

THE OPTIMIZATION AND ANALYSIS OF SYSTEMS FOR
THE DETECTION OF PULSE SIGNALS IN RANDOM NOISE

by

HAROLD V. HANCE

B.S., California Institute of Technology

(1939)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(1951)

Signature of Author_____

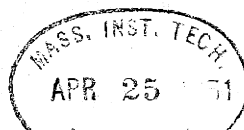
Department of Electrical Engineering
January 2, 1951

Certified by_____

Thesis Supervisor

Chairman, Departmental Committee on Graduate Students

EE
Thesis
1951



ABSTRACT

This investigation is concerned with the problem of the detection of weak pulse signals accompanied by random noise, with particular emphasis upon the case of radar signals. This problem is investigated both theoretically and experimentally, the experimental investigation being devised so as to check the theoretical results in so far as possible.

In the case of radar signals, the information relative to any one radar target arrives in the form of a train of pulses-- N in number. The individual pulses in the pulse train are perturbed by noise, assumed to originate within the receiving system, which causes the receiver output pulses to be randomly distributed in amplitude. In the absence of a target signal, noise pulses will still be present and hence can be mistaken as target signals. Based upon the values of the receiver output signal obtained in N successive observations at a particular range position, the detection system must make a decision and report whether a signal or only noise is present. It is shown that this decision can be based upon the result of a functional operation on the observed receiver output signals, X_1, X_2, \dots, X_N , by an operator $R(X_1, X_2, \dots, X_N)$. This operator is called the detection system function. The effect of noise in the system is to cause two types of errors, as follows:

Type I - with only noise present, the detection system reports a signal;

Type II - with a signal present, the detection system reports no signal.

These errors are of a statistical nature, and hence a probability measure of their effect upon detection system performance is required. A probability measure originally proposed by D. O. North is adopted for this investigation. The performance with respect to type-I errors is expressed by the noise probability, P_N , and with respect to type-II errors by the signal detection probability, P_D .

The system optimization problem considered is the determination of the detection system function that maximizes the detection probability subject to the condition that the noise probability is held constant. Consideration is given first to the analytical representation of the signal as delivered by the radio-frequency portion of the receiving system (including the intermediate-frequency amplifier and filter). It is shown that when certain idealizing assumptions are made, a discrete-sample representation of the signal is possible in which the samples are all statistically independent—a fact which greatly simplifies the analysis.

The optimization theory is developed first for the case of the detection of a single radio-frequency pulse of known phase angle. A geometrical representation of the signal samples is given, which in turn suggests a method of representing geometrically any arbitrary detection process or detection system function. The problem of choosing the

optimum detection system function is then solved by an application of the theory of mathematical statistics, following which the validity of this solution is demonstrated. It is found that the best detection system in this case is one in which observation is made of only the component of the receiver output signal in phase with the pulse signal to be detected, the so-called coherent detection system. The performance characteristics of this type of detection system are analyzed and the results presented in graphical form.

When the phase angle is unknown, the system defined by the optimization theory requires the observation of the amplitude of the receiver output signal without regard for its phase angle, this being the conventional envelope detection process. The results of the analysis for this case are found to differ but slightly from those for the coherent detector.

When N , the number of pulses available for detection, is greater than one and the phase angle of each is known, the optimum detection process consists of coherent addition or integration of the signals prior to rectification, followed by coherent detection. The minimum signal strength required for detection in this case is smaller than that required for the detection of a single pulse by the factor N . When the phase angles of the signals are unknown but constant, coherent integration is still called for, but envelope detection is required in place of coherent detection.

Finally, the case of detection of a train of pulses having random phases is considered. It is found that a straightforward application of the optimization theory does not lead to a unique solution of the problem, the optimum detection system being defined in this case only if the strength of the target signal to be detected is specified. In this case, the maximum likelihood principle of mathematical statistics is used to obtain a unique solution, although the detection system thus defined is not an optimum system in a strict sense. This solution requires the addition of the receiver output pulses following envelope detection. An approximate analysis based upon the central limit theorem shows that the minimum detectable signal strength varies inversely with N for N very small, but as N increases, the dependence approaches asymptotically an inverse square-root law.

Another process of signal integration is analyzed which differs from the ideal integrator in that the signals decay exponentially with time. It is found that the minimum detectable signal strength for this case is greater than that for the ideal integrator by approximately one-half decibel.

With the exception of the coherent integration and detection systems, all of the above-mentioned systems were studied experimentally. The experimental results obtained differ from the theoretical results by less than one-half decibel in all but a very few instances. The range of experimental conditions investigated is considered to be adequate to establish the validity of the theory to within the limits of the experimental accuracy.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to his thesis supervisor, Professor J. B. Wiesner, for many helpful suggestions regarding the subject matter of this investigation and for continued interest and encouragement. Gratitude is expressed to Dr. R. M. Ashby, through whom the author's interest in this problem was first aroused as a result of many stimulating and enlightening discussions, for much encouragement and helpful advice.

The author also wishes to acknowledge his indebtedness to the Naval Research Laboratory, whose liberal policy of employee education made it possible to carry out a portion of this investigation as a part of his official duties while employed at the Naval Research Laboratory Field Station, Boston, Mass., and to the Naval Air Development Center for permission to complete this work under a similar arrangement. Sincere thanks are due especially to Dr. H. Krutter under whose general supervision this investigation was conducted. The author wishes also to express appreciation to his colleagues at the Naval Air Development Center for their interest and helpful suggestions, to E. R. Shepherd and T. Ward who constructed much of the experimental equipment, and to his wife who assisted greatly by typing and editing the original manuscript and the final text.

TABLE OF CONTENTS

	<u>Page</u>
Abstract	ii
Acknowledgements	vi
Chap. I. INTRODUCTION	1
1.1 Introductory Discussion of the Optimization Problem	1
1.2 Background of the Optimization Problem	6
1.3 Background of the Development of Electrical Methods of Signal Integration	10
Chap. II. THEORY OF OPTIMUM SIGNAL DETECTION SYSTEMS .	16
2.1 Mathematical Formulation of System Optimization Problem	16
2.2 Discrete Sample Representation of Signals and Noise	19
2.3 Solution of the System Optimization Problem .	26
2.4 Example 1--Optimum Critical Boundary for Detection of a Single Radio-Frequency Pulse of Known Phase Angle	39
2.5 Example 2--Optimum Critical Boundary for Detection of a Single Radio-Frequency Pulse of Unknown Phase Angle	44
2.6 Example 3--Detection of a Sequence of N Radio- Frequency Pulses All of Equal Amplitude and Known Phase Angle	57
2.7 Example 4--Detection of a Sequence of N Radio- Frequency Pulses of Equal Amplitude with Constant but Unknown Phase Angle	59
2.8 Example 5--Detection of a Sequence of N Radio- Frequency Pulses All Having Equal Ampli- tudes but Random Phase Angles	62

	<u>Page</u>
Chap. III. FURTHER ANALYSIS OF SIGNAL DETECTION SYSTEMS	71
3.1 Type-I Integrator	72
3.2 Type-II Integrator	87
Chap. IV. DESCRIPTION OF THE EXPERIMENTAL SYSTEM . . .	97
4.1 General Description	97
4.2 Equipment Details	104
Chap. V. EXPERIMENTAL PROCEDURE AND RESULTS	128
5.1 Detection Without Integration	128
5.2 Detection With Type-I Integrator	134
5.3 Detection With Type-II Integrator	136
5.4 Significance of the Experimental Results . . .	140
Appendix A--References	141
Appendix B--Biographical Sketch of Author	144

LIST OF ILLUSTRATIONS

<u>Fig. No.</u>		<u>Page</u>
2.1	Amplitude versus Frequency Response Characteristics	20
2.2	Quadrature Components of Limited-Bandwidth Noise and Representative Ordinates	20
2.3	Vector Representation of Signal and Noise Currents	28
2.4	Representation of Detection System Function . .	28
2.5	Relationship of Sample Points to Critical Boundary	32
2.6	Relationship of Probability Density Contours to Critical Boundary	32
2.7	Two Alternative Critical Boundaries	36
2.8	Optimum Critical Boundary for Detection of a Single Pulse of Known Phase Angle, θ	36
2.9	Plot of the Function $P(u) = (2\pi)^{-1/2} \int_u^{\infty} \exp(-\frac{t^2}{2}) dt$	43
2.10	Signal Detection Probability versus Signal Strength--Single Pulse, Coherent Detection .	45
2.11	Threshold Signal Level versus Noise Probability	46
2.12	Plot of the Function $-\text{Log}_{10} P_N = (Z_0^2/2) \text{Log}_{10} \epsilon$	51
2.13	Signal Detection Probability versus Signal Strength--Single Pulse, Envelope Detection .	52
2.14	Threshold Signal Level versus Pulse-Train Number--No Integration	56
3.1	Block Diagram of Type-I Integrator	73
3.2	Response of Type-I Integrator to a Succession of Pulses	73
3.3	Mean Value and Standard Deviation of Envelope Function versus Signal Strength	77

<u>Fig. No.</u>		<u>Page</u>
3.4	Mean Value of Envelope Function in Central Measure versus Signal Strength	81
3.5	Threshold Signal Level Versus Pulse-Train Number--Type-I Integration, Noise Probability = 10^{-4}	83
3.6	Threshold Signal Level versus Pulse-Train Number--Type-I Integration, Noise Probability = 10^{-6}	84
3.7	Threshold Signal Level versus Pulse-Train Number--Type-I Integration, Noise Probability = 10^{-8}	85
3.8	Block Diagram of Type-II Integrator	89
3.9	Response of a Type-II Integrator to a Succession of Pulses	89
3.10	Standard Deviation of Output of Type-II Integrator versus Signal Strength	94
3.11	Threshold Signal Level versus Pulse-Train Number--Type-II Integration, Noise Probability = 10^{-6}	96
4.1	Signal Detection Study System	98
4.2	Timing Diagram for Signal Detection Study System	99
4.3	General View of Experimental System	100
4.4	Trigger and Gate Generator Unit	106
4.5	Pulse Train Control Unit	107
4.6	I.F. Pulse and C.W. Signal Generator Unit . . .	109
4.7	High-Gain I.F.-Amplifier Strip	113
4.8	Multi-Bandwidth Unit	114
4.9	Amplitude - Frequency Response Characteristic for 0.1 Mc Bandwidth Channel	116
4.10	Overall Amplitude Response Characteristic for Pulsed I.F. Signals	117

<u>Fig. No.</u>		<u>Page</u>
4.11	Block Diagram of Signal Integrator Unit	119
4.12	Response of Modified Type-I Integrator to Pulse Signals in Noise	125
4.13	Response of Type-II Integrator to Pulse Signals in Noise	127
5.1	Signal Detection Probability versus Signal Strength for Single Pulse--Experimental Results	131
5.2	Noise Probability versus Slicing Level--Experi- mental Results	133
5.3	Signal Detection Probability versus Integrator Time Constant--Experimental Results	138

CHAPTER I

INTRODUCTION

1.1 Introductory Discussion of the Optimization Problem.

This investigation is concerned with the problem of the detection of weak pulse signals imbedded in a background of random noise, a problem of importance in relation to certain types of communication systems as well as to radar detection systems. Although many of the results presented here are applicable in both cases, the investigation has been planned primarily with the radar problem in mind. The effects of noise upon the performance of a radar system is of great practical significance because it is one of the principal factors that determine the maximum range at which a specified object can be detected. This range is also influenced by a number of other factors such as transmitted pulse energy, operating frequency, antenna size and location, propagation characteristics of the path between the radar station and the object to be detected, reflection characteristics of the object, characteristics of the receiving system, length of time available for searching the space in which the object may be located, and, finally, the method by which the signal at the output of the receiving system is observed or sensed. All of these factors are dealt with to a greater or lesser extent in the voluminous literature on the subject of radar. It is not within the scope of this investigation to review even cursorily all of these aspects of the problem.

Therefore, it will be assumed that the reader is familiar with this subject at least to the extent that it is covered in a textbook on radar such as Radar System Engineering.⁽¹⁾

In this investigation, attention is confined to the problems of analysis and optimization of the detection capabilities of radar systems with respect to the deleterious effects of noise disturbances originating within the receiving system, assuming that the characteristics of the signals appearing at its input terminals have been determined by prior considerations.

So far, the term "detection" has been used loosely; however, it is readily appreciated that a precise definition of the function of a detection system is necessary before a quantitative investigation is possible. This function is to make decisions as to the presence or absence of objects in each of a number of volume elements or cells in space which taken together constitute the volume to be searched. It is apparent, however, that the character of the signals at the input terminals of the receiving system will depend upon the manner in which the space is searched or scanned. Because it is this input signal that is the starting point in the present investigation, the definition just given is not in a convenient form for application.

An alternative form more suitable for present purposes is easily deduced by recognizing that the volume elements scanned by the detection system are scanned in sequence in

almost every case. Consequently there is usually a simple correspondence between these volume elements and points on the time scale. Because of the repetitive character of the radar signal, information relative to a single object arrives piecemeal at the receiver. Hence it is often more convenient to express the time of occurrence of a received pulse, not in terms of its position on the usual continuous time scale with arbitrary origin, but rather in terms of its position on a time scale whose origin coincides with the beginning of the transmitted pulse from which the received pulse resulted. Alternatively, because of the correspondence between points on this latter time scale and the physical distance or range of the reflecting object or radar target which gives rise to the received signal, it is equally satisfactory to specify the range position of the target. The term range position as used hereafter is to be interpreted in the light of the foregoing discussion.

As a result of the azimuth (or elevation) scanning action of the radar system, the number of pulse signals received from any one radar target is limited, the exact number depending upon the scanning rate and the antenna beamwidth. Such a succession of pulse signals will be called a pulse train. Denoting this number by N , the required alternative definition of the function of the detection system can now be stated: it is to make decisions as to the presence or absence of target signals based upon N successive observations

of the received signal made at each of a number of range positions. More specifically, the radar system must assimilate this piecemeal information and produce a single "yes or no" report for each range position, this report being based upon the signal information received within the period of N pulses. Whatever form it may take, this assimilation process will be called integration. It is one of the objectives of this investigation to determine the best method of signal integration to be used in any specific case.

In the above discussion, the tacit assumption was made that the number of pulses per antenna beamwidth was strictly limited. This is equivalent to requiring that the antenna radiation pattern be strictly confined to a certain angular region, with no radiation occurring outside of these limits. This requirement is seldom satisfied in practice. Furthermore, during the passage of the antenna beam across a radar target, the received signal amplitude will vary in accordance with the radiation pattern of the antenna. Nevertheless, to avoid undue complication in the ensuing analysis, only the case of finite pulse trains of constant amplitude is considered. Even so, the results obtained are expected to be applicable in practice if the antenna beamwidth implied by this idealization is properly chosen in relation to that of the actual antenna under consideration.

Because we are concerned with the detection of weak signals, the maximum usable receiver amplification is to be

desired. For this investigation, the noise originating in the receiver input amplifier and associated circuits as a result of thermal agitation and shot effects is considered to be the limiting factor although, in practice, other sources of interference, such as electrical interference or terrain reflections, may be the limiting factors. For a given noise level in the input circuits, the maximum usable amplification is that which causes the noise at the output of the receiver to produce objectionable effects in the indicating or sensing element of the system. Here again, a quantitative definition must be formulated to express the effects of noise on the detection process. The objectionable effects are twofold. First, noise pulses may be mistaken for target signals and give rise to false alarms. Second, when a signal is present, it may be partially or totally annulled by the random noise disturbance and thus fail to produce a recognizable response. Because of the random character of the noise disturbance, it is apparent that the problem of weak signal detection can be studied adequately only on a statistical basis.

D. O. North⁽²⁾ first advanced a statistical criterion of radar system performance that recognizes these aspects of the problem. Because of the somewhat limited availability of North's report, and because his criterion is used as the basis for the present investigation, his criterion will be

WMO

WMO
WMO

discussed in detail. However, this discussion is deferred until the next chapter in order to permit a brief review of the background of the optimization problem to be given first.

1.2 Background of the Optimization Problem. The statistical approach in the study of radar system performance appears to have been first used by A. V. Haeff in experimental studies conducted at the U. S. Naval Research Laboratory. (3)* In these experiments the minimum detectable signal strength was defined as that for which an observer was correct one-half of the time in reporting the position of the signal as presented upon an A-type cathode-ray oscilloscope display.** A wide range of experimental conditions was covered in this investigation, and much information useful to the radar system designer was obtained. Empirical formulas were developed to represent the experimental results. However, the author did not present a theory to explain the observed results in terms of underlying statistical or psychological factors.

