-point will be determined by solving the equation \( \phi_{Rk}'(s) = 0 \) or

\[
\varepsilon_{R0} = a_{TI} b \int_{X} d^2y [1 + b^2 u_k(y)] (H_R(y))^{-s} - 1 \ln H_R(y), \quad \text{for } k=0,1. \quad (4.7)
\]

The nth derivative of the complex phase for both \( H_1 \) and \( H_0 \) as \( n \geq 2 \) can be written as

\[
\phi_{Rk}^{(n)}(s) = (-1)^n a_{TI} b \int_{X} d^2y [1 + b^2 u_k(y)] (H_R(y))^{k-s} \ln^n H_R(y), \quad (4.8)
\]

for \( k=0,1 \).
and \( \phi_{R0}^n(s_0) = \phi_{R1}^n(s_1) \) for \( n \geq 1 \) where \( s_k \) is the saddle-point under \( H_k \) for \( k = 0,1 \) and satisfies (4.7) and (2.51). Because the function \( H_R(x) \) has the term \( \gamma_0(x) \) in the denominator as given in (4.4), the evaluation of (4.6), (4.7) or (4.8) will involve a double integration. The iteration procedure to search for the saddle point as discussed in Chapter II will be tedious numerically although it can be carried out by the digital computer. In the meantime, the \( n \)th cumulant under hypothesis \( H_k \) for \( k = 0 \) or 1 is

\[
G_{Rk}^n = (-1)^n \frac{d^n}{ds^n} \ln h_{Rk}(s) \bigg|_{s=0}
\]

\[
= (-1)^n n \text{e}_k \int_{A} d^2x [1 + d^2 u_k(x)] \ln^n H_R(x), \text{ for } k=0,1. \quad (4.9)
\]

If all the \( n \)th cumulants for \( n > 2 \) are negligible in comparison with the second cumulant, the distribution of the statistic \( g_R \) can be approximated by a Gaussian function. The mean and variance of the statistic can be calculated from (4.9) at \( n = 1 \) and \( n = 2 \). We have not carried out any numerical example for this ideal detector. However, we will investigate the performance of the ideal detector with a Gaussian image in the next paragraph for the decision whether a single source is located at \( x \) or at \( -x \).
Ideally detector for binary detection in optical communication

In order to transmit the information about the bit "1" or the bit "0", an optical system can be used by focusing a radiant source either on the upper half plane or the lower half plane of a photosensitive surface. The observer will choose between the hypotheses: (H₁) the bit "1" and (H₀) the bit "0" on the basis of the set of photoelectrons ejected from a large number of small areas of the photosensitive surface. The intensity of the image of the light source at point \((x_i)\) of the surface under hypothesis \(H_k\) for \(k = 0\) or \(1\) in the presence of the uniform background light can be described by

\[
I_k(x_i) = I_b + I_S \gamma_k(x_i)
\]  

(4.10)

with \(\gamma_k(x_i) = \gamma(x_{i1}, x_{2i} + (-1)^k x_0)\), for \(k = 0\) or \(1\).

where \(\gamma(x_{i1}, x_{2i})\) is the weighting function and \((0,x_0)\) or \((0,-x_0)\) is the center location of the image on the upper half or the lower half of the photosensitive surface, and \(I_b\) is the uniform intensity of the background light.

The optimum statistic \(g_R\) from the likelihood ratio can be expressed according to (4.3) with \(H_R(x)\) as the ratio of \(I_1(x)\) and \(I_0(x)\) defined by (4.10). \(g_R\) will be compared with the decision level \(g_{R0}\). Hypothesis \(H_1\) will be chosen if \(g_R > g_{R0}\). Otherwise hypothesis \(H_0\) will be chosen. The m.g.f. of the statistic \(g_R\) and its cumulants will be involved with double integration as discussed in the previous
paragraph. However, when the background light is negligible or absent so that \( I_b = 0 \), the distribution of the statistic can be derived directly from the inverse Fourier transform of the characteristic function when Gaussian images are postulated, and the performance of the ideal detector can be analyzed as follows.

The function \( H_R(x) \) described in (4.4) at \( I_b = 0 \) becomes

\[
H_R(x) = \frac{I_1(x)}{I_0(x)}
\]

\[
= \frac{\gamma_1(x)}{\gamma_0(x)}, \quad (4.11)
\]

where for the Gaussian images we have

\[
\gamma_k(x) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} \left[ x_1^2 + (x_2 + (-1)^k x_0)^2 \right]\right\}, \quad (4.12)
\]

for \( k = 0 \) or 1,

and the function \( H_R(x) \) from (4.11) becomes

\[
H_R(x) = \exp(2x_0 x_2/\sigma^2), \quad x_0 > 0. \quad (4.13)
\]

Thus the statistic \( g_R \) will be the sum of the numbers of photoelectrons ejected from each small area \( dA_i \), \( i = 1, 2 \ldots \) or \( N \), weighted by the coordinate \( x_{2i} \) of the center location \( (x_{1i}, x_{2i}) \) of that area. As we pass the limit \( dA_i \to 0 \), the m.g.f. of the optimum statistic can be
written as

\[
h_{Rk}(s) = \exp \left[ \alpha T I S \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx_1 dx_2 \gamma_k(x) (H_R(x))^{-s-1} \right] \\
= \exp \left[ N_s \left( \exp \left\{ \left( \frac{x_0}{\sigma} \right)^2 \right\} \left( (-1)^k + s \right) \right) - 1 \right],
\]

(4.14)

for \( k = 0, 1 \)

with \( N_s = \alpha T I S \) as the total number of photoelectrons emitted from the photosensitive surface during the observation interval \((0, T)\) because of the light from the source.

If we expand the right-hand side of (4.14) into a series of \( \exp \left\{ \left( \frac{x_0}{\sigma} \right)^2 \left( (-1)^k + s \right) \right\} \) we have

\[
h_{Rk}(s) = \exp(-N_s) \sum_{n=0}^{\infty} \frac{N_s^n}{n!} \exp \left[ -M_{nk} s + \frac{n s^2}{2} \right],
\]

(4.15)

for \( k = 0, 1 \)

where \( M_{nk} = -2n \left( \frac{x_0}{\sigma} \right)^2 (1)^k \) for \( n = 0, 1, 2, \ldots \)

\[
\sigma^2_n = 4n \left( \frac{x_0}{\sigma} \right)^2 \quad k = 0, 1
\]

(4.16)

The m.g.f. \( h_{Rk}(s) \) from (4.15) is a convergent power series where the \( n \)th term of the series is the m.g.f. of the Gaussian random variable with mean \( M_{nk} \) and variance \( \sigma^2_n \) given by (4.16) and is weighted by a
factor \( \exp(-N_s) N_s^n / n! \). The p.d.f. of the statistic \( g_R \) can be obtained by taking the inverse Fourier transform of its characteristic function and can be written as

\[
P_k(g_R) = \exp(-N_s) \delta(g_R) + \sum_{n=1}^{\infty} \frac{N_s^n}{n! \sqrt{2\pi} \sigma_n} \exp \left[ -\frac{1}{2\sigma_n^2} \left( \frac{g_R - M_{nk}}{\sigma_n} \right)^2 \right]
\]

for \( k=0,1 \) \hspace{1cm} (4.17)

where \( \delta(g_R) \) is the Dirac delta function\(^3\).

The p.d.f. of the statistic \( g_R \) has the delta function at the origin. The probability that there is no photoelectron ejected at all is \( \exp(-N_s) \).

The distribution of the statistic under hypothesis \( H_k \) for \( k=0 \) or \( 1 \) is

\[
F_k(g_{RO}) = P_r[g_R \leq g_{RO}] = 1 - \int_{g_{RO}}^{\infty} P_k(g_R) \, dg_R
\]

\[
= 1 - \exp(-N_s) \sum_{n=1}^{\infty} \frac{N_s^n}{n!} \text{erfc} \left( \frac{g_{RO} - M_{nk}}{\sigma_n} \right), \quad \text{for } g_{RO} \geq 0
\]

\[
= 1 - \exp(-N_s) \left[ 1 + \sum_{n=1}^{\infty} \frac{N_s^n}{n!} \text{erfc} \left( \frac{g_{RO} - M_{nk}}{\sigma_n} \right) \right],
\]

for \( g_{RO} < 0 \) \hspace{1cm} (4.19)

with \( \sigma_n^2 \) and \( M_{nk} \) given by (4.16) and the function \( \text{erfc}(x) \) defined by (2.44) in Chapter II.
The \( n \)th cumulant of the statistic can be also derived as follows,

\[
C_{Rk}^n = (-1)^n \frac{d^n}{ds^n} \ln h_{Rk}(s)|_{s=0}
\]

\[
= (-1)^n \frac{n}{\sigma} \left( \frac{2 \xi}{\sigma^2} \right)^n \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \exp \left[ -\frac{1}{2\sigma^2} (x_2 + (-1)^k x_0)^2 \right] x_2^n dx_2
\]

\[
= (-1)^n \frac{n}{\sigma^2} \left( \frac{2 \xi}{\sigma^2} \right)^n \mu_{nk}, \quad \text{for } k=0,1, \tag{4.20}
\]

where \( \mu_{nk} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \exp \left[ -\frac{1}{2\sigma^2} (x_2 + (-1)^k x_0)^2 \right] x_2^n dx_2 \)

\[
= \sigma^n E[X - (-1)^k \frac{x_0}{\sigma}]^n
\]

\[
= \sigma^n \sum_{r=0}^{n} \binom{n}{r} (-1)^{kr} \left( \frac{x_0}{\sigma} \right)^r E[X^{n-r}]
\]

is the \( n \)th central moment \( (4) \) of a zero mean Gaussian random variable \( X \) with unit variance and

\[
E[X^n] = 1 \cdot 2 \cdot 3 \ldots (n-1) \text{ for } n \text{ even}\]

\[
= 0 \quad \text{for } n \text{ odd}, \quad \binom{n}{r} = \frac{n!}{r!(n-r)!}
\]

The average error probability according to (2.6) is equal to the Bayes cost when the relative costs \( C_{10} - C_{00} \) and \( C_{01} - C_{11} \) are equal and

\( C_{01} = C_{10} = 1, C_{00} = C_{11} = 0 \). The decision level \( g_{RO} \) will be set at zero when \( \xi = \frac{1}{2} \), and the minimum error probability of the ideal
detector can be calculated from

\[ p_e = \frac{1}{2} Q_0 + \frac{1}{2}(1 - Q_d) \]

\[ = \frac{1}{2} \exp(-N_s) \left[ 1 + 2 \sum_{n=1}^{\infty} \frac{N_s^n}{n!} \text{erfc}(\sqrt{n} \frac{x_0}{\sigma}) \right] \quad (4.21) \]

for \( n=1,2,\ldots \).

At \( N_s = 0.1, 0.5, 1, 2, \text{and } 5 \), the error probabilities of the ideal detector have been calculated and plotted as a function of the ratio \( \frac{2x_0}{\sigma} \) by the solid curves in Fig. 19 where \( 2x_0 \) is the distance separating the images, which are centered at \((0,x_0)\) and \((0,-x_0)\) of the surface. \( \sigma \) is the width of the Gaussian image described in (4.12). The error probability \( p_e \) will be reduced as the distance \( 2x_0 \) between the two images increases. However, \( p_e \) will be limited by a value of \( \frac{1}{2} \exp(-N_s) \) as the ratio \( 2x_0/\sigma \) approaches to infinity. Therefore, the larger the value of \( N_s \), the better the performance of the detector will be.
Figure 19  Average error probability $P_e$ in binary bits detection vs. distance of separation ($2 x_0/\sigma$) between images for ideal detector, counting comparator, or simple detectors in the absence of background light; $N_e = 0.1, 1, 2, 5.$
Simple detector for detection of binary-bits

When background light also passes through the aperture and corrupts the object light, a simple way is to measure the total number of photoelectrons ejected from a finite area \( A' \) of the upper half photosensitive surface centered at \((0,x_0)\). The simple detector described in Chapter II thus can be used here to detect the binary bits. As \( TW >> 1 \), the number of photoelectrons ejected from the area \( A' \) has a Poisson distribution. The conditional probability under hypothesis \( H_k \) for \( k = 0 \) or \( 1 \) is given by (2.3) where the means can be determined from

\[
\bar{n}_{T1} = \alpha T \int \int_{A'} [I_b + I_S \gamma(x_1, x_2 - x_0)] \, dx_1 \, dx_2, \\
\bar{n}_{T0} = \alpha T \int \int_{A'} [I_b + I_S \gamma(x_1, x_2 + x_0)] \, dx_1 \, dx_2. 
\tag{4.22}
\]

For a Gaussian image, with \( \gamma_k(x) \) given in (4.12) and \( A' = \pi R_0^2 \sigma^2 \) is the circular area of the upper half surface where \( x_1^2 + (x_2 - x_0)^2 \leq R_0^2 \sigma^2 \). We have

\[
\bar{n}_{T1} = \alpha T \int \int_{A'} [I_b + I_S \frac{1}{2\pi \sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x_1^2 + (x_2 - x_0)^2)\right)] \, dx_1 \, dx_2 \\
= \alpha \pi R_0^2 \sigma^2 T_I + \alpha T \frac{I_S}{2\pi} \int_0^R \int_0^{2\pi} \exp\left(-\frac{R^2}{2}\right) \, RdRd\theta \\
= N_0 \left[ \frac{1}{2} R_0^2 + D^2 (1 - \exp(-R_0^2/2)) \right],
\]
\[ \bar{n}_{T0} = aT \int_{A} \left[ I_b + I_s \frac{1}{2n\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x_1 + (x_2 + x_0)^2)\right) \right] dx_1 \, dx_2 \]

\[ = a\pi \sigma^2 T I_b + a T \int_{0}^{2\pi} \int_{0}^{R_0} \exp\left[-\frac{R^2}{2} - \frac{2Rx_0}{\sigma} \cos\theta - \frac{2x_0^2}{\sigma^2}\right] R\,dR\,d\theta \]

\[ = N_0 \left[ \frac{1}{2} R_0^2 + D^2 (1 - Q\left(\frac{2x_0}{\sigma}, R_0\right)) \right], \quad (4.23) \]

where \(Q(a,b)\) is the "Q function" (5) for constants \(a\) and \(b\), and

\[ N_0 = 2\pi \sigma^2 a T I_b \]

\[ D^2 = \frac{N_0}{a T I_s} = \frac{N_0}{N_s} \]

\(N_0\) is the average number of photoelectrons ejected from an area \(2\pi \sigma^2\) during the interval \((0,T)\), \(D^2\) is the signal-to-noise ratio as defined in (2.57) and (2.58).

If \(A'\) is a square area, that is, we only consider the area

\[ A' = 4R_0^2 \sigma^2 \text{ for } |x_2 - x_0| \leq R_0\sigma, |x_1| \leq R_0\sigma \text{ in the upper half plane,} \]

the means can be calculated as

\[ \bar{n}_{T1} = a T \int_{-R_0^2}^{R_0^2} \int_{-R_0^2}^{R_0^2} \left\{ I_b + I_s \frac{1}{2n\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x_1^2 + (x_2 - x_0)^2)\right) \right\} dx_1 \, dx_2 \]
\[ R = 4R_0^2aTl_0 + aT \frac{I_S}{2\pi} \left[ \int_{-R_0}^{R_0} \exp\left(-\frac{R^2}{2}\right) dR \right]^2 \]

\[ N_0 \left[ \frac{2R_0^2}{\pi} + D^2(1 - 2 \text{erfc}(R_0))^2 \right] \]

\[ \bar{n}_{T0} = \alpha T \int_{-R_0}^{R_0} \sigma d\bar{x} \int_{x_0 - R_0}^{x_0 + R_0} d\bar{x}_2 \left[ I_{b1} + I_S \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2} \left( x_1^2 + (x_2 + x_0)^2 \right) \right] \right] \]

\[ N_0 \left[ \frac{2R_0^2}{\pi} + D^2(1 - 2 \text{erfc}(R_0))[\text{erfc}\left(\frac{2x_0}{\sigma} - R_0\right)] \] 

\[ - \text{erfc}\left(\frac{2x_0}{\sigma} + R_0\right) \right] \] \hspace{1cm} (4.24)