Experimental investigations conducted at the Massachusetts Institute of Technology Radiation Laboratory somewhat along the line of Haeff's work by Lawson and Ashby considerably extended the scope of the empirical data. (4) In addition, they investigated the performance of the Plan Position

*This work, although done in 1941-1942, was not published until 1946 because of security restrictions.

**Commonly used radar terminology will not be defined here when it is adequately defined in Reference (1).

Indicator (PPI) over a wide range of experimental conditions.

In all of these experiments a human observer was a vital part of the detection system; hence the psychological characteristics of the observer constituted an important but largely unknown factor affecting the results. In an attempt to minimize the randomizing effects of these factors, the role of the observer was reduced to that of making simple decisions. In the Radiation Laboratory experiments a signal was presented for a predetermined length of time on the radar oscilloscope at one of several marked positions. The observer was required to report at which of the positions he thought the signal occurred. For any particular set of test conditions, the fraction of correct statements, after correction for chance correct guesses, was taken as the detection probability. The detection threshold was defined as the signal level at which the detection probability was 90 percent. Using this procedure, a wide range of experimental conditions was covered. It was found that this procedure yielded results reproducible both between different observers and over long periods of time.

Along with this experimental work, considerable emphasis was placed upon the development of a theory of radar detection to account for the observed results.⁽⁴⁾ It was necessary, of course, to make certain assumptions concerning the characteristics of the human observer as they affect the detection of signals. The agreement obtained between theory

and experiment is in most respects quite satisfactory. In addition, Lawson and Uhlenbeck give a solution to the problem of determining the optimum detection system according to the criterion of performance used in the Radiation Laboratory research.

However, the applicability of these results for predicting the performance to be expected of practical radar systems is open to question on the grounds that the detection of signals in practice is not a matter of making a simple selection of one of a number of possible alternatives (the several possible range positions at which the signal is presented) given the a priori knowledge that a signal is present at one position only. It is the author's belief that the approach to the problem advanced by North and adopted for this investigation is considerably more realistic than that used in the Radiation Laboratory work. However, it should be said of the Radiation Laboratory criterion that it is well adapted to making subjective tests, and a large quantity of experimental data was obtained by this method.(4)

In some theoretical investigations, including portions of the work of Lawson and Uhlenbeck, the criterion of detection was formulated only to the extent of assuming that detection was possible if the signal-to-noise ratio of the video signal presented on the display was of the order of magnitude of unity.(5,6) On this basis, it is possible to correlate a substantial portion of the experimental data with

basic statistical phenomena. However, such a detection criterion is obviously inadequate as a basis for a general study of the detection problem as it does not provide for the determination of absolute signal levels required for detection.

Whatever the limitations of these earlier methods of attack, they have served to bring into perspective the relationship between many of the important parameters of a radar system and their effects upon signal detectability. This is especially true of one of the more important aspects of the signal detection process, namely that of signal integration. It is now quite generally appreciated that the ability of the radar display system, or of the observer, to integrate the piecemeal information as supplied by the radar receiver contributes greatly to the performance of a radar detection system. Signal integration as found in conventional radar systems is generally thought of as an averaging process wherein the random noise fluctuations at each point of the range scale are "smoothed out" while a recurrent signal pulse is augmented, thus producing an improvement in signal-to-noise ratio. The analysis of the effects of this "averaging" type of signal integration upon signal detectability on the basis of the signal-to-noise ratio criterion has been carried out by a number of workers. (4,5,7) However, as noted above, these results cannot be used to calculate absolute signal levels required for detection.

1.3 Background of the Development of Electrical Methods of Signal Integration. Because this investigation is concerned in part with the experimental study of the performance characteristics of signal detection systems employing certain techniques of electrical signal integration, a brief account of the development of such techniques will be given. The first form of electrical signal integration to be used appears to have been the so-called range-gated integrator. In this system, signals occurring in a short range-element or gate are accumulated in an electrical energy storage device. In its simplest form, the energy storage device is a capacitor connected to a high-impedance metering circuit. The origin of this type of integrator is not known to the author. This system is capable of extremely long integration periods and has been developed to a high degree of perfection.^(4,8) However, it is of limited utility in search radar applications because it is capable of observing only a single range element at a time. The use of a multiplicity of such elementary integrators to overcome this difficulty was investigated by T. T. Eaton and I. Wolff.⁽⁹⁾ However, because of its complexity, such a system has never found practical application.

Apparently the first suggestion of a method of electrical signal integration capable of integrating simultaneously at each of a large number of range positions was made as early as 1940 by Captain S. M. Tucker, U. S. Navy.⁽¹⁰⁾ In this

method the signal to be integrated is impressed upon the input terminals of a delay line. The output of the delay line is fed back regeneratively via an amplifier to the input terminals of the delay line. The delay through the amplifier and delay line is made exactly equal to the spacing between the pulses. Hence a pulse from the receiver travels through the delay line and the amplifier and arrives back at the input at the same time as the next pulse from the receiver. Thus, properly spaced, repetitive pulses build up according to the well-known theory of regeneration; whereas noise pulses, occurring randomly in time, build up less rapidly. Therefore signals are amplified more than the accompanying noise.

Because of the long delay time around the feedback loop, the system is stable only if the loop gain is less than unity. With this stability requirement satisfied, a single pulse applied to the input terminals of the integrator will be recirculated around the loop but with its amplitude reduced by a constant factor each trip around--a decay process which is exponential in character.

In early experiments with this integration system, conducted by the RCA Laboratories Division of the Radio Corporation of America, the system showed considerable promise. As a result extensive development work, with the objective of producing a system for field tests, was carried out by the RCA Laboratories under contract with the National Defense

Research Committee.⁽¹⁰⁾ For this purpose a lumped-parameter electrical delay line of 710 sections was constructed. This large number of sections was used in an attempt to minimize distortion of the pulse signals resulting from repeated recirculation through the delay line. Using this delay line, it was found in laboratory tests that a reduction of 6 decibels in minimum detectable signal strength was obtainable when observing continuously-repetitive pulses, i.e., pulse trains of very long duration such as would be received by a non-scanning radar system. The principal limiting factors were departures of the delay-frequency and amplitude-frequency characteristics of the delay line from the ideal. Because of these limitations in the delay line, and because it was anticipated that it would be very difficult to overcome them, no field tests were made. No other method of constructing an electrical delay line of the required fidelity was immediately available, and the project was terminated. However, in the final report on this project, mention was made of reports from other laboratories of promising results obtained with the ultrasonic delay line as a high-fidelity delay system. It is this system that was used in the experimental investigations to be described in later chapters.

Suggestions have also been advanced for the use of an electrical or ultrasonic delay line as a signal integrating device in which no external feedback path is used. In 1941 W. W. Hansen proposed to make use of the reflection properties

of an electrical delay line open-circuited at both ends.⁽¹¹⁾ If the attenuation factor of the line is sufficiently small, a signal applied at one end will travel back and forth through the line. When the round-trip delay is made equal to the spacing between pulses, this type of integrator is equivalent to the feedback type previously described. However, because of the attenuation of the signal caused by unavoidable losses in such a line, this system is not practicable.

The use of two-way repeaters inserted in the line to overcome the effects of line attenuation was made the subject of a patent application filed in September, 1942, by T. T. Eaton and D. G. C. Luck.⁽¹²⁾ This system, while workable in principle, is probably more complex to construct than the feedback integrator because of the need for balancing coils or transformers in the two-way amplifiers, and hence does not appear to be of practical interest.

The development of electrical signal integration techniques has not been centered entirely around the use of the delay lines. Techniques for the storage and integration of signals in the form of electrical charges deposited upon dielectric surfaces have been the subject of a number of investigations. The first experiments along this line were reported by T. T. Eaton and I. Wolff,⁽⁹⁾ who made use of the orthicon camera tube as the signal storage device. It was found, however, that the spatial non-uniformity of the

orthicon storage surface severely limited its performance as an integrator. Moreover, because the signal information was presented to the orthicon via a cathode-ray tube and optical system rather than in electrical form, this system was sufficiently complex as to be of doubtful practical value even if the orthicon target uniformity could have been sufficiently improved.

An electrostatic signal storage tube much better adapted than the orthicon tube for use as a signal integrator is the barrier-grid storage tube.⁽¹³⁾ J. V. Harrington and T. F. Rogers have reported the results of an investigation of the use of this tube for signal integration.⁽⁷⁾ They found that this method of signal integration is characterized by the same exponential decay process that was noted above for the delayed-feedback integrator. This decay process is a result of the erasure of stored charge by the scanning beam. The rate of erasure can be controlled, within limits, by adjustment of the scanning-beam current. This erasure factor is analogous to the loop-gain factor for the delayed-feedback integrator. Therefore this type of storage-tube integrator comes within the scope of the analysis to be presented in later chapters. However, for the experimental investigation, the delayed-feedback integrator was chosen because the range of experimental conditions attainable is greater than would have been the case with the storage-tube integrator.

Numerous other types of storage tubes have been described in the literature. However, none of them possesses the exponential decay characteristic, and therefore they do not come within the scope of this analysis.

*How about
amplifier methods
yet*

CHAPTER II

THEORY OF OPTIMUM SIGNAL DETECTION SYSTEMS

2.1 Mathematical Formulation of System Optimization

Problem. In Chapter I the main features of the signal detection problem were discussed in general terms. In this chapter a mathematical formulation of the problem is given, and a general analytical solution to the problem of determining optimum signal detection systems is found. This theory is then applied to several cases of practical interest. In later chapters the predictions of the theory are compared with the results of the experimental phase of this research.

The first step in the development of the theory is to devise a mathematical representation of the detection process. Each target, as it is scanned, is assumed to produce a sequence of radio-frequency signal pulses, N in number, at the input terminals of the detection system. These signal pulses there become mixed with noise. The range position and the instant of time at which they occur are unknown. It is the function of the detection system to examine the signals at each range position and report whether a signal or only noise (no signal) is present.* This report will therefore be based upon at most N signal samples at each range position.

*The meaning of the phrase "each range position" must be clarified before a quantitative analysis is attempted. This matter is taken up in the next section.

A general mathematical representation of the detection process that is consistent with the above ideas is then as follows: Let the N signal (or noise) samples at any one range position be completely characterized by measurable quantities $X_1, X_2, \dots, X_1 \dots X_N$, and let the detection system produce a report, R , which is a single-valued function of the sample values $X_1 \dots X_N$, i.e., define a detection system function

$$R = R(X_1, X_2 \dots X_N) \quad (2.1)$$

Then establish a critical level, R_0 , and when $R < R_0$ accept the hypothesis that only noise is present; conversely, if $R > R_0$, accept the hypothesis that a signal is present. Because of the presence of noise, there will be two possible types of error at each range position:

- type I - with only noise present, the detection system reports a signal;
- type II - with a signal present, the detection system reports no signal.

Type I errors are, of course, false alarms; and, for obvious reasons, their average rate of occurrence must be held to a low value, determined by the operational use to which the system is to be put. It will be seen later that, once the system function $R(X_1 \dots X_N)$ is chosen, the rate of occurrence of type I errors can be controlled by the choice of the critical level, R_0 . A suitable measure of

detection system performance with respect to type I errors is, therefore, the average false alarm rate, hereafter denoted by n_f .

A measure of system performance with respect to type II errors must also be adopted. It can be seen at once that the rate of occurrence of errors of this type will not be a suitable measure. Because errors of this type can occur only when target signals are present, their rate of occurrence would depend upon the rate of occurrence of signals and hence upon factors other than the parameters of the detection system itself--an obviously unsatisfactory situation. If, instead, a probability measure of type II errors is adopted, this logical difficulty is avoided because the probability of a type II error is just the fractional part of the total number of signals that occur which do not produce responses. The probability of detecting a signal is one minus the probability of a type II error, because when a signal is present either it will or will not be detected, and these two possibilities are mutually exclusive. The probability of detection of a signal, denoted by P_D , is used as the measure of system performance with respect to type II errors.

An optimum detection system can now be defined as one which maximizes the detection probability, P_D , when the false alarm rate, n_f , is some specified constant value. It remains now to discover how to determine this optimum system but, before a solution to this problem is attempted, certain additional points should be clarified.

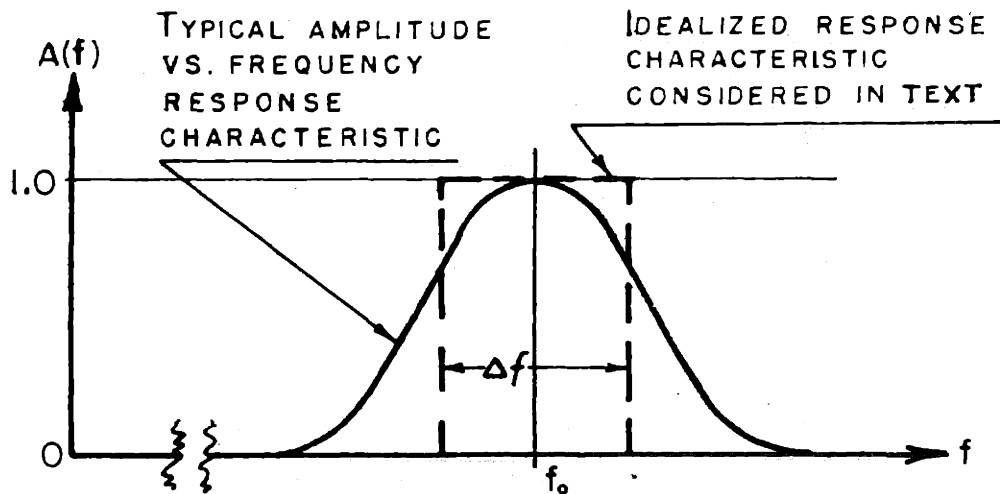
2.2 Discrete Sample Representation of Signals and Noise.

In the foregoing discussion, it has been stated that the operation $R(X_1 \dots X_N)$ should be applied at "each range position" without having defined the meaning of this phrase. To clarify this aspect of the problem it is necessary to consider the nature of the signal delivered by the radio-frequency portion of the receiver (including the intermediate-frequency amplifier). The overall amplitude versus frequency response characteristic, $A(f)$, is almost universally a curve somewhat as shown by the solid-line curve in Fig. 2.1, i.e., an approximately symmetrical curve centered at frequency f_0 and of width, Δ_f , small relative to f_0 . For this case the amplifier output current resulting from noise originating in the amplifier input circuits has a power spectrum density, $w(f)$, given by

$$w(f) = A^2(f) \quad (2.2)$$

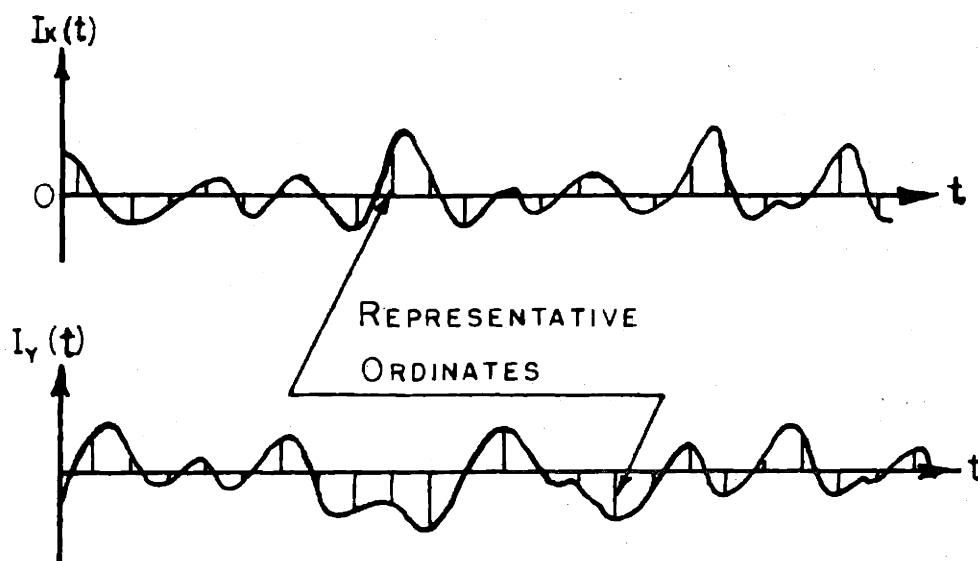
In order to simplify the following discussion, the actual amplifier is assumed to have been replaced by an idealized amplifier having a uniform response characteristic of width, Δ_f , as shown in Fig. 2.1 by the dashed-line curve. This idealized amplifier will produce a noise current, $I_N(t)$, having the same mean square value, Ψ_0 , as that for the original amplifier if

$$\Delta f = \frac{\int_0^{\infty} w(f) df}{W(f_0)} \quad (2.3)$$



AMPLITUDE VS. FREQUENCY
RESPONSE CHARACTERISTICS

Fig. 2.1



QUADRATURE COMPONENTS OF
LIMITED-BANDWIDTH NOISE AND
REPRESENTATIVE ORDINATES

Fig. 2.2

Rice has shown that this noise current can be represented as follows:*

$$I_N(t) = I_X(t) \cos \omega_0 t + I_Y(t) \sin \omega_0 t \quad (2.4)$$

where

$$\left. \begin{aligned} I_X(t) &= \sum_{n=1}^K c_n \cos[(\omega_n - \omega_0)t - \varphi_n] \\ I_Y(t) &= \sum_{n=1}^K c_n \sin[(\omega_n - \omega_0)t - \varphi_n] \end{aligned} \right\} \quad (2.5)$$

I_X and I_Y can be considered as modulation factors for the sinusoidal and cosinusoidal terms and are, of course, random functions of time. In these expressions the C_n 's are constants proportional to $A(f)$ and the φ_n 's are randomly distributed from 0 to 2π . For any particular set of C_n 's and φ_n 's, it is seen that I_X and I_Y will be functions of time defined by a sum of sinusoidal terms having frequencies $(1/2\pi)(\omega_n - \omega_0)$, $n = 1, 2, \dots, k$, and having amplitudes proportional to $A(f)$. Hence, I_X and I_Y will contain frequency components lying between zero and $(1/2\pi)(\omega_k - \omega_0)$ where k is the largest value of n for which $C_n \neq 0$. Hence the highest frequency component in I_X and I_Y is $\Delta f/2$. Furthermore, Rice has shown that I_X and I_Y are statistically independent variables with normal distributions having zero mean and variance ψ_0 .