The observer will compare the number of photoelectrons \( n \) ejected from the area \( A' \) with a decision level \( g_{R0} \) such that \( H_1 \) will be chosen if \( n > g_{R0} \) and \( H_0 \) otherwise. The decision level can be determined according to (2.33) and

\[ g_{R0} = (\ln \Lambda_{s0} + \bar{n}_{T1} - \bar{n}_{T0})[\ln(\bar{n}_{T1}/\bar{n}_{T0})]^{-1}, \] \hspace{1cm} (4.25)

where \( \Lambda_{s0} \) is the decision level on the likelihood ratio and can be calculated according to (2.2) when the cost functions are given. For the Neyman Pearson strategy, randomization must be used, and \( g_{R0} \) is then an integer, which can be determined from (2.74), where \( Q_0 \) is preassigned. The reliability and the average error probability can be calculated by using (2.73) and (2.6). As the background light also passes through the aperture, there is a certain optimum size of the observation area.
at which for a fixed value of the false-alarm probability, the detection probability is maximum, as we have discussed in Chapter II.

The result and discussion would be the same if the observation area \( A' \) were taken from the lower half of the photosensitive surface and centered at \((0, -x_0)\). We only have to change the weighting function by

\[
\gamma_k(x) = \gamma(x_1, x_2 - (-1)^k x_0), \text{ for } k=0,1, \quad (4.26)
\]

where the mean numbers \( \bar{n}_{T0} \) and \( \bar{n}_{T1} \) observed from the lower half plane will be the same as given by (4.22), (4.23) or (4.24).

In order to compare with the ideal detector just discussed, we shall investigate the performance of the simple detector in the absence of background light. The observation area \( A' \) will be now taken as the whole upper (or lower) half plane. For a Gaussian image with \( \gamma_k(x) \) given by (4.12) the mean numbers can be written as

\[
\bar{n}_{T1} = N_s \int_{-\infty}^{\infty} dx_1 \int_{0}^{\infty} dx_2 \frac{1}{2\pi\sigma^2} \exp \left[ - \frac{1}{2\sigma^2} (x_1^2 + (x_2 - x_0)^2) \right] \]

\[
= N_s [1 - \text{erfc}(\frac{x_0}{\sigma})],
\]

\[
\bar{n}_{T0} = N_s \int_{-\infty}^{\infty} dx_1 \int_{0}^{\infty} dx_2 \frac{1}{2\pi\sigma^2} \exp \left[ - \frac{1}{2\sigma^2} (x_1^2 + (x_2 + x_0)^2) \right] \]

\[
= N_s \text{erfc}(\frac{x_0}{\sigma}), \quad (4.27)
\]
When the relative costs are equal and \( \xi = \frac{1}{2} \), the decision level is
\[
\varepsilon_{R_0} = (\bar{n}_{T1} - \bar{n}_{T0})[\ln(\bar{n}_{T1}/\bar{n}_{T0})]^{-1}
\]
for the Bayes strategy. The error probability \( P_e \) is calculated according to (2.77) and plotted as a function of the ratio \( (2x_0/\sigma) \) by the dashed curves in Fig. 19 at
\( N_s = 0.1, 0.5, 1, 2, \) and \( 5 \). As the ratio \( \frac{2x_0}{\sigma} \) approaches infinity, where \( \bar{n}_{T1} \rightarrow N_s \) and \( \bar{n}_{T0} \rightarrow 0 \), the error probability \( P_e \) will be also limited by \( \frac{1}{2} \exp(-N_s) \), which is half of the probability for \( n=0 \) with the average number \( N_s \).

The error probability \( P_e \) of the simple detector in the absence of the background light can be further reduced if the optimum observation area is considered. That is, instead of taking the whole upper half plane, the observation area will be moved up from the center line by a distance \( x_d \) such that the error probability according to (2.77) is minimum. In other words, for the Gaussian image we can take the proper integral range of \( x_2 \) from (4.27) as \( (x_d, \infty) \) instead of \( (0, \infty) \) as \( x_d > 0 \) so that the probability \( P_e \) can be further minimized. For \( N_s = 5 \), the typical values of \( x_d \) at different values of ratio \( (x_0/\sigma) \) are listed in table 2. The minimized \( P_e \) is plotted as a function of \( (2x_0/\sigma) \) in Fig. 20 where the error probabilities of the ideal detector and the simple detector at \( x_d=0 \) are also plotted.

When the background light also passes through the aperture, the observation area \( A' \) for Gaussian image discussed in (4.23) or (4.24) can be used. With the square area \( A' = 4\sigma^2 \) at \( R_0=1 \), the means are calculated from (4.24), and the error probabilities of the simple detector are plotted as a function of \( N_s \) for \( N_0 = 0, 0.1, 0.5, \) and
at $x_0/\sigma = 2$ in Fig. 21 by the dashed curves. In Fig. 22, the error probabilities are also plotted as a function of $N_s$ at $x_0/\sigma = 1$ and $x_0/\sigma = 2$ for the square area $A' = 4\sigma^2$ by the dashed curves. When the mean number $\bar{n}_{T_0}$ is fixed at 0.5, 1 or 2, the error probability of the simple detector is then plotted as a function of the number difference $\Delta N = \bar{n}_{T_1} - \bar{n}_{T_0}$ in Fig. 23 by the dashed curves.
Figure 20  Average error probability $P_e$ in binary bits detection vs. distance $(2 \times_0 / \sigma)$ for ideal detection and simple detector with fixed area or optimum area in the absence of the background light, $N_b = 5$. 


**Simple Detector**

- **Optimum area**

**Ideal Detector**

$N_s = 5$

$P_e$ vs $2x_0/\sigma$ graph:

- Simple Detector (fixed area)
- Simple Detector (optimum area)
- Ideal Detector
Figure 21  Average error probability $P_e$ in binary detection vs. average number $N_s$ for simple detector, counting comparator with square observation area $A' = 4\sigma^2$ at $R_0 = 1$ and $x_0/\sigma = 2$, $N_0 = 0, 0.1, 0.5, 1$. 
Simple Detector

Counting Comparator

$N_0 = 1$

$P_e$

$N_S$

$0.01$

$0.001$

$0$

$1$

$2$

$3$

$4$

$5$

$6$

$7$
Figure 22  Average error probability $P_e$ in binary detection vs. average number $N_s$ for simple detector and counting comparator with square observation area $A' = 4\sigma^2$ at $R_0 = 1$ and $N_0 = 0.1$; $x_0/\sigma = 1, 2$. 
Simple Detector
Counting Comparator

$N_0 = 0.1$

$\frac{x_0}{\sigma} = 1$

$\frac{x_0}{\sigma} = 2$

$P_e$

$N_s$

$0.01$

$0.001$

$0$ $1$ $2$ $3$ $4$ $5$ $6$ $7$
Figure 23  Average error probability $P_e$ in binary detection vs. average number difference $\Delta N$ for simple detector and counting comparator where $\Delta N = \bar{n}_{T1} - \bar{n}_{T0}$ with $\bar{n}_{Tk}$ as the mean number of photoelectrons observed from a finite area of one-half of the surface under hypothesis $H_k$ for $k=0, 1$; $\bar{n}_{T0} = 0.5, 1, 2$. 
Counting comparator for detection of binary bits

Another simple receiver for detecting the radiant source located at either \((0,x_0)\) or \((0,-x_0)\), similar to the receiver discussed by Peters and Arguello\(^{(2)}\), measures the total number \(n_u\) of photoelectrons emitted from the area \(A'\) centered at \((0,x_0)\) in the upper half photosensitive surface and compares it with the total number \(n_L\) of photoelectrons emitted from the area \(A''\) centered at \((0,-x_0)\) of the lower half plane \((A'=A'')\). The observer will choose \(H_1\) if \(n_u > n_L\) and \(H_0\) if \(n_u < n_L\). When \(n_u = n_L\) hypothesis \(H_1\) will be chosen with a probability \(f\). The false-alarm probability and the detection probability according to (2.9) and (2.10) can be written as

\[
Q_0 = \Pr[n_u > n_L | H_0] + f \Pr[n_u = n_L | H_0]
\]

\[
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_0(n+m) P_1(n) + f \sum_{n=0}^{\infty} P_0(n) P_1(n),
\]

\[
Q_d = \Pr[n_u > n_L | H_1] + f \Pr[n_u = n_L | H_1]
\]

\[
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_1(n+m) P_0(n) + f \sum_{n=0}^{\infty} P_1(n) P_0(n),
\]

where \(P_k(n) = \frac{n^n}{n!} e^{-\bar{n}_{Tk}}\), for \(k=0,1\),

\[\text{(4.28)}\]

with means \(\bar{n}_{T1}\) and \(\bar{n}_{T0}\) given in (4.22), (4.23), or (4.24).

(4.28) can be expressed in simpler form. Since
\[
\sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_0(n+m) P_1(n)
\]
\[
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \tilde{n}_{T_0}^{n+m} e^{-\tilde{n}_{T_0}} \tilde{n}_{T_1}^n e^{-\tilde{n}_{T_1}} / n! (m+n)!
\]
\[
= e^{-(\tilde{n}_{T_1} + \tilde{n}_{T_0})} \sum_{m=1}^{\infty} \left( \frac{\tilde{n}_{T_0}}{\tilde{n}_{T_1}} \right)^{m} I_m(ab)
\]
\[
= 1 - Q(a,b) \quad ,
\]

where \(I_m(x) = \sum_{n=0}^{\infty} (x/2)^{m+2n} / n! (n+m)!\)

is the modified Bessel function \(^{(6)}\) for \(m=0,1,2...\) and

\[
Q(a,b) = 1 - e^{-\frac{1}{2}(a^2 + b^2)} \sum_{m=1}^{\infty} \left( \frac{b}{a} \right)^m I_m(ab)
\]

is the "Q function" \(^{(5)}\) with

\[
a = \sqrt{2n_{T_1}}, \quad b = \sqrt{2n_{T_0}} .
\]

Also we can write

\[
\sum_{n=0}^{\infty} P_1(n) P_0(n) = \sum_{n=0}^{\infty} \left( \tilde{n}_{T_0} \tilde{n}_{T_1} \right)^n e^{-\left( \tilde{n}_{T_0} + \tilde{n}_{T_1} \right)/(n!)}^2
\]
\[
= -\frac{1}{2} (a^2 + b^2) e^{-I_0(ab)} .
\]
Therefore, for equal prior probabilities or $\xi = \frac{1}{2}$, the average error probability for making the decisions is

$$
P_e = \frac{1}{2} Q_0 + \frac{1}{2} (1 - Q_d)
$$

$$
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_0(n+m) P_1(n) + \frac{1}{2} \sum_{n=0}^{\infty} P_0(n) P_1(n)
$$

$$
= 1 - Q(a, b) + \frac{1}{2} \exp\left(-\frac{1}{2}(a^2 + b^2)\right) I_0(ab).
$$

(4.32)

The error probability $P_e$ from (4.32) of the counting comparator has been evaluated in the absence of the background light for the Gaussian images with the mean numbers under $H_1$ and $H_0$ as given by (4.27). The error probabilities $P_e$ of this receiver are plotted as a function of the ratio $(2x_0/\sigma)$ at $N_s = 0.1, 0.5, 1, 2, \text{and } 5$ in Fig. 19. The error probability decreases as the ratio $(2x_0/\sigma)$ increases. As $2x_0/\sigma$ approaches infinity, the error probability will be limited by $\frac{1}{2} \exp(-N_s)$.

When the background light also passes through the aperture, the error probabilities of this receiver are also calculated with a square observation area $A' = 4\sigma^2$ with $R_0 = 1$ and plotted as a function of $N_s$ in Fig. 21 by the solid curves at $N_0 = 0, 0.1, 0.5, \text{and } 1$ for $x_0/\sigma = 2$. In Fig. 22, the error probabilities are plotted as a function of $N_s$ when the same observation area $A' = 4\sigma^2$ is used and the average background number $N_0$ is fixed at 0.1 for $x_0/\sigma = 1$ and 2. In Fig. 23 the error probabilities are plotted as a function of the number
difference $\Delta N$ by the solid curves at fixed values of average number $\overline{n}_{T0}$, 0.5, 1, and 2. As the size of the square area $A'$ is varied, the error probability $P_e$ changes. The error probabilities of the counting comparator are plotted as a function of the length $R_0$ in Fig. 24 at different values of $N_0$, 0.1, 0.5, and 1.
Figure 24  Error probability $P_e$ of the counting comparator vs. the length $R_0$ for the rectangular observation area; $N_s = 5$. The curves are indexed by the parameter $N_0$ defined in (2.57).
\( N_s = 5 \)
\( \sigma_0 / \sigma = 3.5 \)

![Graph showing the relationship between \( P_e \) and \( R_0 \) for different values of \( N_0 \).]
The investigations of the detectors of binary bits that we have just discussed are based on the primary assumption that we know that the image is located at either \((0, x_0)\) or \((0, -x_0)\) exactly. We would like to investigate the performances of those detectors when the distance between the image is still \(2x_0\), but the images are mislocated with a shift of \(\epsilon\). That is the images are actually located at \((0, x_0')\) and \((0, -x_0'')\) under hypothesis \(H_1\) and \(H_0\) where \(x_0' = x_0 + \epsilon, x_0'' = x_0 - \epsilon\) or \(x_0' = x_0 - \epsilon, x_0'' = x_0 + \epsilon\) for \(\epsilon > 0\).

When the ideal detector is used, the expression of the function \(H_R(x_i)\) from (4.13) can still be used, where the weighting function \(\gamma_k(x_i)\) from (4.12) will be replaced with \(x_0 = x_0'\) for \(k = 1\) and \(x_0 = x_0''\) for \(k = 0\). We can derive the m.g.f. of the optimum statistic \(g_R\) following (4.14), and the expression of the m.g.f. \(h_{Rk}(s)\) from (4.15) can still be used with \(M_{nk}\) now replaced by

\[
M_{nk} = -2n \frac{x_0 x_0'}{\sigma^2} (-1)^k, \text{ for } k=1
\]

\[
= -2n \frac{x_0 x_0''}{\sigma^2} (-1)^k, \text{ for } k=0, n=0,1,2... \quad (4.33)
\]

and \(\sigma_n^2\) unchanged.

The false-alarm probability and detection probability with zero decision level can be expressed according to (4.19) as
\[ Q_0 = P_r\{g_R > 0|H_0\} \]
\[ = \exp(-N_S) \sum_{n=1}^{\infty} \frac{N_S^n}{n!} \operatorname{erfc}\left(\sqrt{n} \frac{x_0^n}{\sigma}\right) , \]
\[ Q_d = P_r\{g_R > 0|H_1\} \]
\[ = \exp(-N_S) \sum_{n=1}^{\infty} \frac{N_S^n}{n!} (1 - \operatorname{erfc}\left(\sqrt{n} \frac{x_0}{\sigma}\right)) . \quad (4.34) \]

The average error probability for equal prior probability or \( \xi = \frac{1}{2} \) can be written as

\[ P_e = \frac{1}{2} Q_0 + \frac{1}{2}(1 - Q_d) \]
\[ = \frac{1}{2} \exp(-N_S)[1 + \sum_{n=1}^{\infty} \frac{N_S^n}{n!} (\operatorname{erfc}\left(\sqrt{n} \frac{x_0^n}{\sigma}\right) + \operatorname{erfc}\left(\sqrt{n} \frac{x_0}{\sigma}\right))] . \quad (4.35) \]

The error probability will be the same for either \( x_0' = x_0 + \epsilon \)
\( x_0'' = x_0 - \epsilon \) or \( x_0' = x_0 - \epsilon \), \( x_0'' = x_0 + \epsilon \). That is, (4.35) is the average error probability of the ideal detector when the images are mislocated with a shift either \( \epsilon \) or \(-\epsilon\).