*See Reference (14), especially sections 2.8 and 3.7.

Now Shannon has proved that any function of time, $f(t)$, limited to a band from zero to W cycles per second is completely determined by giving its ordinates at a series of discrete points spaced $1/2W$ seconds apart. (15) Applying this result to I_X and I_Y , it follows that any noise or signal waveform delivered by an amplifier of the type under consideration can be precisely specified by two sets of ordinates, their respective values being given at a series of discrete points spaced $1/\Delta f$ seconds apart. This representation is illustrated in Fig. 2.2. In the development of the theory of optimum detection systems, it will be found convenient to use this discrete sample representation. A further property of this representation is of importance in relation to the statistics of these sample values. Rice has derived the autocorrelation function of the variables $I_X(t)$ and $I_Y(t)$. (14) For the idealized case under consideration, it is

$$\rho_X(\tau) = \rho_Y(\tau) = \frac{\sin(\pi \Delta f \tau)}{(\pi \Delta f \tau)} \quad (2.6)$$

Hence the autocorrelation function is zero for $\tau = n/\Delta F$, n any integer. Since this is the spacing between the sample ordinates, it follows that all of the sample values are statistically independent. This fact greatly simplifies the statistical analysis.

The meaning that should be assigned to the expression "each range position" is now quite clear. As used hereafter,

this expression will mean "each of a sequence of points in time separated by an interval $1/\Delta f$ ". This is not intended to imply that the detection system must actually operate on the basis of sampled information. It merely means that nothing of value in the original signal would be ignored if it did so operate. Hence an analysis based upon this sampled information is still completely general.

The false alarm rate, n_f , can now be expressed very simply in terms of the statistics of the samples and a hypothetical sampling rate $\Delta f = 1/(\text{interval between samples})$. A false alarm can be considered to occur whenever the sample values associated with any one range position are such as to make $R(X_1 \dots X_N) > R_0$. Let P_N be the probability of this event happening, hereafter called simply the noise probability. Then the false alarm rate is given by

$$n_f = P_N \times \Delta f / N \quad (2.7)$$

where N is the number of successive radar repetition periods over which the received signals are "integrated". Because of this simple relationship, it is equally satisfactory for purpose of analysis to characterize the false alarm performance by P_N as by n_f , although the quantity of ultimate interest to the system designer is, of course, n_f .

This probability measure of false alarm performance appears to have been first used by North.⁽²⁾ The idea of a false alarm rate and its connection with the noise probability

was not introduced until somewhat later. This alternative approach has been used in some unpublished work, of which a paper presented at the 1950 Convention of the Institute of Radio Engineers is typical.⁽¹⁶⁾ Apparently, however, the equivalence between these two measures of false alarm performance has heretofore been accepted largely on intuitive grounds.

It is appropriate at this point to note the consequences of replacing the actual receiver filter by a rectangular filter. Using as a criterion the ratio of the peak signal power to average noise power at the output of the receiver filter, North showed that the optimum filter for the reception of rectangular pulses could be replaced by a rectangular filter of the proper bandwidth with a loss of only one decibel in signal-to-noise ratio.⁽²⁾ If this substitution were not made in this investigation, a very considerable elaboration of the theory would be required. In the first place, it is to be noted that there would be no definite upper limit to the frequencies contained in the functions $I_X(t)$ and $I_Y(t)$ since the C_n 's appearing in Eqs. 2.5 would approach zero asymptotically at frequencies widely separated from the band-center frequency, f_0 . Hence the representation of $I_N(t)$ by the use of sample ordinates would be exact only if the spacing between the sampling points were much less than $1/\Delta f$. As a further result, the sample ordinates would no longer

be statistically independent, thus greatly complicating the statistical analysis.

Returning now to the question of system optimization, the problem can be formulated as follows: It is required to determine the system function $R(X_1 \dots X_N)$ which maximizes the detection probability, P_D , while the noise probability, P_N , is held constant at some specified value. It appears that this problem was first investigated by North by a method which led to considerable analytical complexity.⁽¹⁷⁾ He found it necessary to use certain mathematical approximations in his analysis with the consequence that his results apply only to the case of the detection of signals small compared to noise. In the theory developed here, a different line of approach is used which, in addition to yielding general results in a straightforward manner, also serves to throw additional light upon the nature of the signal detection process. Furthermore, in this investigation the starting point is the signal available at the output of the i.f. amplifier prior to rectification, rather than the video signal after rectification, the starting point in North's analysis.

The basic mathematical and probability theory used in this new approach is an adaptation of the general theory of tests of statistical hypotheses, as developed by J. Neyman and E. S. Pearson.⁽¹⁸⁾ In this theory one is concerned with the problem of deciding how best to choose between two

alternative statistical hypotheses on the basis of a finite set of sample values or observations obtained by some specified sampling process. An elementary discussion of this problem can be found in most introductory texts on mathematical statistics. However, a much more adequate treatment is available in The Advanced Theory of Statistics, Volume II by M. G. Kendall. (19)

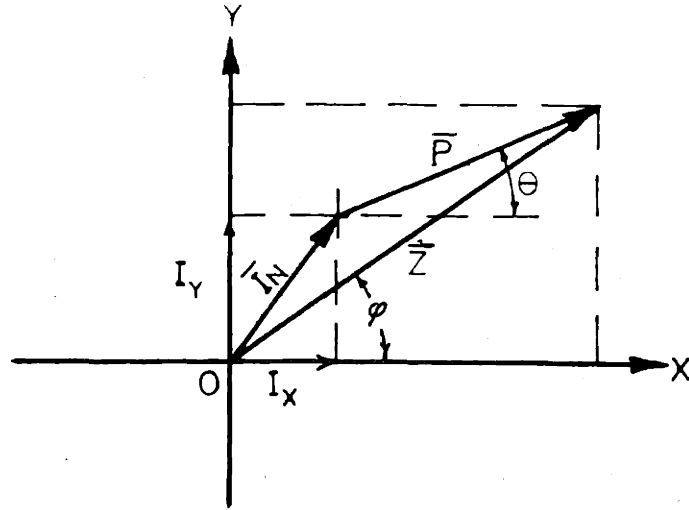
In this chapter the required theory is developed in terms of the signal detection problem. Also, instead of developing the theory initially in its most general form, some specialized cases are first considered in order to illustrate the basic ideas.

2.3 Solution of the System Optimization Problem. Consider first the case where only a single observation is made at each range position. This situation arises in the case of a radar system which scans the field of view very rapidly and in which there is no memory from one scan to the next. Let us suppose that the target signal and accompanying noise are observed at the output of a radio-frequency amplifier having the idealized bandpass characteristic described above. In the following discussion the unqualified term "signal" is used to refer to the time-varying receiver output-current without implying the presence of a target signal. Thus this "signal" may be entirely a result of noise disturbances. It is the function of the detection system to decide whether or not a target signal is present.

It is necessary now to decide what characteristics of the receiver output signal should be measured at each range position to serve as a basis for the detection of target signals. It was noted above that the signal at the output of the radio-frequency amplifier can be represented as the sum of two sinusoidal quantities whose phases are in quadrature and whose amplitudes are independently and normally distributed. This analytical representation leads to the usual geometrical representation, shown in Fig. 2.3, where the two vectors, I_x and I_y , represent the amplitudes of the two sinusoidal quadrature components of the noise disturbance.⁽²⁰⁾ If, in addition, a target signal of amplitude P is present, it can be represented by a third vector, \bar{P} , as shown. Its phase angle, Θ , is determined by the time of arrival of the target signal and by the instant of time chosen as the origin of the time scale. The resultant receiver output signal is represented by the vector \bar{Z} .

Because of the presence of noise in the receiver, it is not possible to measure P and Θ but, instead, only the magnitude, Z , and angle, φ , of the resultant of the signal and noise components. This resultant is equally well defined by its X and Y components, and this latter representation is used in the ensuing discussion.

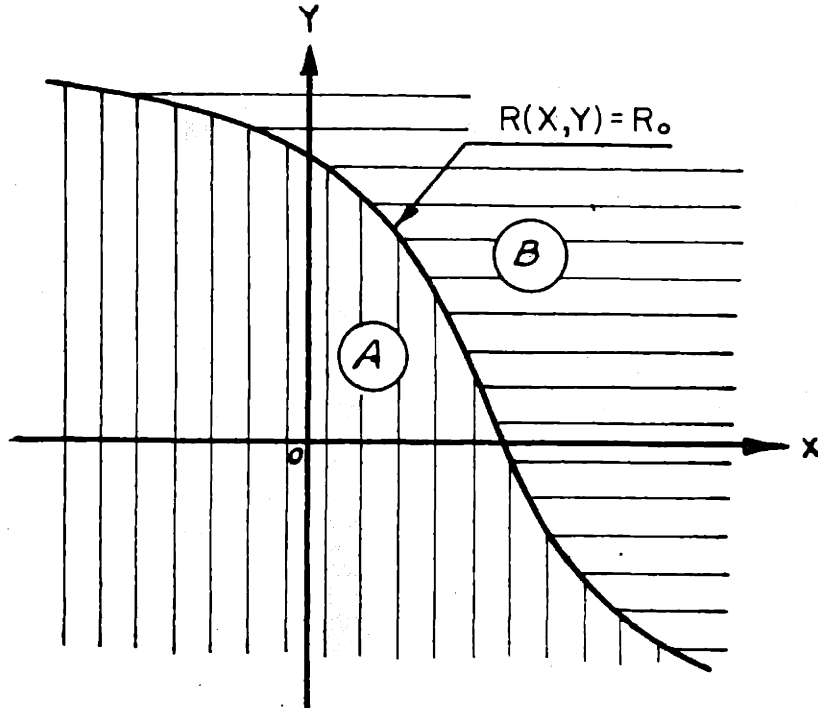
FIG 2.3



VECTOR REPRESENTATION
OF SIGNAL AND NOISE
CURRENTS

Fig. 2.3

FIG 2.4



REPRESENTATION OF
DETECTION SYSTEM
FUNCTION

Fig. 2.4

Following Rice, ⁽¹⁴⁾ the probability densities of the component variables,

$$\begin{aligned} X(t) &= I_X(t) + P \cos \theta \\ Y(t) &= I_Y(t) + P \sin \theta \end{aligned} \tag{2.8}$$

can now be written. They are

$$\left. \begin{aligned} p(X) &= (2\pi\psi_0)^{-1/2} \exp - \frac{(X - P \cos \theta)^2}{2\psi_0} \\ p(Y) &= (2\pi\psi_0)^{-1/2} \exp - \frac{(Y - P \sin \theta)^2}{2\psi_0} \end{aligned} \right\} \tag{2.9}$$

where ψ_0 is the mean square value of the noise current.

If, then, at each range position, the receiver radio-frequency output signal is resolved into two components, X and Y, these components completely characterize this signal and hence contain all of the data that will be useful in deciding when a target signal is present. It remains now to find the function $R(X,Y)$, which provides the optimum performance in the sense already discussed. To see how to proceed, it is helpful to consider the meaning of Eq. 2.1 in relation to the vector diagram of Fig. 2.3. In this example, Eq. 2.1 simplifies to

$$R(X,Y) = R_0 \tag{2.10}$$

because there are only two observables at each range position, i.e., the two signal components X and Y. R_0 is, of course, the critical level discussed at the beginning of this chapter. If now we regard X and Y as continuous variables, then for each different choice of the function $R(X,Y)$ Eq. 2.10 determines a curve in the X - Y plane. Fig. 2.4 shows a plot in

the $X - Y$ plane of the curve defined by one possible function. If $R(X,Y)$ is a continuous function, then for all points (X,Y) on one side of this line, $R(X,Y) < R_0$. Suppose that this corresponds to region "A" in Fig. 2.4.

Similarly for all points on the other side of this line, $R(X,Y) > R_0$, and this corresponds to region "B". It follows that when the observed receiver output components X and Y for a particular range position are such as to make $R(X,Y) < R_0$, the point in the $X - Y$ plane with coordinates (X,Y) will lie to one side of the line defined by Eq. 2.10, and when the values are such as to make $R(X,Y) > R_0$, the point will lie on the other side. Therefore, the decision to be made at each range position between the two alternatives, "signal" or "noise", can be based upon the position in the $X - Y$ plane of the point defined by the observed receiver output components.

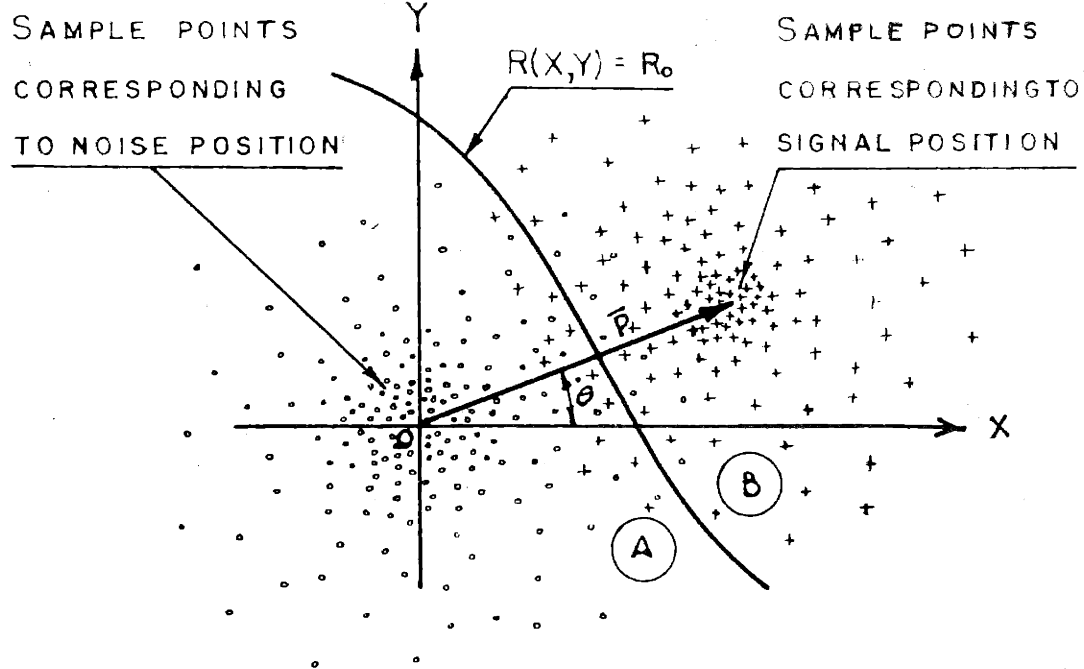
In the language of mathematical statistics, the $X - Y$ plane is the sample space of the random variables X and Y , and a point (X,Y) is a sample point. In the present example, the sample space is the entire $X - Y$ plane, as both X and Y can (in theory, at least) assume any arbitrarily large value, either positive or negative. The choice of a detection system function can, therefore, be regarded as a division of the sample space into two regions, a "signal on" region and a "signal off" region. The boundary separating these two

regions will be designated as the critical boundary in conformity with the terminology of mathematical statistics.