When the counting comparator is used, with the distance between the images fixed at \( 2x_0 \), the mean numbers due to a shift \( \epsilon \) or \(-\epsilon\) of the actual image can be expressed as
\[
\bar{n}_{T1\pm} = N_s [1 - \text{erfc} \left( \frac{x_0 \pm \varepsilon}{\sigma} \right)],
\]
\[
\bar{n}_{T0\pm} = N_s \text{erfc} \left( \frac{x_0 \pm \varepsilon}{\sigma} \right),
\]

(4.36)

where \( \bar{n}_{T1\pm}, \bar{n}_{T0\pm} \) are the means for a positive shift \( \varepsilon \) and \( \bar{n}_{T1-}, \bar{n}_{T0-} \) for a negative shift \( -\varepsilon \). The false-alarm probability and the detection probability of the counting comparator when the images are mislocated at \((0,x_0)\) and \((0,-x_0)\) under hypothesis \( H_1 \) and \( H_0 \), while actually located at \((0,x_0+\varepsilon)\) and \((0,-x_0+\varepsilon)\) or \((0,x_0-\varepsilon)\) and \((0,-x_0-\varepsilon)\), can be written as

\[
Q_{0\pm} = \Pr[n_u > n_L | H_0] + f \Pr[n_u = n_L | H_0]
\]
\[
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_{1\pm}(n) P_{0\pm}(n+m) + f \sum_{n=0}^{\infty} P_{1\pm}(n) P_{0\pm}(n),
\]

\[
Q_{d\pm} = \Pr[n_u > n_L | H_1] + f \Pr[n_u = n_L | H_1]
\]
\[
= 1 - \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P_{1\pm}(n) P_{0\pm}(n+m) + f \sum_{n=0}^{\infty} P_{1\pm}(n) P_{0\pm}(n),
\]

(4.37)

where \( P_{k\pm}(n) = \frac{\bar{n}_{Tk\pm}}{n!}, \) for \( k=0,1 \).

\( Q_{0\pm} \) and \( Q_{d\pm} \) are the false-alarm probability and the detection
probability for a positive shift $e$ while $Q_{0-}$, $Q_{d-}$ are the probabilities for a negative shift $-e$.

For equal prior probability $\xi = \frac{1}{2}$ and $f = \frac{1}{2}$, the average error probability will be the same for either a shift $e$ or $-e$ and can be written as

$$P_{e\pm} = \frac{1}{2} Q_{0\pm} + \frac{1}{2} (1 - Q_{d\pm})$$

$$= \frac{1}{2} \left[ \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_{1\pm}(n) P_{0\pm}(n+m) + \frac{1}{2} \sum_{n=0}^{\infty} P_{1\pm}(n) P_{0\pm}(n) \right]$$

$$+ \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} P_{1\pm}(n) P_{0\pm}(n+m) + \frac{1}{2} \sum_{n=0}^{\infty} P_{1\pm}(n) P_{0\pm}(n) \right]$$

$$= 1 - \frac{1}{2} \left( Q(a_-^+, b_-^+) + Q(a_+^-, b_-^-) \right) - \frac{1}{2} \left( a_-^2 + b_-^2 \right) I_0(a_-^+ b_-^-)$$

$$+ e^{- \frac{1}{2} (a_+^2 + b_-^2) I_0(a_+^+ b_-^-)} \right], \tag{4.38}$$

where $a_\pm = \sqrt{2n_{1\pm}}$, $b_\pm = \sqrt{2n_{0\pm}}$ and $Q(a,b)$, $I_0(ab)$ are given in (4.29).

$P_{e\pm}$ is the error probability for a shift of $e$ and $P_{e\pm}$ is the error probability for a shift of $-e$. $P_{e\pm} = P_{e\pm}$ as we can see from (4.38).

The error probabilities of both the ideal detector and the counting comparator are calculated at different values of shift $e$ when (4.35) and (4.38) are used. In Fig. 25, the error probabilities of these two detectors are plotted as a function of the ratio $|e|/x_0$ at $N_s = 2, 5$ and $8$ with $x_0/\sigma = 1$. In Fig. 26 the error probabilities are plotted as a function of the ratio $|e|/x_0$ at $x_0/\sigma = 1.5, 2$ and $4$ with fixed value of $N_s$ at $5$. 
Figure 25  Average error probability $P_e$ in binary bits detection vs. the ratio $|e|/x_0$ due to the mislocation of the images for the ideal detector or the counting comparator at $x_0/\sigma = 1$; $N_s = 2, 5, 8$. 
Figure 26 Average error probability $P_e$ in binary bits detection vs. the ratio $|c|/x_0$ due to the mislocation of the images for the ideal detector and the counting comparator for fixed value $N_g = 5; x_0/\sigma = 1.5, 2, 4$. 
$N_s = 5$

- Ideal Detector
- Counting Comparator

$x_0 / \sigma = 1.5$

$x_0 / \sigma = 2$

$x_0 / \sigma = 4$

$P_e$

$|\epsilon| / x_0$
Discussion

In the absence of uniform background light, registering the locations of the ejected photoelectrons for the detection of either one of the two Gaussian images, as we have discussed in this chapter, not only utilizes more of the information, but also effectively reduces the mean value of the optimum statistic under $H_0$ and thus improves the performance of the ideal detector. As for the simple detector, the gain by optimizing the observation area will be limited because varying the area will affect both the mean numbers $\bar{n}_T^0$ and $\bar{n}_T^1$ as defined in (4.27). Therefore, the performance of the ideal detector is better than that of the simple detector, as we can see from Fig. 19 and Fig. 20. When the counting comparator is used, the observed numbers of the photoelectrons ejected from both the upper half surface and the lower half surface will be used, while for the simple detector only the number of photoelectrons ejected from the upper half surface (or the lower half surface) will be utilized for making decision. In other words, the counting comparator utilizes more information than the simple detector, but less information than the ideal detector. Thus the performance is better than that of the simple detector, but not as good as that of the ideal detector as shown in Fig. 19.

When background light also passes through the aperture, the average error probability for making decisions will be increased. In the meantime, the observer must define a finite observation area when any one of the detectors will be used. For the simple detector or the counting comparator, the finite observation area can be chosen such
that for a fixed false-alarm probability, the detection probability of
the detector will be as large as possible. The proper size of the
square area for the counting comparator to be used at $0.1 \leq N_0 \leq 1$ is
$R_0 = 2$ or $A' = 16\sigma^2$ as we can see from Fig. 24. For a finite square
area $4\sigma^2$, the average error probabilities for both simple detector and
counting comparator will increase as the background light increases or
the ratio $x_0/\sigma$ decreases as shown in Fig. 21, Fig. 22, and Fig. 23.

When Gaussian images are postulated, the performance of the
detectors for detecting binary bits will also depend on the prior
knowledge about the locations of the images. The average probability
of making errors in each decision will be increased if the images for
bit "1" and bit "0" are mislocated. The severity of mislocation will
also depend on the distance $2x_0$ separating the images. For example,
when the ratio $x_0/\sigma = 4$, the effect on either the ideal detector or
the counting comparator due to mislocation will be insignificant until
the ratio of shift $|\varepsilon|/x_0$ is greater than 1.2, as we can see from Fig.
26. Furthermore, the effect on the counting comparator of mislocation
of the images will be little smaller than that on the ideal detector
as shown in Fig. 25 or Fig. 26 and at $x_0/\sigma = 1.5$; for example, the
average error probability of the counting comparator becomes smaller
than that of the ideal detector when $|\varepsilon|/x_0$ is greater than 1.6, as
shown in Figure 26,
Table 2 - Optimum Distance for Simple Detector

<table>
<thead>
<tr>
<th>$x_0/\sigma$</th>
<th>0.1</th>
<th>0.5</th>
<th>0.75</th>
<th>1.0</th>
<th>1.25</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_d/\sigma$</td>
<td>0.16</td>
<td>0.52</td>
<td>1.09</td>
<td>1.1</td>
<td>1.1</td>
<td>1.08</td>
<td>1.04</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Footnotes
Chapter IV

1. Helstrom [27].

2. Peters and Arguello [28].

3. Papoulis [29], p. 155, eq. (5.77).

4. Papoulis [29], p. 146.

5. Helstrom [10], Appendix F.

References


APPENDIX A

Asymptotic Expansion Approximation of the Tail Distribution by the Method of Steepest Descent

In detection theory, the evaluation of the cumulative distribution for a statistic is often important, and it is required to reach a certain accuracy. For example, when the Neyman-Pearson criterion is used, the decision level of the likelihood-ratio detector must be calculated from the preassigned false-alarm probability $Q_0$. As $Q_0$ is small, there may be a serious error in the decision level if the cumulative distribution cannot be accurately calculated. The statistic we will discuss in this Appendix is of the type of the sum of $N$ independent random variables and can be written in general as

$$Z = \sum_{i=1}^{N} X_i \quad i = 1, 2, \ldots, N.$$  \hspace{1cm} (A1)

When $\{X_i\}$ are identical random variables and $N$ is large, the central limit theorem (1),(2) can be applied. If $\{X_i\}$ are independent but not identical, the distribution of $Z$ may be dominated by one of the random variables and can be approximated by the Gaussian distribution when conditions such as Lyapunov's condition (3) are satisfied. On the other hand, the distribution of the statistic $Z$ can be derived from the inverse Fourier transform of its characteristic function when the inverse integral can be approximated by the asymptotic expansion series through the steepest-descent method. For detailed treatment of the steepest-descent approximation we refer to the works of many
authors (4-7). For applications of the steepest-descent method in statistics we refer to the works of the authors (8-10). Here we will discuss the asymptotic expansion approximation of the integral with only one saddle point.

The characteristic function (11) of the random variable $Z$ is defined by

$$h(\omega) = \mathbb{E}[e^{i\omega Z}] = \int_{-\infty}^{\infty} e^{-i\omega Z} dF(Z),$$  \hspace{1cm} (A2)

where $F(Z) = \Pr[Z \leq Z_0]$ is the distribution of the random variable $Z$.

The probability that $Z$ exceeds a value $Z_0$ can be written as

$$Q(Z_0) = \Pr[Z > Z_0] = 1 - \int_{-\infty}^{\infty} U(Z_0 - Z) dF(Z),$$  \hspace{1cm} (A3)

where $U(x)$ is the step function.

If we take the Fourier transform of $Q(Z_0)$ by discarding the portion of the integral which oscillates with infinite rapidity, we have

$$\mathbb{F}[Q(Z_0)] = \int_{-\infty}^{\infty} e^{-i\omega Z_0} Q(Z_0) dZ = 2\pi \delta(\omega) - \frac{1}{i\omega} h(\omega),$$  \hspace{1cm} (A4)

where $\delta(\omega)$ is the Dirac delta function.

The probability $Q(Z_0)$ can then be expressed in terms of $h(\omega)$ by taking the inverse Fourier transform of (A4). That is
\[ Q(Z_0) = \mathcal{F}^{-1}\mathcal{F}[Q(Z_0)] = 1 - \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega Z_0} \frac{h(-\omega)}{i\omega} \, d\omega \]  

(A5)

where \( \mathcal{F}, \mathcal{F}^{-1} \) are the Fourier transform and inverse Fourier transform operators respectively.

When the integral on the right-hand side of (A5) cannot be evaluated exactly, the method of steepest-descent can be used to approximate the tail probability \( Q(Z_0) \). The integral from (A5) can be extended to a contour integral as

\[ Q(Z_0) = 1 - \frac{1}{2\pi i} \oint_{c_1} \frac{1}{s} \exp[Z_0 \phi(s)] \, ds. \]  

(A6)

The contour \( c_1 \) of the integration is a straight line running from \( \alpha_0 - i\omega \) to \( \alpha_0 + i\omega \) for \( \alpha_0 > 0 \) in the complex plane \( s \) with real values of \( Z_0 \). Here

\[ \phi(s) = s + Z_0^{-1} \ln h(s) \]  

(A7)

is the complex phase of the integral, and \( h(s) = E[e^{-sZ}] \) is the m.g.f. of the random variable \( Z \) as defined by (A2) with \( s = i\omega \). The integral in (A6) can be approximated by taking the integration path on the complex plane so that the imaginary part of \( \phi(s) \) is constant (path of the steepest descent) along the path, which also passes through the real saddle point \( s_0 \) upon which the modulus of \( e^{Z_0 \phi(s)} \) decreases most rapidly. The saddle point \( s_0 \) can be determined from the equation
Since the origin is also a singular point of the integral, the contributions from taking the integral path through the saddle point $s_0$ and around the singularity $s = 0$ cannot be treated separately when $s_0$ is small or near to the origin. Rice\textsuperscript{(9)} presented a more general discussion for cases involving more than one saddle point. In our case the complex phase $\phi(s)$ has only one saddle point as discussed in Appendix B. If $\phi(s)$ is analytic in the neighborhood of the origin, $\phi(s)$ behaves much like a second-order polynomial in $s$. This suggests changing the variable of the integral from $s$ to $u$ in such a way that

\[
\varphi(u) = u^2 - 2u_0 u = \phi(s) \tag{A9}
\]

is the new complex phase\textsuperscript{(12)} in the $u$ plane with $u_0$ as the new saddle point corresponding to $s_0$ in the $s$-plane, and $\varphi(u_0) = \phi(s_0)$. The new saddle point $u_0$ can be determined from (A9) as

\[
u_0 = \text{sign}(s_0) \left(-\phi(s_0)\right)^{1/4}, \tag{A10}\]

where the function sign($x$) = 1 for $x > 0$, sign($x$) = -1 for $x < 0$, so that $u_0$ and $s_0$ have the same sign. $\phi(s_0)$ is the complex phase at $s = s_0$ and is assumed to be non-positive. Thus, integral on the right-hand side of (A6) can be expanded into a uniform asymptotic
series of powers $z_0^{-1}$ as

$$\int_{c_1}^{1} \frac{1}{s} \exp[Z_0 \phi(s)] ds$$

$$= \int_{c_1'}^{1} \frac{1}{u} f(u) \exp[Z_0 (u^2 - 2 u_0 u)] du$$

$$= k_0(Z_0) + k_1(Z_0) \sum_{n=1}^{\infty} A_n Z_0^{-n}, \quad (A11)$$

where

$$f(u) = u ds/sdu,$$

$$k_0(Z_0) = \int_{c_1'}^{1} u^{-1} \exp[Z_0 (u^2 - 2 u_0 u)] du,$$

$$k_1(Z_0) = \int_{c_1'}^{1} \exp[Z_0 (u^2 - 2 u_0 u)] du,$$

with $c_1'$ running from $\alpha' - i\infty$ to $\alpha' + i\infty$ for $\alpha' > 0$ in the complex $u$ plane. $A_n$ is the coefficient of the $n$th term in the expansion and can be determined from the Ursell method as discussed by Rice (13).