A geometrical representation of the probabilities of type I and type II errors in relation to the choice of the function $R(X,Y)$ will now be given. Suppose that a large, but finite, number of observations of the receiver output signal is made at a range position where no target signal is present. These observations, when plotted as sample points in the $X - Y$ plane, will be distributed somewhat as shown by the small circles appearing in Fig. 2.5, i.e., they will be scattered over the $X - Y$ plane but will be most dense in the vicinity of the origin, 0.

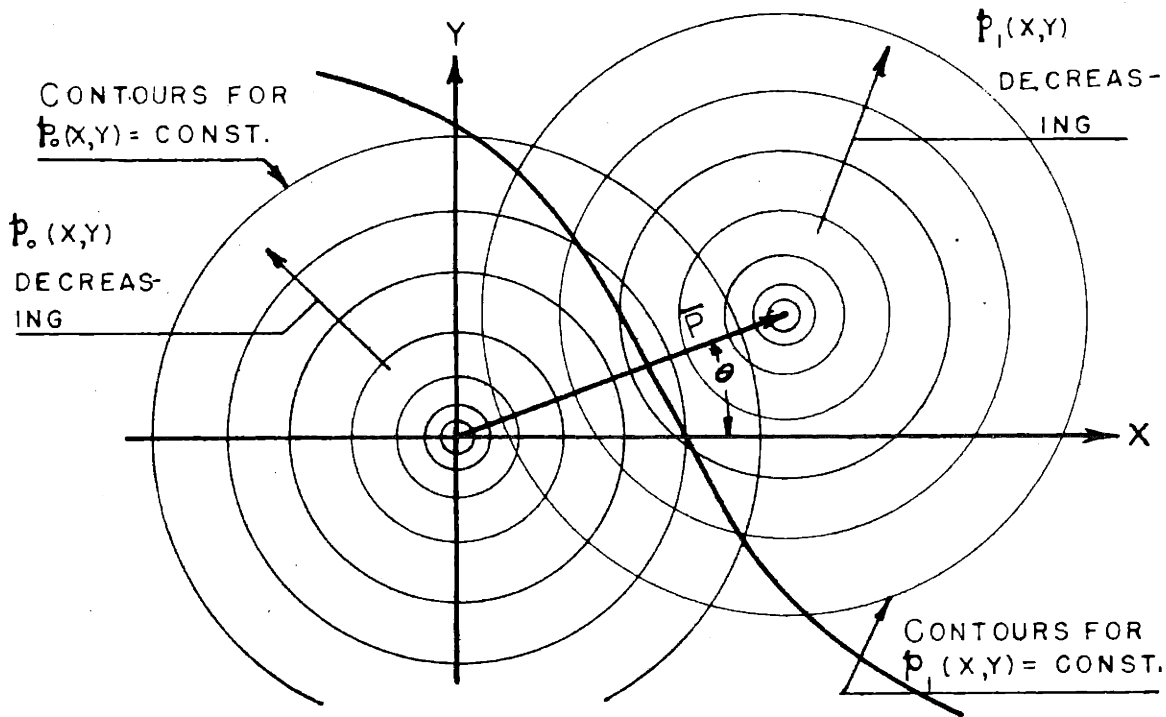
If similar observations are made at a range position at which is located a target signal of amplitude P and phase angle θ , the sample points as shown by the crosses in Fig. 2.5 will be concentrated about the tip of the vector \bar{P} .

A possible critical boundary for discriminating between target signals and noise is shown in Fig. 2.5. In this instance, since the majority of the sample points for the noise position fall to the left of the boundary in region "A", this should be taken as the "off" region. Similarly, region "B" is to be taken as the "on" region since most of the sample points for the signal position fall into this region. Therefore, a type I error occurs whenever a sample point for a noise-only range position falls into region "B", while a type II error occurs when a sample point for the signal position falls into region "A".



RELATIONSHIP OF
SAMPLE POINTS TO
CRITICAL BOUNDARY

Fig. 2.5



RELATIONSHIP OF PROBABILITY
DENSITY CONTOURS TO
CRITICAL BOUNDARY

Fig. 2.6

FIG. 2.5

FIG. 2.6

In order to discuss the probabilities of occurrence of these two types of error, the scatter diagram of Fig. 2.5 can be replaced by a diagram showing contours of constant probability density for observations at a noise-only range position and at a target signal position. The nature of these contours is shown in Fig. 2.6. In this diagram, $p_0(X,Y)$ is the probability density function of the sample values at a noise-only range position, while $p_1(X,Y)$ is the corresponding function for the target signal position. In the following analysis, a subscript "o" denotes the distribution function for the noise-only case, while a subscript "1" is used when the target signal is present.

These two families of contours can be thought of as representing a probability surface whose height above the X - Y plane at any point is $p(X,Y)$. The probability of a type I error is then given by

$$P_I = P_N = \iint_{\text{Region "B"}} p_0(X,Y) dXdY \quad (2.11)$$

This is just the volume bounded by the probability surface $p_0(X,Y)$, the cylindrical surface $R(X,Y) = R_0$ and region "B" of the X - Y plane. The probability of type II error is given by

$$P_{II} = \iint_{\text{Region "A"}} p_1(X,Y) dXdY \quad (2.12)$$

This is the volume bounded by the probability surface $p_1(X,Y)$, the cylindrical surface $R(X,Y)$ and region "A" of the X - Y plane.

Now the detection probability, P_D , is of greater practical interest than P_{II} and is related to it by the equation

$$P_D = 1 - P_{II} \quad (2.13)$$

Hence an equation equivalent to Eq. 2.12 is

$$P_D = 1 - \iint_{\text{Region "A"}} p_1(X,Y) dXdY = \iint_{\text{Region "B"}} p_1(X,Y) dXdY \quad (2.14)$$

The optimization problem can now be stated as follows:
Given the probability distribution functions $p_0(X,Y)$ and $p_1(X,Y)$ and a constant $P_N < 1$, it is required to find the critical boundary defined by $R(X,Y) = R_0$ which maximizes P_D . The solution to this problem was obtained by Neyman and Pearson by an application of the Calculus of Variations. (18)
It will suffice for present purposes to state this solution, which is very simple in form, and then verify that it possesses the desired maximizing properties. The optimum critical boundary is defined by

$$R(X,Y) = \frac{p_1(X,Y)}{p_0(X,Y)} = R_0 \quad (2.15)$$

Thus the optimum critical boundary is determined by the noise and signal-plus-noise probability density functions. The constant R_0 is as yet undetermined. It is clear, however, that

corresponding to any particular choice of $R(X,Y)$ each different choice of the critical level R_0 will determine a different critical boundary. Hence a value of R_0 can be found that will make the noise probability, P_N , assume any preassigned value (less than unity, of course). The probability density functions p_1 and p_0 are also known as likelihood functions, and their ratio is known as the likelihood ratio.

To show that Eq. 2.15 is the required solution, it is only necessary to show that if any other function $R'(X,Y)$ is used to determine a boundary according to an equation

$$R'(X,Y) = R'_0 \quad (2.16)$$

and R'_0 is chosen so that the noise probability has the same value in the two cases, then the detection probability will be smaller with this alternative boundary. In Fig. 2.7, two possible alternative critical boundaries are shown. Curve R is supposed to have been determined by the optimizing relation (Eq. 2.15), while curve R' is the alternative critical boundary corresponding to Eq. 2.16. Corresponding to the critical region "B" of Fig. 2.5, there are now two critical regions, one for each of the two curves R and R' . In general, these two regions will have a portion in common, represented in Fig. 2.7 by w_1 . There will also be two regions not shared, denoted by w_2 and w_3 . Since the noise probability is required to be the same for each of these two boundaries, it is

FIG. 2.7

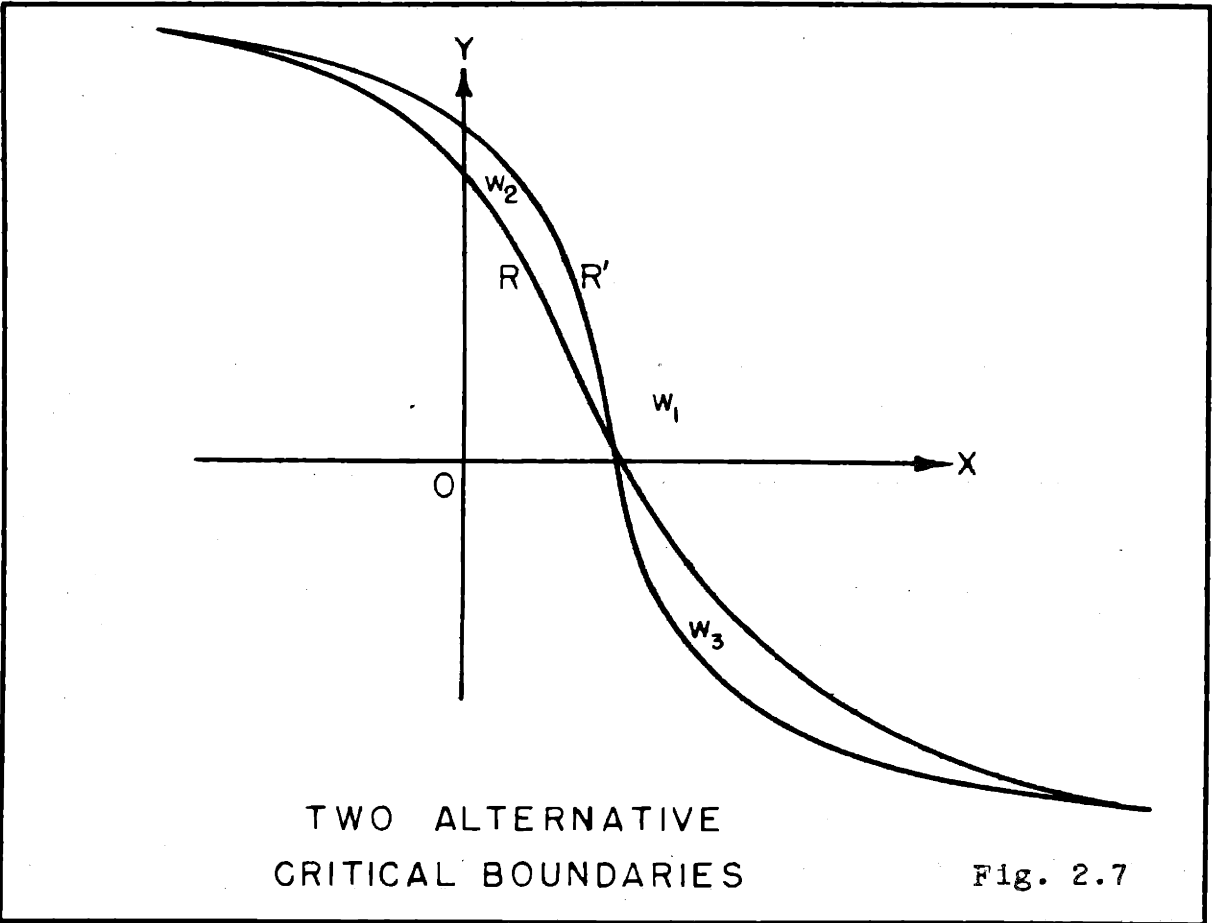
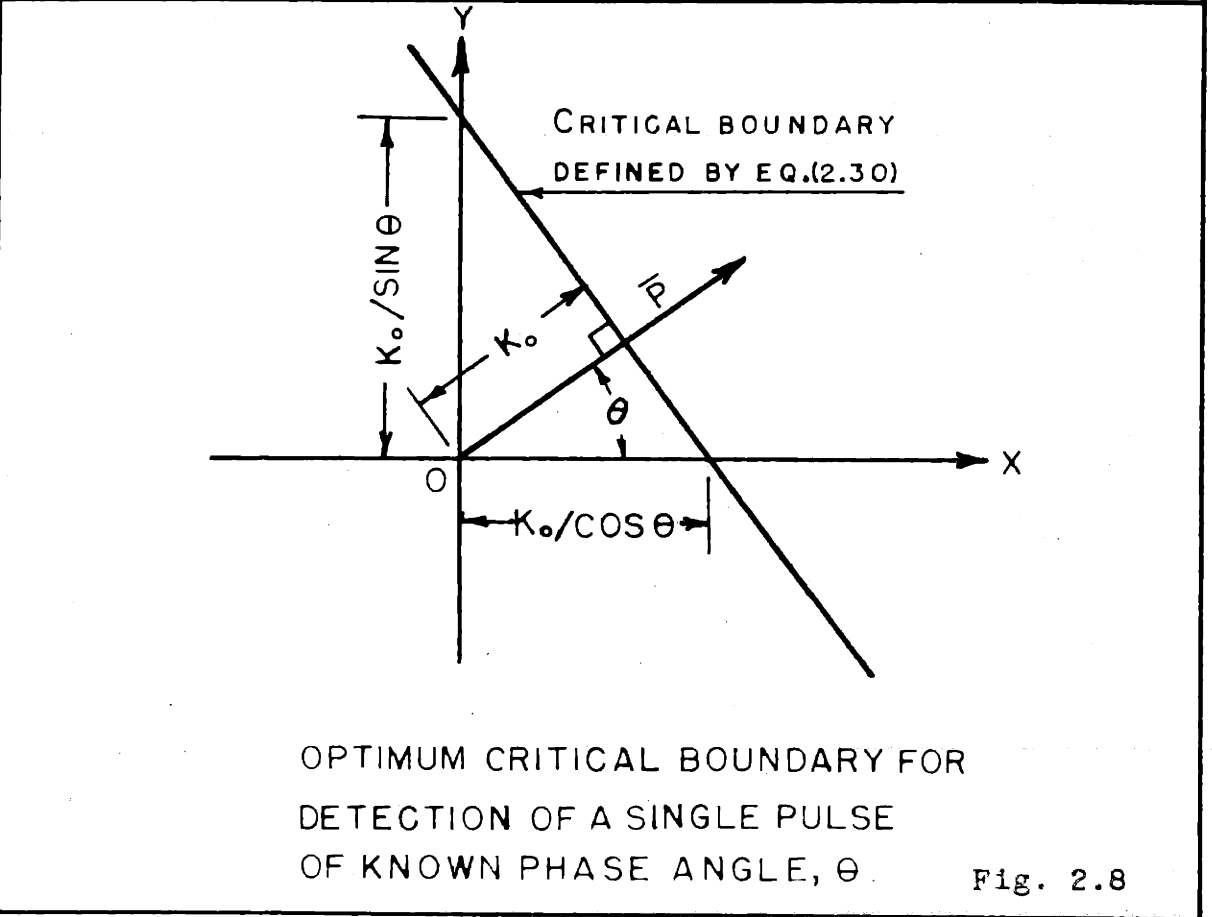


FIG. 2.8



necessary that these boundaries be related so as to satisfy the equation

$$\iint_{w_1 + w_2} p_0(X,Y) dXdY = P_N = \iint_{w_1 + w_3} p_0(X,Y) dXdY \quad (2.17)$$

Now each of these two integrals can be split into two parts as follows:

$$\iint_{w_1} p_0 dXdY + \iint_{w_2} p_0 dXdY = \iint_{w_1} p_0 dXdY + \iint_{w_3} p_0 dXdY \quad (2.18)$$

From which it follows that

$$\iint_{w_2} p_0 dXdY = \iint_{w_3} p_0 dXdY \quad (2.19)$$

That is, when observations are made at a noise position, the probability of the sample point falling in region w_2 is equal to the probability of its falling in region w_3 .

Now it has already been shown that $R(X,Y) > R_0$ for points (X,Y) lying to one side of curve R , while $R(X,Y) < R_0$ for points (X,Y) lying on the other side. Hence, if region $w_1 + w_2$ is the "signal-on" region, then in this region

$$R(X,Y) = \frac{p_1(X,Y)}{p_0(X,Y)} > R_0$$

$$\text{or } p_1(X,Y) > R_0 p_0(X,Y) \quad (2.20)$$

Similarly in region w_3 ,

$$p_1(X,Y) < R_0 p_0(X,Y) \quad (2.21)$$

It follows, then, from Eqs. 2.19, 2.20 and 2.21 that

$$\iint_{w_2} p_1 dX dY > \iint_{w_3} p_1 dX dY \quad (2.22)$$

Adding $\iint_{w_1} p_1 dX dY$ to both sides of this inequality yields

$$\iint_{w_1+w_2} p_1 dX dY > \iint_{w_1+w_3} p_1 dX dY \quad (2.23)$$

Here the left-hand member of the inequality is the detection probability for the case of the supposed optimum boundary and is greater than the right-hand member, which is the detection probability for any alternative boundary. Hence, the boundary defined by Eq. 2.15 does have the required maximizing property.

In the above discussion, it may appear that the particular type of deformation used in going from boundary R to R' is of a somewhat specialized character, since the boundaries are shown as having a point of intersection. Because of the constraint on P_N , it is seen that any deformation which tends to increase P_N must be counteracted by one having the opposite effect. Therefore, the type illustrated in Fig. 2.7 is the simplest admissible. Furthermore, any arbitrarily complex deformation can be arrived at by a succession of deformations of this simplest type, which indicates the generality of the proposition proved for this special case.

So far the discussion has been confined to the solution of the two-dimensional case. It was shown by Neyman and Pearson⁽¹⁸⁾ that this solution applies without change to the general case where N observations are available, provided the two-dimensional probability density functions are replaced by the joint probability density function for the N samples. Thus in the general case the critical boundary is defined by

$$R(X_1 \dots X_N) = \frac{p_1(X_1 \dots X_N)}{p_0(X_1 \dots X_N)} = R_0 \quad (2.24)$$

If $N > 3$, this equation can be interpreted geometrically only as a hyper-surface in a multi-dimensional space.

2.4 Example 1--Optimum Critical Boundary for Detection of a Single Radio-Frequency Pulse of Known Phase Angle. It is now possible to complete the solution of the two-dimensional case under consideration. Using Eqs. 2.9, $p_1(X,Y)$ can be written as

$$p_1(X,Y) = p_1(X)p_1(Y) = (2\pi\psi_0)^{-1} \exp -\frac{1}{2\psi_0} [(X-P\cos\theta)^2 + (Y-P\sin\theta)^2] \quad (2.25)$$

Upon expanding the squared terms in the exponent, this becomes

$$p_1(X,Y) = (2\pi\psi_0)^{-1} \exp -\frac{1}{2\psi_0} [X^2 + Y^2 + P^2 - 2P(X\cos\theta + Y\sin\theta)] \quad (2.26)$$

from which $p_0(X,Y)$ is obtained by placing $P = 0$.

$$p_0(X,Y) = (2\pi\psi_0)^{-1} \exp -\frac{1}{2\psi_0} [X^2 + Y^2] \quad (2.27)$$

Therefore, the required critical boundary is defined by

$$R(X,Y) = \frac{p_1(X,Y)}{p_0(X,Y)} = \exp\left\{\frac{1}{2\psi_0} [zP(X \cos \theta + Y \sin \theta) - P^2]\right\} = R_0 \quad (2.28)$$

Since $R(X,Y)$ is constant on the boundary, its logarithm will be also. Hence, an equivalent definition of the boundary is

$$\frac{1}{2\psi_0} [zP(X \cos \theta + Y \sin \theta) - P^2] = \ln R_0 \quad (2.29)$$

After multiplying through by ψ_0/P and transposing all constant terms to the right-hand side, the equation defining the boundary becomes

$$X \cos \theta + Y \sin \theta = K_0 \quad (2.30)$$

where K_0 is a function of R_0 , P and ψ_0 . It will be seen later, however, that K_0 can be determined directly and hence that R_0 need not be determined at all. For any specified value of θ , it can be easily verified that Eq. 2.30 defines a straight line at a distance K_0 from the origin and perpendicular to the vector \bar{P} , as illustrated in Fig. 2.8.

For a specified value of the noise probability, P_N , K_0 could be determined by use of Eq. 2.11 by expressing the value of the integral in terms of K_0 and θ . Fortunately, a much simpler approach is available in this case. It is observed from Eqs. 2.25 and 2.27 that the contours of constant probability density for $p_0(X,Y)$ and $p_1(X,Y)$ are circles centered respectively at the origin and at the terminus of the vector \bar{P} . Hence the statistics of the problem are unchanged if θ is set

equal to zero. Eq. 2.30 then becomes

$$X = K_0 \quad (2.31)$$

This means that the component of the receiver output signal in phase with the target signal is the quantity that should be observed. If, then, at any particular range position, the observed $X < K_0$, it must be presumed that only noise is present; while if $X > K_0$, a signal is presumed to be present. In this case, K_0 is some particular value of X , say X_0 , which is to be chosen so as to yield the desired noise probability, P_N . Hence X_0 is determined in terms of P_N by the following equation:

$$P_N = \int_{X_0}^{\infty} p_0(X) dX = (2\pi\psi_0)^{-1/2} \int_{X_0}^{\infty} \exp - \frac{X^2}{2\psi_0} dX \quad (2.32)$$

The detection probability is then given by

$$P_D = \int_{X_0}^{\infty} p_1(X) dX = (2\pi\psi_0)^{-1/2} \int_{X_0}^{\infty} \exp - \frac{1}{2\psi_0} [X-P]^2 dX \quad (2.33)$$

These expressions can be simplified somewhat by defining a new random variable

$$x = X/\psi_0^{1/2} \quad (2.34)$$

and by expressing the signal strength through the dimensionless quantity

$$\lambda = P^2/2\psi_0 \quad (2.35)$$

which is the ratio of the mean square signal current, $P^2/2$, to the mean square noise current, ψ_0 . Eqs. 2.32 and 2.33 then become

$$P_N = (2\pi)^{-1/2} \int_{x_0}^{\infty} \exp - \frac{x^2}{2} dx \quad (2.36)$$

and
$$P_D = (2\pi)^{-1/2} \int_{x_0}^{\infty} \exp - \frac{1}{2} [x - \sqrt{2\lambda}] dx \quad (2.37)$$

Eq. 2.36 can be expressed in terms of the probability integral

$$W(u) = (2\pi)^{-1/2} \int_{-u}^u \exp - \frac{t^2}{2} dt \quad (2.38)$$

by making use of the fact that the integrand is an even function and noting that $W(\infty) = 1.0$.* Thus

$$P_N = \frac{1}{2} - \frac{1}{2}W(x_0) = \frac{1}{2} [1 - W(x_0)] \quad (2.39)$$

from which

$$x_0 = W^{-1}(1-2P_N) \quad (2.40)$$

where $W^{-1}(z)$ is the function inverse to $W(z)$.

P_D can be expressed in terms of the probability integral, $W(u)$, by making the change of variable $u = (x - \sqrt{2\lambda})$ in Eq. 2.37. This yields

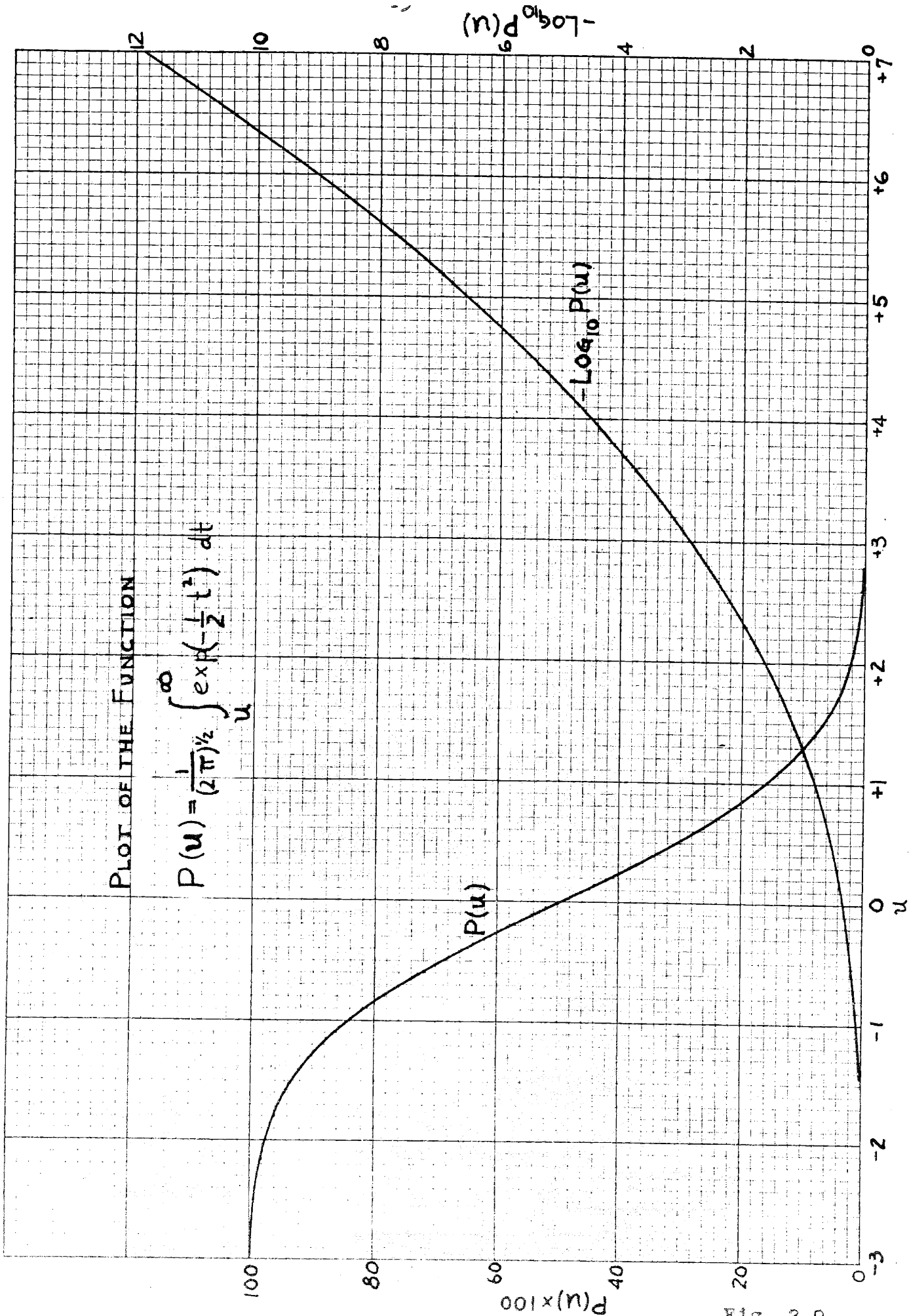
$$P_D = (2\pi)^{-1/2} \int_{u_0}^{\infty} \exp - \frac{u^2}{2} du = \frac{1}{2} [1 - W(u_0)] \quad (2.41)$$

A plot of the common logarithm of the function $P(u) = \frac{1}{2} [1 - W(u)]$ versus u appears in Fig. 2.9. This curve suffices for the determination of x_0 according to Eq. 2.40. For the determination of P_D as a function of λ according to Eq. 2.41, a curve of $P(u) \times 100$ versus u is also given in Fig. 2.9. In this case $u = x_0 - \sqrt{2\lambda}$.

*A tabulation of this integral is available in reference 21. The probability integral is closely related to the error function

$$\Phi(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp - t^2 dt$$

but the probability function, $W(u)$, is the more convenient form for use in the present problem.



PLOT OF THE FUNCTION

$$P(u) = \frac{1}{(2\pi)^{1/2}} \int_u^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt$$

$P(u)$

$-\text{Log}_{10} P(u)$

$-\text{Log}_{10} P(u)$

Fig. 2.9

It is seen from Fig. 2.9 that P_N is very strongly dependent upon x_0 . In Fig. 2.10, P_D is plotted as a function of λ with P_N as the parameter. It is seen that P_D is relatively insensitive to P_N and hence to x_0 . This latter relation is better illustrated by the dashed-line curves in Fig. 2.11, where the signal strength, λ_p , required for specified values of P_D is plotted as a function of P_N . Here p is the percent detectability and is equal to $P_D \times 100$ percent.

2.5 Example 2--Optimum Critical Boundary for Detection of a Single Radio-Frequency Pulse of Unknown Phase Angle.

It is seen from the preceding example that for each different phase angle of the signal to be detected there is a different optimum critical boundary, as typified in Fig. 2.8. In the usual case, this phase angle is not known in advance--any value between zero and 2π radians being equally probable. Therefore, the phase angle, θ , appearing in Fig. 2.8 has the character of a random variable with a probability density function

$$p(\theta) = \frac{1}{2\pi} \quad (2.42)$$

For this case, the joint probability distribution function of the sample values X and Y is no longer given by Eq. 2.26 as this must now be regarded as a conditional distribution function, $p(X, Y | \theta)$, for X and Y when θ assumes a specified value. The joint distribution function of X and Y alone, as required for application of the optimization theory, can be found by first obtaining the joint distribution $p(X, Y, \theta)$ and then