Thus, the probability from (A6) can be approximated by the uniform asymptotic expansion series as

$$Q(Z_0) = 1 - E(Z_0) - I(Z_0), \quad (A12)$$

where
\[ E(Z_0) = \begin{cases} 1 - \text{erfc} \left[ (-2 Z_0 \phi(Z_0))^2 \right], & \text{for } s_0 < 0 \\ \text{erfc} \left[ (-2 Z_0 \phi(Z_0))^2 \right], & \text{for } s_0 > 0 \end{cases} \]

and

\[
I(Z_0) = \frac{\exp[Z_0 \phi(s_0)]}{[2\pi Z_0 \phi^{(2)}(s_0)]^{1/2}} \sum_{k=0}^{\infty} \sum_{n=0}^{k} \left\{ \left( \frac{-2}{Z_0 \phi^{(2)}(s_0)} \right)^{k} \right\}^{s_0-1} \\
\times \sum_{n=0}^{2k} (-s_0)^{-2k+n} \sum_{\ell=0}^{n} A_{\ell, n} \left( \frac{1}{2} \right)^{n-k+\ell} \\
\times \text{sign}(s_0) \left( \frac{1}{2} \right)^{k} \left( \frac{-\phi^{(2)}(s_0)}{2 \phi(s_0)} \right)^{1/2} (Z_0 \phi(s_0))^{-k}) \}
\]

where

\[
\text{erfc} y = \frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} \exp(-x^2/2) \, dx
\]

and

\[
A_{\ell, n} = \begin{cases} 0, & \text{for } n < \ell \text{ or } \ell = 0, n > 1 \\ 1, & \text{for } n = \ell = 0 \end{cases}
\]

\[
A_{\ell+1, n+1} = \frac{-2}{n+1} \sum_{m=1}^{n-k+1} \frac{m \phi^{(m+n)}(s_0)}{(m+2) \phi^{(2)}(s_0)} A_{\ell, n-m+1}
\]

\[
\left( \frac{1}{2} \right)_m = \frac{1}{2} \left( \frac{1}{2} + 1 \right) \left( \frac{1}{2} + 2 \right) \cdots \left( \frac{1}{2} + m - 1 \right)
\]

A_{\ell, n} are the coefficients which can be calculated by the recurrence relation through the nth derivative \( \phi^{(n)}(s_0) \) of the complex phase, where
\[ \phi^n(s) = z_0^{-1} \left( \frac{d}{ds} \right)^n \ln h(s), \text{ for } n \geq 2, \quad (A13) \]

and is evaluated at the saddle point \( s = s_0 \). This scheme can be easily programmed for a digital computer. The tail distribution \( Q(Z_0) \) is obtained by adding up the terms in the asymptotic expansion given by (A12) until they become insignificantly small or they stop decreasing and begin to increase.

When \( Z \) is a discrete random variable, the tail distribution can be written as

\[ Q(Z_0) = \sum_{Z > Z_0}^{\infty} p(Z), \quad (A14) \]

where \( p(Z) \) is the probability of the random variable \( Z \) and the m.g.f. can be written as

\[ h_d(s) = E[e^{-sZ}] = \sum_{Z = 0}^{\infty} p(Z) e^{-sZ}. \quad (A15) \]

The calculation of the tail distribution is simpler and more accurate if one first approximates the probability \( p(Z) \) and then adds up the probabilities for all \( Z > Z_0 \) as given by (A14). For example, when \( Z \) takes only integral values, we can express the probability by the contour integral and approximate it by the uniform asymptotic series as
\[ p(Z) = \frac{1}{2\pi i} \int_{c_d} e^{Z \phi_d(s)} ds \]

\[ \approx \frac{\exp[Z \phi_d(s_0)]}{[2\pi Z \phi_d''(s_0)]^{1/2}} \sum_{m=0}^{\infty} \left( \frac{-2}{Z \phi_d''(s_0)} \right)^m A_{l,2m} \left( \frac{1}{2} \right)^{l+m}, \quad (A16) \]

where \( c_d \) is the contour running from \( a'' - i\pi \) to \( a'' + i\pi \), \( a'' > 0 \) and

\[ \phi_d(s) = \left( \ln h_d(s) + s \right)^{-1}. \quad (A17) \]

The saddle point \( s_0 \), the coefficients, \( A_{l,2m} \) and the derivative of the complex phase \( \phi_d(s) \) can be determined as before.

To illustrate the applications of the asymptotic-expansion approximation, three well-known distributions, Gaussian, gamma, and Poisson will be discussed as follows:

**Gaussian distribution**

The p.d.f. of Gaussian distribution is given by

\[ p(Z) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left[ -\frac{1}{2\sigma^2} (Z - m)^2 \right], \quad (A18) \]

with \( m, \sigma \) as the mean and variance. The m.g.f. \( h(s) \) is given by

\[ h(s) = \exp\left[ -s m + \frac{\sigma^2 s^2}{2} \right]. \quad (A19) \]
The saddle point can be determined from \( \frac{d}{ds} \phi(s) = 0 \) or by (A8). We have

\[
s_0 = (m - Z_0)/\sigma^2. \tag{A20}
\]

The complex phase and its nth derivatives evaluated at \( s_0 \) are

\[
\phi(s_0) = Z_0^{-1} \ln h(s_0) + s_0 = Z_0^{-1}(m - Z_0)^2/2\sigma^2,
\]

\[
\phi''(s_0) = \sigma^2/Z_0,
\]

\[
\phi'(s_0) = \phi^n(s_0) = 0, \text{ for } n \geq 2. \tag{A21}
\]

Thus, the coefficient \( A_{x,n} = 0 \) for all \( x \), \( n \neq 0 \) and \( I(g_0) \) vanishes; also, the tail distribution becomes

\[
Q(Z_0) = 1 - E(Z_0) = \text{erfc} \left( \frac{m - Z_0}{\sigma} \right), \text{ for } s_0 < 0
\]

\[
= 1 - \text{erfc} \left( \frac{m - Z_0}{\sigma} \right), \text{ for } s_0 > 0 \tag{A22}
\]

as it should be for Gaussian distribution.

**Gamma distribution**

The p.d.f. of a gamma distribution is given by
\[ p(Z) = \frac{\lambda}{\Gamma(\beta)} (\lambda Z)^{\beta-1} e^{-\lambda Z}, \quad \beta > 0, \lambda > 0, Z > 0 \]
\[ = 0 \quad \text{otherwise}. \quad (A23) \]

The m.g.f. (15) is given by
\[ h(s) = \left(1 + \frac{s}{\lambda}\right)^{-\beta}. \quad (A24) \]

The saddle point \( s_0 \) is determined from (A8) or
\[ s_0 = Z_0^{-1} \beta - \lambda, \quad Z_0 > 0. \quad (A25) \]

The complex phase and its nth derivative evaluated at \( s_0 \) are
\[ \phi(s_0) = Z_0^{-1} \left[-\beta \ln(\beta/Z_0\lambda) + \beta - \lambda Z_0\right], \]
\[ \phi^n(s_0) = (-1)^n (n-1)! \left(Z_0/\beta\right)^{n-1}, \quad \text{for } n \geq 2. \quad (A26) \]

The special case at \( \lambda = 1, \beta = 1 \) is the exponential distribution where
\[ p(Z) = e^{-Z}, \quad Z > 0 \]
\[ = 0 \quad \text{otherwise}. \quad (A27) \]

The complex phase and its derivatives of the exponential distribution evaluated at the saddle point from (A25) and (A26) are
\[ s_0 = z_0^{-1} - 1 , \]
\[ \phi(s_0) = z_0^{-1} (1 + \ln z_0) - 1 , \]
\[ \phi^n(s_0) = (-1)^n z_0^{n-1} (n - 1)! \text{ for } n \geq 2. \] (A28)

The numerical calculation of the tail distribution by using the asymptotic expansion series given in (A12) for the exponential distribution has been carried out and compared with the exact value. The percentage of relative error is listed in Table 3.