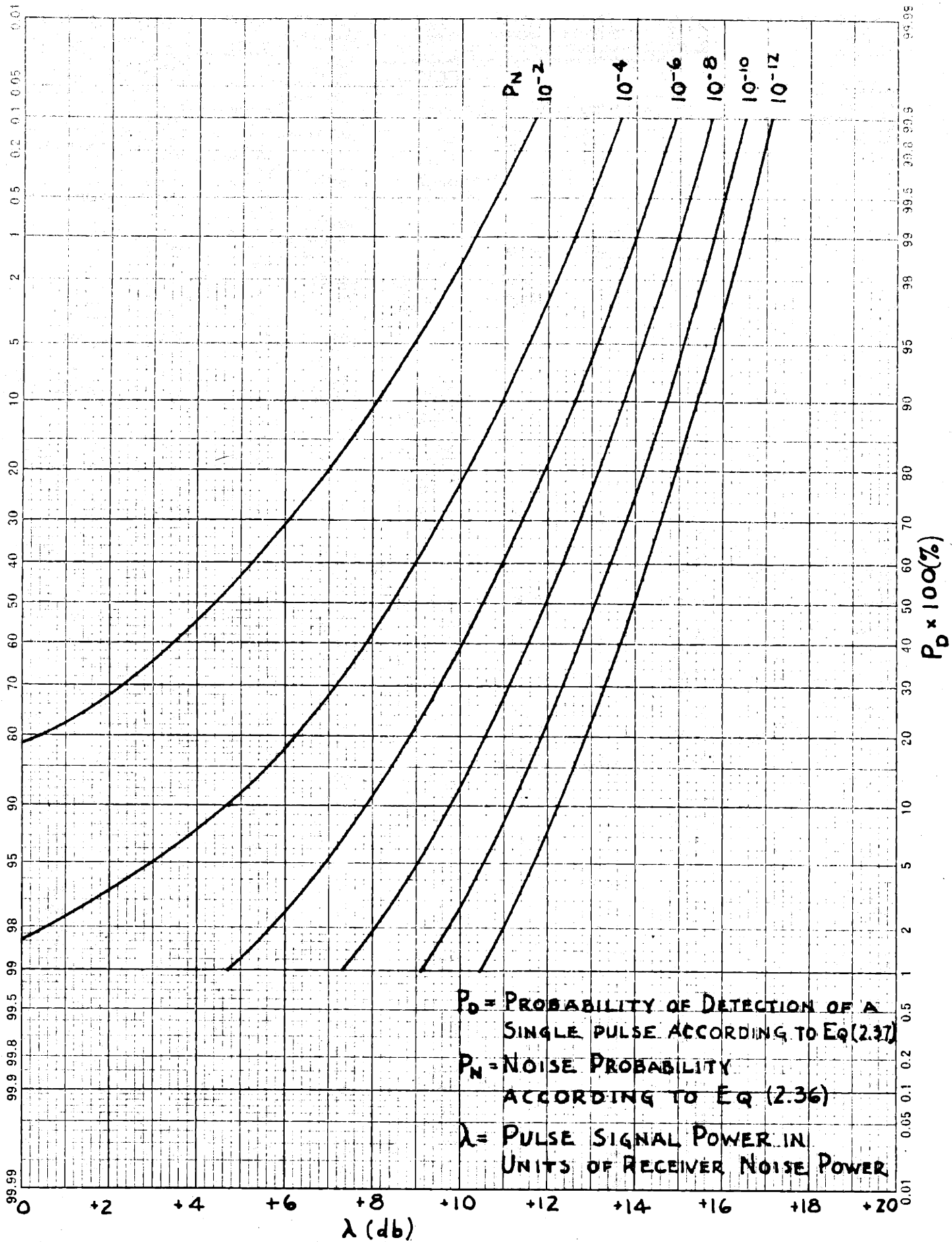


Fig. 2.10

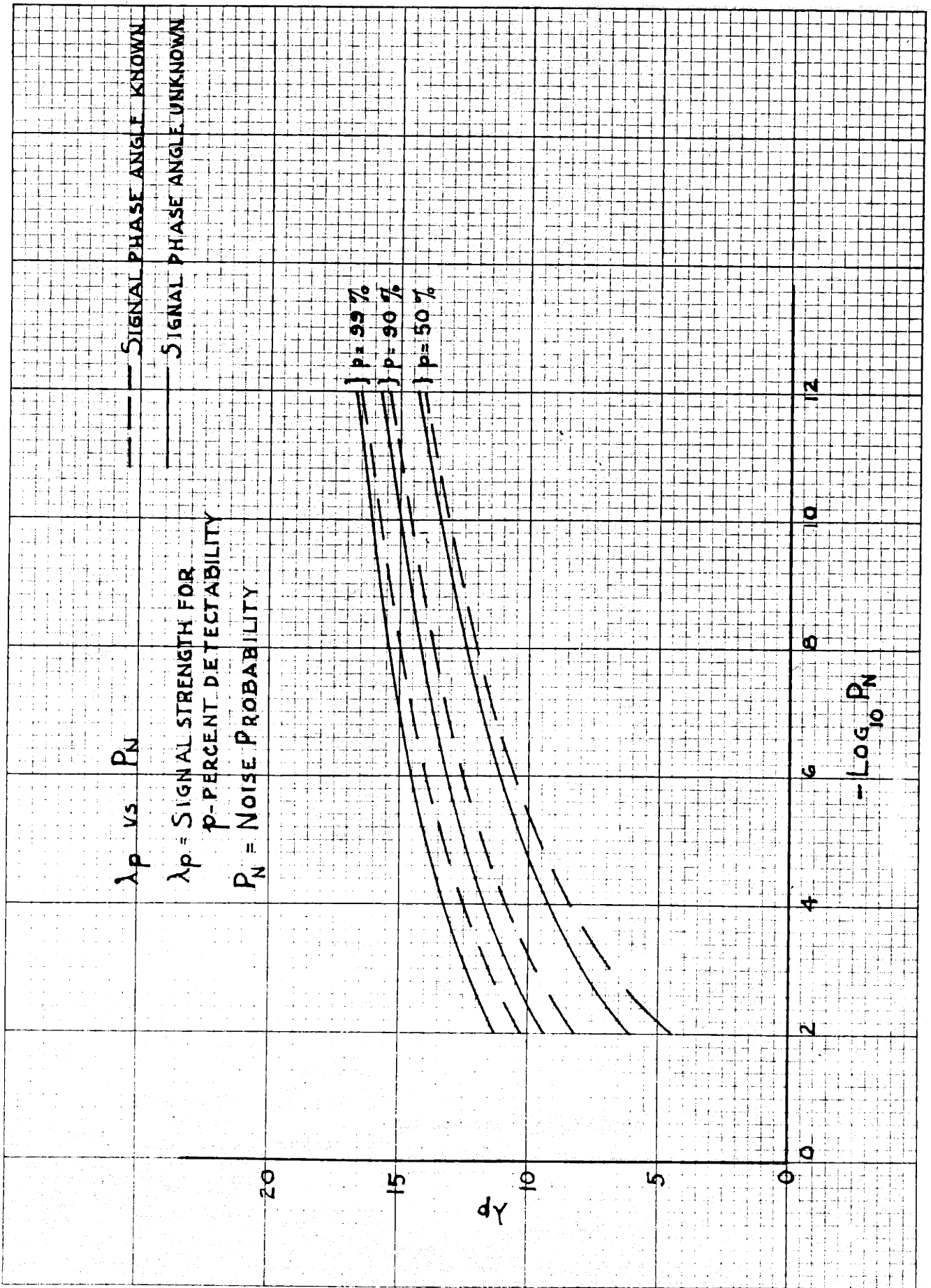


Fig. 2.11

integrating over the range of θ . Thus

$$p_1(X, Y) = \int_0^{2\pi} p(X, Y, \theta) d\theta \quad (2.43)$$

Now $p(X, Y, \theta)$ can be expressed in terms of the conditional distribution function, $p(X, Y|\theta)$, and $p(\theta)$ as follows: (22)

$$p(X, Y, \theta) = p(X, Y|\theta)p(\theta) \quad (2.44)$$

Hence Eq. 2.43 can be written as

$$p_1(X, Y) = \int_0^{2\pi} p(X, Y|\theta)p(\theta) d\theta \quad (2.45)$$

Using Eq. 2.26 for $p(X, Y|\theta)$ and Eq. 2.42 for $p(\theta)$ there results

$$p_1(X, Y) = \frac{1}{2\pi\psi_0} \exp\left\{-\frac{1}{2\psi_0^2}[X^2+Y^2+P^2]\right\} \int_0^{2\pi} \exp\left[\frac{P}{\psi_0}(X\cos\theta+Y\sin\theta)\right] \frac{d\theta}{2\pi} \quad (2.46)$$

The integral can be put into a recognizable form as follows:

Making use of the identity

$$X \cos \theta + Y \sin \theta = \sqrt{X^2 + Y^2} \cos(\theta - \varphi) \quad (2.47)$$

where $\varphi = \tan^{-1} \frac{Y}{X}$, the integral becomes

$$\frac{1}{2\pi} \int_0^{2\pi} \exp\left[\frac{P}{\psi_0} \sqrt{X^2 + Y^2} \cos(\theta - \varphi)\right] d\theta$$

which is of the form

$$\frac{1}{2\pi} \int_0^{2\pi} \exp[A \cos(\theta - \varphi)] d\theta$$

Making the change of variable $\alpha = \theta - \varphi$, the integral becomes

$$\frac{1}{2\pi} \int_{-\varphi}^{2\pi - \varphi} \exp[A \cos \alpha] d\alpha$$

Because the integrand is periodic with a period 2π , the value of the integral is unchanged if the limits of integration are changed to zero and 2π , giving for the integral

$$\frac{1}{2\pi} \int_0^{2\pi} \exp[A \cos \alpha] d\alpha.$$

This integral defines the Bessel function of imaginary argument, $I_0(A)$. Hence Eq. 2.46 becomes

$$P_1(X, Y) = \frac{1}{2\pi\psi_0} \exp\left\{-\frac{1}{2\psi_0}[X^2 + Y^2 + P^2]\right\} I_0\left(\frac{P}{\psi_0}\sqrt{X^2 + Y^2}\right) \quad (2.48)$$

$p_0(X, Y)$ can now be obtained simply by placing $P = 0$ in this expression and noting that $I_0(0) = 1$. Hence the optimum critical boundary is defined by

$$R(x, Y) = \frac{P_1(x, Y)}{P_0(x, Y)} = \exp\left[\frac{P^2}{2\psi_0}\right] I_0\left(\frac{P}{\psi_0}\sqrt{X^2 + Y^2}\right) = R_0 \quad (2.49)$$

It is seen that X and Y appear only in the combination $\sqrt{X^2 + Y^2}$ which is the envelope of the receiver radio-frequency output signal. Therefore, the best detection system in this case is based upon the observation of the envelope function.

For the further analysis of this case, it is now expedient to introduce a normalized variable

$$Z = \sqrt{X^2 + Y^2} / \psi_0^{1/2} \quad (2.50)$$

and, as before, take

$$\lambda = P^2 / 2\psi_0 \quad (2.35)$$

Eq. 2.49 then becomes

$$\exp[-\lambda] I_0(\sqrt{2\lambda} Z) = R_0$$

Transposing the constant factor $\exp[-\lambda]$ to the right-hand side of this equation, it becomes

$$I_0(\sqrt{2\lambda} Z) = k_0 \quad (2.51)$$

where k_0 is an undetermined constant.

Since the left-hand member of this equation is a constant on the optimum critical boundary, the argument of the

function $I_0(\sqrt{2\lambda} Z)$ must be also. Therefore, the required critical boundary is defined by

$$Z = Z_0 \quad (2.52)$$

where the value of Z_0 is to be determined so as to establish the required noise probability, P_N . The general method of procedure to be followed from here on in order to obtain quantitative results is analogous to that used in the previous example and, in fact, has already been carried out by other workers.⁽¹⁶⁾ However, the main results are presented here for the sake of completeness and to facilitate comparison with the results of the previous example. The probability distribution function of the envelope function, Z , is well known.^(2,14)

It is given by

$$p_i(z) = z \exp\left[-\frac{z^2}{2} - \lambda\right] I_0(\sqrt{2\lambda} Z) \quad (2.53)$$

Hence

$$P_N = \int_{Z_0}^{\infty} p_i(z) dz = \int_{Z_0}^{\infty} z \exp\left[-\frac{z^2}{2}\right] dz = \exp\left[-\frac{Z_0^2}{2}\right]$$

from which the following expression for the critical level, Z_0 , is obtained:

$$Z_0 = \sqrt{-2 \ln P_N} \quad (2.54)$$

Also

$$P_D = \int_{Z_0}^{\infty} p_i(z, \lambda) dz = \int_{Z_0}^{\infty} \exp\left[-\frac{z^2}{2} - \lambda\right] I_0(\sqrt{2\lambda} Z) dz \quad (2.55)$$

This integral cannot be expressed in terms of elementary functions. However, Rice has published in graphical form the results of numerical computation of this integral.⁽¹⁴⁾ He has very kindly supplied the author with an enlarged reproduction

of the published graph as well as a graph to an expanded scale. These graphs were further supplemented by an extensive tabulation of this integral supplied by J. I. Marcum of the Rand Corporation. With the help of these graphs and the curve of Fig. 2.12, which is a plot of Eq. 2.54, the curves of Fig. 2.13 were constructed. This latter figure shows the dependence of detection probability, P_D , upon signal strength, λ , for several values of the noise probability, P_N .

The circled points appearing in Fig. 2.13 were obtained experimentally. They are shown here for purposes of comparison with the theoretical results, although a discussion of the experimental techniques used must be reserved for a later chapter. To permit this comparison to be made, one additional factor regarding the theoretical calculation requires explanation. The signal detection probabilities have thus far been calculated in terms of the peak signal power at the output of the radio-frequency amplifier, the signal power being measured in units of average noise power at this point. However, it is usually of greater practical interest to specify system performance in terms of signal power at the input terminals of the receiver as measured in units of receiver noise power referred to the input terminals. When the receiver bandwidth is equal to the reciprocal of the pulse length, as required for maximum signal-to-noise ratio at the output of the radio-frequency amplifier, a steady-state condition is not reached during the pulse. The exact amount by which the output pulse

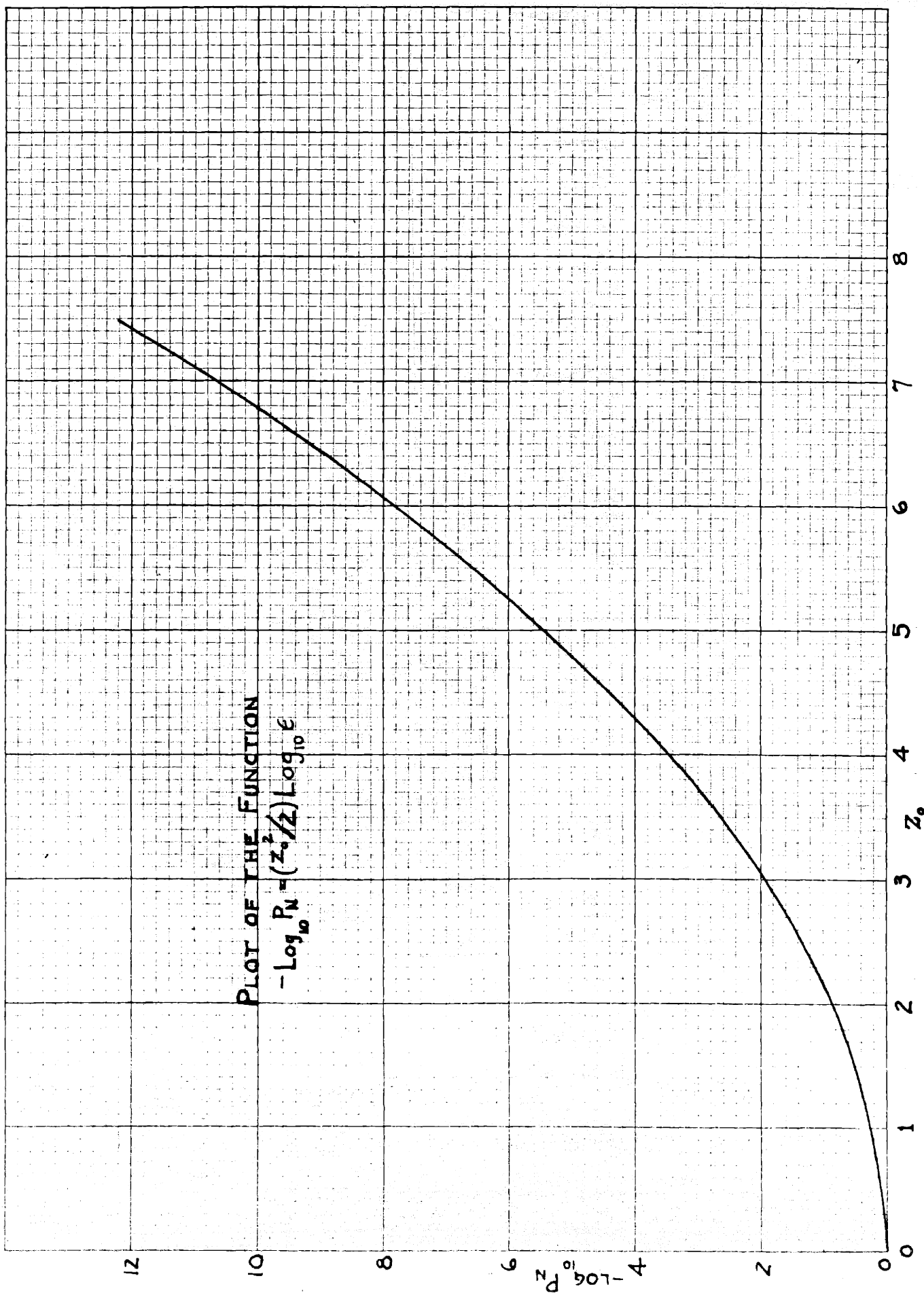


Fig. 2.12

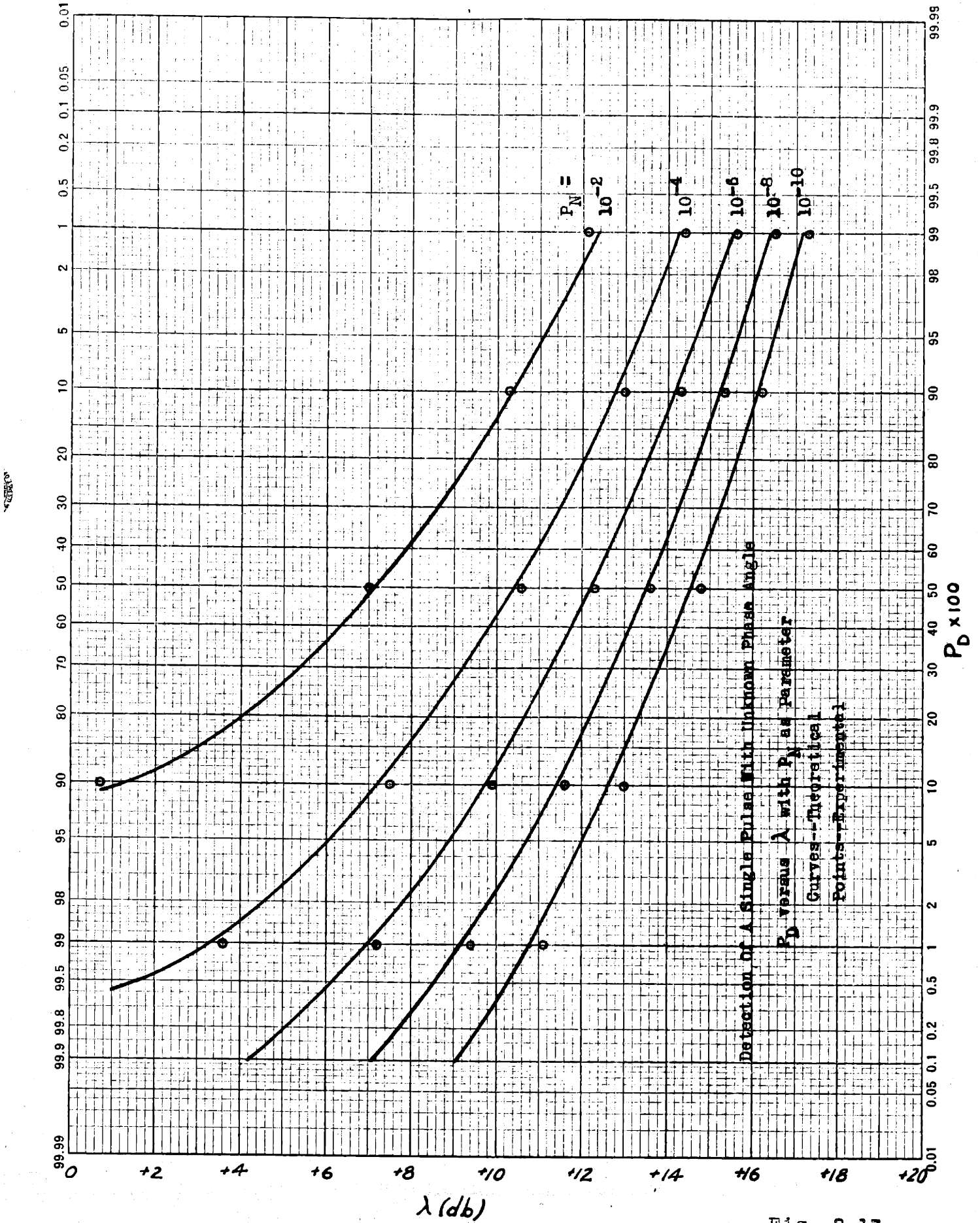


Fig. 2.13

falls short of reaching the steady-state value depends somewhat upon the shape of the receiver bandpass characteristic and the shape of the pulse. For the experimental system used in the present investigation, this factor is approximately one decibel. Consequently, to facilitate comparison between experiment and theory, the computed values of signal-to-noise ratio have been increased by one decibel. The agreement between theory and experiment is seen to be within 0.2 decibel for most of the points, with a maximum discrepancy of 0.4 decibel.

The results for the present case are compared with those for the previous one in Fig. 2.11, in which λ_p is plotted as a function of P_N for several values of p . The dashed-line and solid-line curves refer respectively to detection with known and with unknown phase angles. It is seen that λ_p is larger in the latter case by an average amount of only about one decibel.

A little reflection shows that the optimum critical boundary defined by $Z = \sqrt{X^2 + Y^2} / \psi_0^{1/2} = Z_0$ is intuitively sound. Referring to Fig. 2.8, it is seen that if θ is allowed to assume a succession of values between zero and 2π , then the critical boundary will sweep out a region bounded by a circle centered at the origin. This circle is, in fact, the envelope of the family of curves defined by Eq. 2.30 having θ as a parameter. Regarding as the critical region the portion of the X - Y plane on the side of the critical boundary away from the

origin, 0, it is seen that the region bounded by the circle includes a portion of the critical region corresponding to each value assumed by θ . The curve defined by $Z = Z_0$ is, of course, just such a circle.

Viewed in this light, the determination of the optimum critical boundary appears almost trivially simple. However, in a later example it will be found that a critical boundary that is optimum in the strict sense may be difficult, if not impossible, to determine. In such an event, the critical boundary defined by the envelope of curves in a manner analogous to that followed above, while not an optimum critical boundary, will still serve as a useful compromise.

Before leaving the subject of signal detection without integration, it is of interest to consider the effect upon threshold signal level of increasing the pulse train number, N , when detection is considered to occur if at least one pulse in the pulse train exceeds the critical level. Such a detection system has been considered elsewhere.⁽¹⁶⁾ It is discussed here as an example of an intuitively conceived system whose performance falls far short of that of an optimum system.

When the pulse train number, N , is larger than unity, each of the individual pulses in the pulse train is an independent event in the statistical sense because the amplitudes of the individual pulses are independently distributed. Let the number of pulses in the pulse train be N , and let the probability that at least one of these pulses will exceed

the slicing level be $P_D^{(N)}$. This is just one minus the probability that none will do so, which, in turn, is $(1 - P_D^{(1)})^N$ where $P_D^{(1)}$ is the probability of a single pulse exceeding the slicing level. Stated in a formula,

$$1 - P_D^{(N)} = [1 - P_D^{(1)}]^N$$

or

$$P_D^{(1)} = 1 - [1 - P_D^{(N)}]^{1/N}$$

Using $P_D^{(N)} = 0.9$, corresponding values of $P_D^{(1)}$ were calculated for several values of N . Then from the curves of Fig. 2.13 corresponding values of λ were obtained for several values of P_N , these being the λ_{90} values to be expected under the assumed conditions. The dependence of λ_{90} upon N thus determined is shown by the solid-line curves in Fig. 2.14 for several values of P_N . The results of similar calculations based upon the experimental data for $N = 1$ are shown by the crosses in Fig. 2.14, through which the dotted-line curves are drawn. In addition, experimental runs were made with values of N of 3, 10, 30, 100, 300 and 1000. The results obtained are shown by the circles appearing in Fig. 2.14. The agreement with the results of the extrapolation from the experimental data for the case of $N = 1$ is seen to be very good. Agreement between the theoretical and experimental results is again seen to be within 0.2 decibel for most of the points, with the maximum discrepancy being about 0.5 decibel. It is noted that λ_{90} does decrease somewhat with increasing N . However, it will be found later that the rate of decrease is small compared to that for systems employing integration.

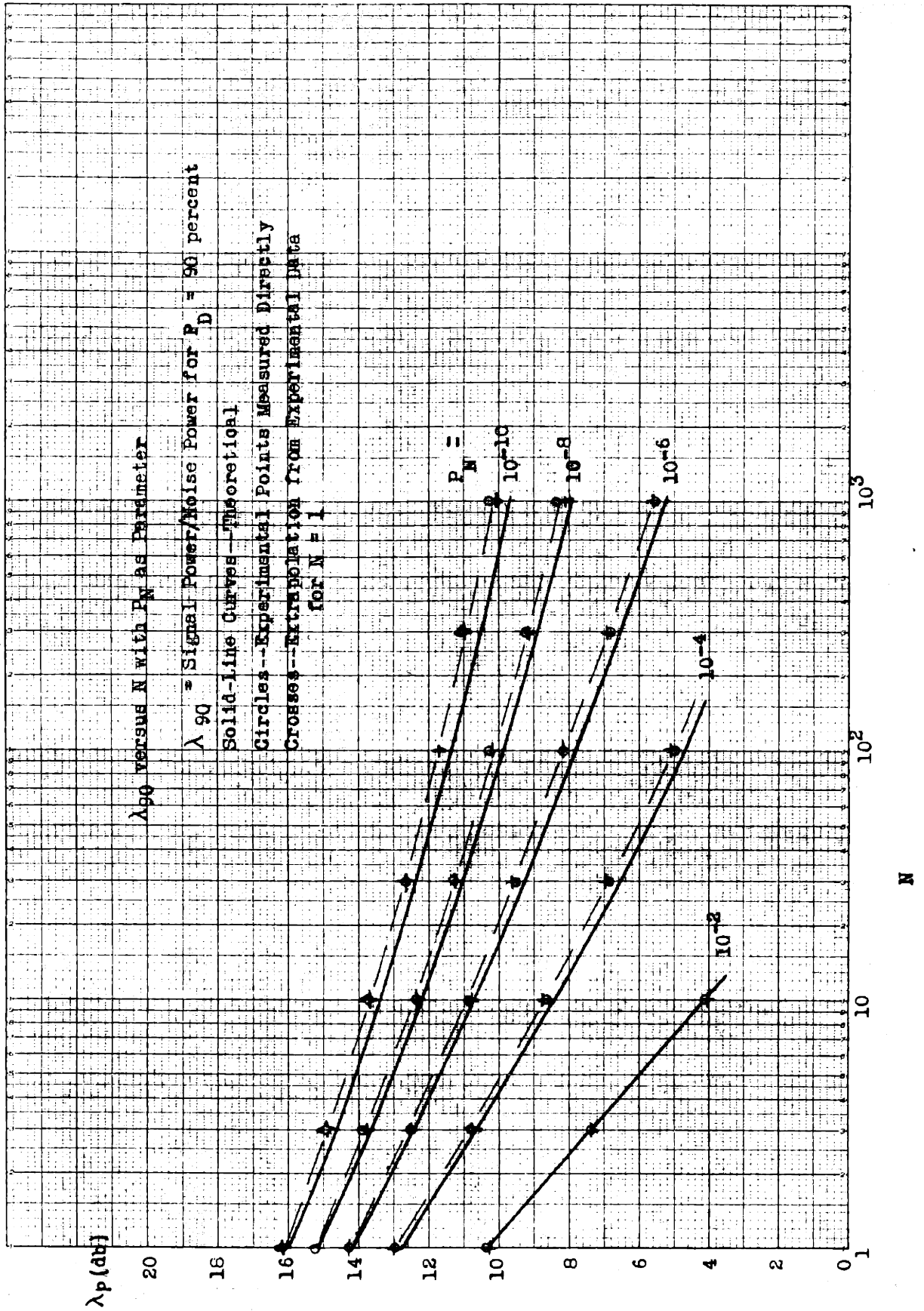


Fig. 2.14

2.6 Example 3--Detection of a Sequence of N Radio-Frequency Pulses All of Equal Amplitude and Known Phase Angle. For this case, the joint probability density function for the 2N observations, $X_1 \dots X_N$ and $Y_1 \dots Y_N$, is required. It is assumed that the amplitude, P , and phase angle, θ , of the signal pulse do not change from one pulse to the next. It is further assumed that the 2N observations at each range position are all statistically independent. Therefore, the required probability density function is a product of terms of the form of Eq. 2.25 and is given by

$$p_i(X_1 \dots X_N, Y_1 \dots Y_N | P, \theta) = \prod_{i=1}^N (2\pi\psi_0)^{-N} \exp\left\{-\frac{1}{2\psi_0} [(X_i - P\cos\theta)^2 + (Y_i - P\sin\theta)^2]\right\} \quad (2.56)$$

Expanding the squared terms in the exponent, this equation can be rewritten as

$$p_i(X_i, Y_i | P, \theta) = (2\pi\psi_0)^{-N} \exp\left\{-\frac{1}{2\psi_0} \sum_{i=1}^N [X_i^2 + Y_i^2 + P^2 - 2P(X_i\cos\theta + Y_i\sin\theta)]\right\} \quad (2.57)$$

where a simplified notation for the probability density function has been used in the left-hand member. At a noise-only position, $P = 0$ and hence

$$p_0(X_i, Y_i) = (2\pi\psi_0)^{-N} \exp\left\{-\frac{1}{2\psi_0} \sum_{i=1}^N [X_i^2 + Y_i^2]\right\} \quad (2.58)$$

The optimum critical boundary is now found by use of Eq. 2.24.

It is

$$R(X_i, Y_i) = \exp\left\{\frac{P}{\psi_0} \sum_{i=1}^N [X_i\cos\theta + Y_i\sin\theta] - \frac{NP^2}{2\psi_0}\right\} = R_0. \quad (2.59)$$

Again, the argument of the exponential function is constant on the optimum critical boundary. Setting this argument equal to a constant and transposing constant terms, the equation defining the critical boundary becomes

$$\sum_{i=1}^N (X_i \cos \theta + Y_i \sin \theta) = K_0 \quad (2.60)$$

Now $X_1 \cos \theta$ is the projection of the component X_1 onto the signal vector \bar{P} and similarly for $Y_1 \sin \theta$. This, however, is equivalent to placing $\theta = 0$ initially and observing only that component of the receiver output which is in phase with the signal to be detected. Hence an equally good definition of the optimum system function is

$$R(X_1, Y_1) = \sum_{i=1}^N X_i = K_0 \quad (2.61)$$

where X_1 is the in-phase component of the receiver output signal.

To determine signal and noise probabilities in this case, it is necessary to obtain an expression for the probability density function of $R(X_1) = \sum_{i=1}^N X_i$. Since the X_i are all distributed normally about the same mean, P , and all have the same variance, ψ_0 , the resultant distribution function will also be normal but with a mean, NP , and a variance, $N\psi_0$. Hence

$$p(R) = (2\pi N\psi_0)^{-1/2} \exp \left[-\frac{(R-NP)^2}{2N\psi_0} \right] \quad (2.62)$$

In terms of the dimensionless variables

$$\gamma = R/(N\psi_0)^{1/2}$$

and

$$\lambda = P^2/2\psi_0 \quad (2.63)$$

the distribution function becomes

$$p(r) = (2\pi)^{-1/2} \exp \left[-\frac{(r - \sqrt{2N\lambda})^2}{2} \right] \quad (2.64)$$

Comparing this result with the distribution function for a single observation, obtained simply by placing $N = 1$ in this expression, it is seen that the effect of integration of N observations is to produce a composite signal having $N\lambda$ as its effective power. Therefore, the results obtained for Example 1 are applicable in this example if λ is replaced by $\lambda' = N\lambda$. Furthermore, since $N\lambda$ is a quantity proportional to the total signal energy in the pulse train, it follows that the total signal energy is the quantity that determines the detection probability in this particular case. It is immaterial whether this energy is received in a single pulse or is divided among a large number of pulses.

2.7 Example 4--Detection of a Sequence of N Radio-Frequency Pulses of Equal Amplitude with Constant but Unknown Phase Angle.