**Poisson distribution**

\[ p(z) = e^{-\lambda} \lambda^{z/z!} , \quad z = 0, 1, \ldots \]
\[ = 0 \quad \text{otherwise} . \] (A29)

The m.g.f. is

\[ h_d(s) = \exp[\lambda(e^{-s} - 1)] . \] (A30)

The saddle point \( s_0 \) is solved from (A8) or

\[ s_0 = \ln (\lambda/z) . \] (A31)

The complex phase and its nth derivatives evaluated at \( s_0 \) are
\[ \phi(s_0) = 1 - \lambda Z^{-1} + \ln(\lambda Z^{-1}) , \]

\[ \phi^n(s_0) = (-1)^n, \text{ for } n \geq 2. \quad (A32) \]

The tail distribution has been evaluated by first approximating the probability \( p(Z) \) with the asymptotic expansion series given in (A16) and then adding up the probabilities for all \( Z > Z_0 \). The relative error by comparing with the exact value is listed in Table 4.
### Table 3 - Percentage Error of Exponential Distribution

<table>
<thead>
<tr>
<th>$Z_0$</th>
<th>0.05</th>
<th>0.3</th>
<th>0.6</th>
<th>0.7</th>
<th>2.5</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>0.0072</td>
<td>0.0194</td>
<td>0.3243</td>
<td>0.3310</td>
<td>-0.0174</td>
<td>-0.0866</td>
<td>-0.0866</td>
<td>0.2039</td>
<td>0.6614</td>
</tr>
</tbody>
</table>

### Table 4 - Percentage Error of Poisson Distribution

<table>
<thead>
<tr>
<th>$Z_0$</th>
<th>6</th>
<th>10</th>
<th>14</th>
<th>18</th>
<th>22</th>
<th>26</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>-0.0061</td>
<td>-0.0062</td>
<td>-0.0087</td>
<td>-0.0049</td>
<td>-0.0043</td>
<td>-0.0058</td>
<td>-0.0045</td>
<td>-0.0028</td>
</tr>
</tbody>
</table>
Footnotes

2. Cramér [31], Chapter VI.
4. Erdélyi [33], Chapter I, II.
5. Jeffreys [34], Chapter I, II.
6. Ursell [35].
7. Bleistein [36].
8. Daniels [8].
9. Rice [7].
10. Mazo and Salz [37].
11. Lukacs [38], Chapter I, eq (1.36).
12. Rice [7], Appendix B.
14. Daniels [8], § 8.
15. Lukacs [38], Chapter I.
Appendix B

Iteration Procedure for Finding the Saddle Point in the Asymptotic Expansion Series

In this Appendix we will show that there is only one real saddle point of the contour integral given by (A6) in Appendix A for approximating the tail distribution of the optimum statistic $g$ or the threshold statistic $g_0$ from (2.18) or (2.26). The formulas used for searching for the saddle point when the Neyman-Pearson criterion is applied will be given.

The cumulant-generating function for the ideal detector is given by (2.20) from Chapter II as

$$c_k(s) = \ln h_k(s) = a T I_b \int_A h^k(x)[(H(x))^{-s}-1] d^2 x, \quad k = 0, 1.$$  \hspace{1cm} (B1)

The m.g.f. $M(s, g_0)$ of the positive statistic $g$ is defined by

$$M(s, g_0) = \mathbb{E}[e^{-s(g-g_0)}] = \int_{-\infty}^{\infty} e^{-s(g-g_0)} dF_k(g) = \exp[g_0 s + c_k(s)] = \exp[g_0 \phi_k(s)], \quad k = 0, 1.$$  \hspace{1cm} (B2)

and can be expressed as a function of $\phi_k(s)$ as given in (2.48).

198
As we investigate the complex phase \( \phi_k(s) \) for \( k = 0, 1 \) and its first and second derivatives from (2.53)

\[
\begin{align*}
\phi_k'(s) &= -\varepsilon_0^{-1} a T I_b \int_A [H(\chi)]^{k-s} \ln H(\chi) d^2\chi + 1, \\
\phi_k''(s) &= \varepsilon_0^{-1} a T I_b \int_A [H(\chi)]^{k-s} \ln^2 H(\chi) d^2\chi,
\end{align*}
\]

\( k = 0, 1, \)

we have \( \phi_k'(-\infty) = -\infty, \phi_k'(\infty) = 1 \) and \( \phi_k''(s) > 0 \) for all real values of \( s \). Then \( \phi_k(s) \) is a non-decreasing function of real \( s \). Furthermore, the cumulant-generating function \( C_k(s) \) exists in \( -\infty < s < \infty \).

Therefore, according to the theorems 6.1 and 6.2 of Daniels(1), there is only one real root, which is simple, of \( \phi_k(s) = 0 \) for \( \varepsilon_0 > 0 \) and no more real roots otherwise. This is true also for the threshold statistic \( \varepsilon_0^* \).

When the Neyman-Pearson strategy is used, the false-alarm probability \( Q_0 \) is preassigned. To search for the saddle point \( s_0 \) under \( H_0 \), we must arbitrarily pick the initial values \( s_{ON} \) and \( s_{OP} \) in such a way that the false-alarm probability calculated at \( s_{ON} \) or \( s_{OP} \) by using the asymptotic series from (A12) will be \( Q_{ON} \) or \( Q_{OP} \) and

\[
Q_{ON} < Q_0 < Q_{OP}.
\]
A trial value \( s_0^T \) of the saddle point then will be determined from the equation

\[
s_0^T = s_{ON} + \frac{s_{OP} - s_{ON}}{Q_{OP} - Q_{ON}} (Q_0 - Q_{ON}) . \tag{B3}
\]

The trial value of the decision level \( g_0^T \) and probability \( Q_0^T \) can be calculated by substituting \( s_0^T \) into (2.62) and (2.55) for the ideal detector or (2.69) and (2.55) for the threshold detector. If \( Q_0^T > Q_0 \), the values of \( s_{OP} \) and \( Q_{OP} \) will be replaced by the values \( s_0^T \) and \( Q_0^T \). If \( Q_0^T < Q_0 \), the values of \( s_{ON} \) and \( Q_{ON} \) will be replaced instead. Thus, a new trial value \( s_0^T \) will be obtained according to (B3), and the procedure will be repeated until the calculated values \( Q_0^T \) approaches the preassigned value \( Q_0 \) within a tolerable error.

The decision level \( g_0 \) is obtained from the final trial value \( g_0^T \). For the ideal detector, the saddle point \( s_1 \) under \( H_1 \) can be determined from (2.51). For the threshold detector, the saddle point \( s_1 \) under \( H_1 \) must be searched for again by iteration. That is, after the decision level \( g_{60} \) is obtained, we must arbitrarily pick the initial values \( s_{1N} \) and \( s_{1P} \) in such a way that the decision levels calculated from (2.69) are \( g_N \) and \( g_P \) and \( g_N < g_{60} < g_P \). The trial value of \( s_1^T \) will be determined from

\[
s_1^T = s_{1N} + \frac{s_{1P} - s_{1N}}{g_P - g_N} (g_{60} - g_N) . \tag{B4}
\]

By substituting the calculated value \( s_1^T \) into (2.69), the trial value \( g_{60}^T \) is determined. If \( g_{60}^T > g_{60} \), \( g_P \) and \( s_{1P} \) will be replaced by \( g_{60}^T \).
and $s_{1T}$, otherwise $g_N$ and $s_{1N}$ will be replaced instead. A new trial value $s_{1T}$ will be determined again from (B4). Repeat the procedure until the calculated value $g_{0T}$ approaches the decision level $g_{00}$ within a tolerable error. The final trial value $s_{1T}$ will be used to calculate the detection probability $Q_d$ for the threshold detector by using (2.55).

For calculating the integral, Simpson's rule has been used for the numerical integrations. The number of points used in the integration procedure depends on the relative error the system requires. In our calculations we use about 50 points at most times. More points can be used when higher accuracy is required. Furthermore, when the saddle point $s_k$ ($k = 0$, or 1) is near zero, double precision may be required to calculate the saddle point $s_k$ so that the term $I_k(g_0)$ in (2.55) can be evaluated accurately.

The typical behavior of the saddle point $s_0$ of the ideal detector for a Gaussian image is given in Fig. 27 where the values of $s_0$ are plotted as a function of signal-to-noise ratio $D^2$ at $N_0 = 0.5$ and 5, with preassigned false-alarm probabilities $Q_0 = 10^{-3}$ and $10^{-5}$. 
Figure 27  Saddle point $s_0$ vs. the signal-to-noise ratio $D^2$ for a Gaussian image; $N_0 = 0.5, 5$, $Q_0 = 10^{-3}, 10^{-5}$.
Footnotes

1. Daniels[8], §6.