In the example just considered, it was assumed that the phase angle of the signal was known. If instead it is unknown but constant for all pulses, Eq. 2.57 must be regarded as a conditional distribution function in which θ is a random variable. To obtain the probability density function for X_i and Y_i alone, it is necessary to multiply Eq. 2.57 by $p(\theta)$ and integrate over the range of θ . Again, taking $p(\theta) = 1/2\pi$ and the range zero to 2π , the required probability density function is

$$p(X_i, Y_i | P) = (2\pi \psi_0)^{-N} \int_0^{2\pi} \exp \left\{ -\frac{1}{2\psi_0} \sum_{i=1}^N [X_i^2 + Y_i^2 + P^2 - 2P(X_i \cos \theta + Y_i \sin \theta)] \right\} \frac{d\theta}{2\pi} \quad (2.65)$$

The part of the summation dependent upon θ can be rewritten

as
$$\sum_{i=1}^N (X_i \cos \theta + Y_i \sin \theta) = (\cos \theta) \sum_{i=1}^N X_i + (\sin \theta) \sum_{i=1}^N Y_i$$

$$= \left\{ \left(\sum_{i=1}^N X_i \right)^2 + \left(\sum_{i=1}^N Y_i \right)^2 \right\}^{1/2} \cos(\theta - \varphi)$$

where
$$\varphi = \tan^{-1} \frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N X_i}$$

Carrying out the integration yields

$$P_i(X_i, Y_i | P) = (2\pi \psi_0)^{-N} \exp \left\{ -\frac{1}{2\psi_0} \sum_{i=1}^N [X_i^2 + Y_i^2] - \frac{NP^2}{2\psi_0} \right\} I_0 \left\{ \frac{P}{\psi_0} \sqrt{\left(\sum_{i=1}^N X_i \right)^2 + \left(\sum_{i=1}^N Y_i \right)^2} \right\} \quad (2.66)$$

Setting $P = 0$ gives

$$P_0(X_i, Y_i | 0) = (2\pi \psi_0)^{-N} \exp \left\{ -\frac{1}{2\psi_0} \sum_{i=1}^N [X_i^2 + Y_i^2] \right\} \quad (2.67)$$

The optimum critical boundary is then defined by

$$R(X_i, Y_i) = \exp \left\{ -\frac{NP^2}{2\psi_0} \right\} I_0 \left\{ \frac{P}{\psi_0} \sqrt{\left(\sum_{i=1}^N X_i \right)^2 + \left(\sum_{i=1}^N Y_i \right)^2} \right\} = R_0.$$

Setting the argument of the function in the left-hand member equal to a constant and transposing constant terms gives as the definition of the optimum critical boundary

$$\sqrt{\left(\sum_{i=1}^N X_i \right)^2 + \left(\sum_{i=1}^N Y_i \right)^2} = Z_0. \quad (2.68)$$

A simple physical interpretation of this expression can now be given. With reference to Fig. 2.3, $\sum_{i=1}^N X_i$ is the sum of the components of the signal vectors and the noise vectors along the X-axis, which, of course, is the component along the X-axis of the resultant of a sum of vectors, $\bar{Z} = \bar{Z}_1 + \bar{Z}_2 + \dots + \bar{Z}_N$, and similarly for $\sum_{i=1}^N Y_i$. Therefore, the left-hand member of

Eq. 2.68 is the magnitude of the vector \bar{Z} . The operation defined by Eq. 2.68 can thus be replaced by the operation of vector addition. It is now recalled that these vectors represent sinusoidally-varying currents, and that the operation of vector addition corresponds to a linear superpositioning of these currents. Hence, it is seen that a detection system equivalent to the one defined by Eq. 2.68 would consist of some means for storing the radio-frequency signal current from one observation period to the next, adding the stored signal to the incoming signals, storing the resultant and repeating the entire process until all of the N signal samples have been added. Such a process is usually referred to as radio-frequency coherent integration.⁽⁴⁾ It is now apparent that the summation process derived in the preceding example can be replaced by coherent integration of the type just described. The results of Example 2 can be used for the present case if λ is replaced by $\lambda' = N\lambda$. This means that if the signal strength required for p -percent detectability in the case of a single pulse is $\lambda_p^{(1)}$, then for the case of N pulses the signal strength required is $\lambda_p^{(N)} = \lambda_p^{(1)}/N$. This dependence of threshold signal level upon the number of pulses available for integration is so simple as scarcely to require a graphical representation.

One of the basic assumptions in the present example is that all of the target signals are received with exactly the same phase angle. Obviously, if the target signal currents during successive observations have only a haphazard phase

relationship, they will tend to nullify one another as often as they reinforce. Unfortunately, in any practical radar system, the coherence requirement places very severe limitations upon allowable target velocities because the effect of target motion is to cause a shift in the phase of the target signal from one observation to the next. Hence, coherent integration, while of theoretical interest, has found no practical application. It now remains to consider the case of target signals with random phases.

2.8 Example 5--Detection of a Sequence of N Radio-Frequency Pulses All Having Equal Amplitudes but Random Phase Angles.

In this case, it is still possible to observe the phase of the signal as perturbed by noise, that is, phase is still a measurable characteristic of the signal. The observation of signal phase is, of course, subject to an ambiguity of an integral number of cycles. It is sufficient for present purposes to assume that the phase angle, θ_1 , for the 1th observation is confined to the range zero to 2π . The uncertainty in θ_1 can be expressed by regarding θ_1 as a random variable having a probability density of $1/2\pi$.

The joint probability density function for the N pairs of observations (X_i, Y_i) , subject to specified values of θ_1 , is as follows:

$$p_i(X_i, Y_i | \theta_i) = (2\pi\Psi_0)^{-N} \exp\left\{-\frac{1}{2\Psi_0} \sum_{i=1}^N [(X_i - P\cos\theta_i)^2 + (Y_i - P\sin\theta_i)^2]\right\} \quad (2.69)$$

To obtain the probability distribution function for the variables X_1 and Y_1 alone, it is necessary to multiply by $\frac{d\theta_1}{2\pi} \frac{d\theta_2}{2\pi} \dots \frac{d\theta_N}{2\pi}$ and integrate over the range zero to 2π for each θ_1 . This repeated integral is resolvable into a product of simple integrals, each of which is identical in form to that occurring in Eq. 2.46. The result of this integration is given by Eq. 2.48. Therefore, the required probability density function is

$$p_1(X_i, Y_i) = (2\pi\psi_0)^{-N} \prod_{i=1}^N \left[\exp\left\{-\frac{1}{2\psi_0} [X_i^2 + Y_i^2 + P^2]\right\} I_0\left\{\frac{P}{\psi_0} \sqrt{X_i^2 + Y_i^2}\right\} \right] \quad (2.70)$$

Hence the optimum critical boundary is defined by

$$R(X_i, Y_i) = \frac{p_1(X_i, Y_i)}{p_0(X_i, Y_i)} = \exp\left\{-\frac{NP^2}{2\psi_0}\right\} \prod_{i=1}^N I_0\left(\frac{P}{\psi_0} \sqrt{X_i^2 + Y_i^2}\right) = R_0 \quad (2.71)$$

Because the variables X_1 and Y_1 are not contained within the argument of a single function, it is not possible to proceed as before to simplify this expression. However, because the left-hand member of this equation is a constant on the critical boundary, its logarithm is also. Thus the repeated product can be replaced by a summation as follows:

$$\sum_{i=1}^N \ln I_0\left(\frac{P}{\psi_0} \sqrt{X_i^2 + Y_i^2}\right) = K_0 \quad (2.72)$$

where all of the constant terms have been included in k_0 .

The quantity $\sqrt{X_1^2 + Y_1^2}$ is, of course, the magnitude of the envelope function Z_1 . Using the dimensionless variables defined by Eqs. 2.50 and 2.35, Eq. 2.72 becomes

$$\sum_{i=1}^N \ln I_0(\sqrt{2\lambda} Z_1) = k_0 \quad (2.73)$$

It does not appear to be possible to further simplify this expression. Consequently, no geometrical representation of the optimum critical boundary suggests itself.

It is seen that the detection process defined by Eq. 2.73 requires that each signal sample, z_1 , be multiplied by the factor $\sqrt{2\lambda}$ and the product weighted by the operation $\ln I_0(x)$. The quantities resulting from these operations are then summed. All of these operations can, of course, be approximated physically to any desired degree of accuracy, and, in this sense, the detection system defined by Eq. 2.73 is physically realizable. However, it is to be noted that the optimum detection process is not determined until λ , the strength of the signal to be detected, is specified.

This situation is analogous to that encountered in Example 1 where it was found necessary to specify the phase angle, θ , in order to determine the optimum detection process. In Example 2 a probability density function was assumed for the unspecified parameter, θ , which permitted its elimination by integration. In addition to the parameter θ , λ was also contained in the original probability density functions for the variables X and Y. However, in that example, as in all of the others considered, it happened that the process of integrating to eliminate θ also disposed of λ . Since this did not occur in the present case, the same procedure used to eliminate θ suggests itself, that is, λ might be regarded as an unknown parameter having the character of a random variable to which a probability density function could be assigned.

Eq. 2.70 could then be multiplied by this function and the result integrated over the range of λ . However, the question would then arise of what probability density function to assign to λ . This function would depend, of course, upon such factors as the types of radar targets to be encountered and their distance from the detection system. Because of the very great range of conditions that occurs in practice, it appears that this procedure cannot be used to obtain generally applicable results, although its use in special cases might be practicable.

It is appropriate at this point to mention that the appearance, in some cases, of unspecified parameters in the equation defining the critical boundary was familiar to Neyman and Pearson.⁽¹⁸⁾ They also recognized the possibility of making use of a priori knowledge of the statistics of the parameter values to overcome this difficulty. Although no specific method of so doing was described, the method used in the foregoing analysis is in accordance with this basic idea.

The fundamental difficulty in the present example is the lack of knowledge of the statistics of the parameter λ . Evidently what is required in this situation is a detection process independent of λ (or any other unspecified parameters). It is relevant at this point to recall the discussion in Example 2 of the use as a critical boundary of the envelope of the critical boundaries corresponding to the range of values

of the parameter θ . Clearly, since the envelope was generated by allowing the parameter θ to vary over its entire range, the envelope is described by a function which is independent of θ . That this independence will also be obtained in the case where several parameters are involved can be seen by considering the method of determining the equation of the envelope of a family of surfaces defined by an equation of the form

$$f(X_1, X_2 \dots X_N; \alpha_1, \alpha_2, \dots \alpha_c) = 0 \quad (2.74)$$

where $X_1 \dots X_N$ are independent variables and $\alpha_1, \dots \alpha_c$ are parameters. The envelope of the family, if one exists, is found as follows: Differentiate Eq. 2.74 partially with respect to each of the parameters to obtain the equations

$$\begin{aligned} \partial f / \partial \alpha_1 &= 0 \\ \partial f / \partial \alpha_2 &= 0 \\ &\dots \dots \dots \\ \partial f / \partial \alpha_c &= 0 \end{aligned} \quad (2.75)$$

and use these equations to eliminate the α_i in Eq. 2.74. The resulting equation

$$g(X_1, X_2 \dots X_N) = 0 \quad (2.76)$$

is the equation of the envelope of the family*, and clearly it is independent of the parameters. Although a critical boundary defined by this envelope is not an optimum boundary in the strict sense, it at least satisfies one's intuitional

*Almost any text on advanced calculus can be consulted for a derivation of this method.

requirements for a good compromise in view of the infinite number of different boundaries which exists, corresponding to the range of values that the unspecified parameters can assume.

It is interesting to note that this method of determining critical boundaries was advanced by Neyman and Pearson on intuitive grounds prior to the development of their general theory of most efficient tests of statistical hypotheses. It is called the maximum likelihood principle. The interpretation of this principle in terms of the envelope of a family of curves was not discovered until after the general theory was developed. Neyman and Pearson also showed that if there exists an optimum critical boundary which is independent of the parameters of the original distribution function, then this boundary will coincide with the one given by the principle of maximum likelihood.

The envelope critical boundary will now be determined for the problem at hand. It is convenient to begin with the expression for the optimum critical boundary subject to specified values of phase angles and amplitude of the target signals. This expression can be obtained from Eq. 2.69 by the usual method. It is

$$\frac{P}{\Psi_0} \sum_{i=1}^N (X_i \cos \theta_i + Y_i \sin \theta_i) - \frac{NP^2}{2\Psi_0} = K_0 \quad (2.77)$$

By partial differentiation with respect to θ_i , the following is obtained:

$$-X_i \sin \theta_i + Y_i \cos \theta_i = 0, \quad i = 1, 2, \dots, N. \quad (2.78)$$

By partial differentiation with respect to P, there is obtained the equation

$$P = \frac{1}{N} \sum_{i=1}^N (X_i \cos \theta_i + Y_i \sin \theta_i) \quad (2.79)$$

Now Eq. 2.77 can be transposed to the form

$$P^2 - \frac{2P}{N} \sum_{i=1}^N (X_i \cos \theta_i + Y_i \sin \theta_i) = -\frac{2\psi_0 K_0}{N}$$

Completing the square in the left-hand member gives

$$\left[P - \frac{1}{N} \sum_{i=1}^N (X_i \cos \theta_i + Y_i \sin \theta_i) \right]^2 = \left\{ \frac{1}{N} \sum_{i=1}^N (X_i \cos \theta_i + Y_i \sin \theta_i) \right\}^2 - \frac{2\psi_0 K_0}{N} \quad (2.80)$$

By virtue of Eq. 2.79, the left-hand member of Eq. 2.80 is zero. Hence Eq. 2.80 becomes

$$\sum_{i=1}^N [X_i \cos \theta_i + Y_i \sin \theta_i] = \sqrt{2\psi_0 K_0} = k_0 \text{ say} \quad (2.81)$$

Eq. 2.78 can now be used to eliminate the θ_1 's from Eq. 2.81.

Begin by transposing so as to bring terms in θ_1 to the left-hand side as follows:

$$X_1 \cos \theta_1 + Y_1 \sin \theta_1 = k_0 - \sum_{i=2}^N \{X_i \cos \theta_i + Y_i \sin \theta_i\}$$

Square both sides of this equation and add to it the square of Eq. 2.78 with $i = 1$. This gives

$$X_1^2 + Y_1^2 = \left\{ k_0 - \sum_{i=2}^N [X_i \cos \theta_i + Y_i \sin \theta_i] \right\}^2$$

Thus θ_1 has been eliminated. Taking the square root and transposing so as to bring only terms in $i = 2$ to the left-hand side allows the same process of squaring and adding to be repeated to eliminate θ_2 . This procedure can be continued until all θ_i 's are eliminated, giving finally

$$(X_1^2 + Y_1^2)^{\frac{1}{2}} + (X_2^2 + Y_2^2)^{\frac{1}{2}} + \dots + (X_N^2 + Y_N^2)^{\frac{1}{2}} = k_0 \quad (2.82)$$

as the envelope of the family of surfaces defined by Eq. 2.69.

The quantity $(X_1^2 + Y_1^2)^{\frac{1}{2}}$ is, of course, the magnitude of the receiver output signal, or, in other words, the amplitude of the envelope function Z_1 . Hence the detection system defined by Eq. 2.82 is equally well defined by the equation

$$\sum_{i=1}^N Z_i = Z_0 \quad (2.83)$$

Therefore, according to the maximum likelihood principle, algebraic summation of the sample values at the output terminal of the envelope detector of the receiver is the ^{best} operation for discriminating between signals and noise. This operation can be regarded as a simple averaging process since $\sum_{i=1}^N Z_i = N\bar{Z}_N$ where \bar{Z}_N is the average of the values of N successive samples.*

Although the maximum likelihood principle has been used successfully to derive a detection system, it provides no measure of goodness of the system. An analogous problem exists in comparing different tests of statistical hypotheses. To deal with this problem, Neyman and Pearson introduced the concept of the power of a test. From the statistical standpoint, this measure of goodness is identical to the detection probability, P_D , as defined above. Thus two distinct detection systems will be considered equally good if they both detect a given strength of signal with the same probability when both are adjusted for equal noise probabilities.

*The symbol " $\bar{\quad}$ " placed over an algebraic quantity is used to denote the average value of that quantity. The more conventional " $\bar{\quad}$ " used in statistical analysis is avoided because it has already been used to indicate vector quantities.

An analytical and experimental investigation of the performance characteristics of signal detection systems employing signal integration is the subject of the remaining chapters.

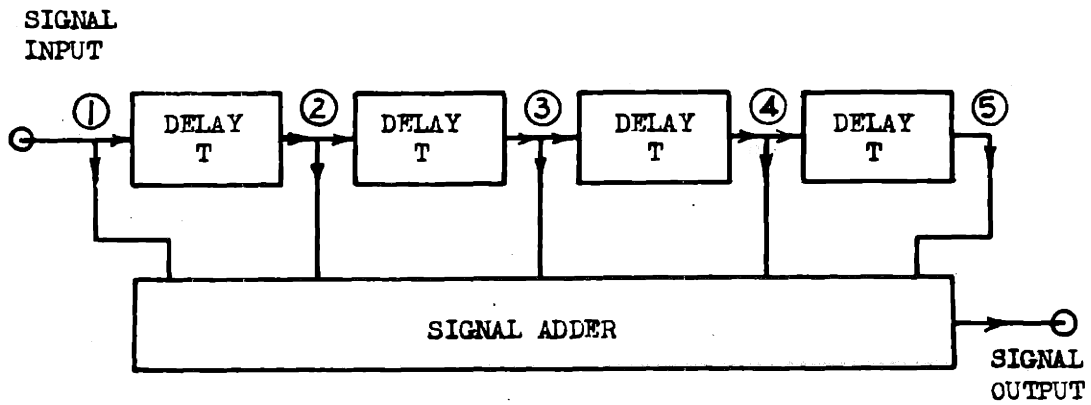
CHAPTER III

FURTHER ANALYSIS OF SIGNAL DETECTION SYSTEMS

It was shown in the preceding chapter that the detection of a train of radio-frequency pulses accompanied by random noise is best accomplished by performing the elementary operation of adding the individual pulses. When the radio-frequency pulses are coherent in phase, the addition should be done prior to envelope detection. The analysis for that case was found to be exceedingly simple, primarily because the probability density function for the amplitudes of the component signals was normal. When phase coherence is lacking, the addition must be done after envelope detection. In this case the amplitudes of the individual component signals are no longer distributed normally. For this and other reasons which will soon become apparent, the analysis of this case is more complex than for the phase-coherent case.

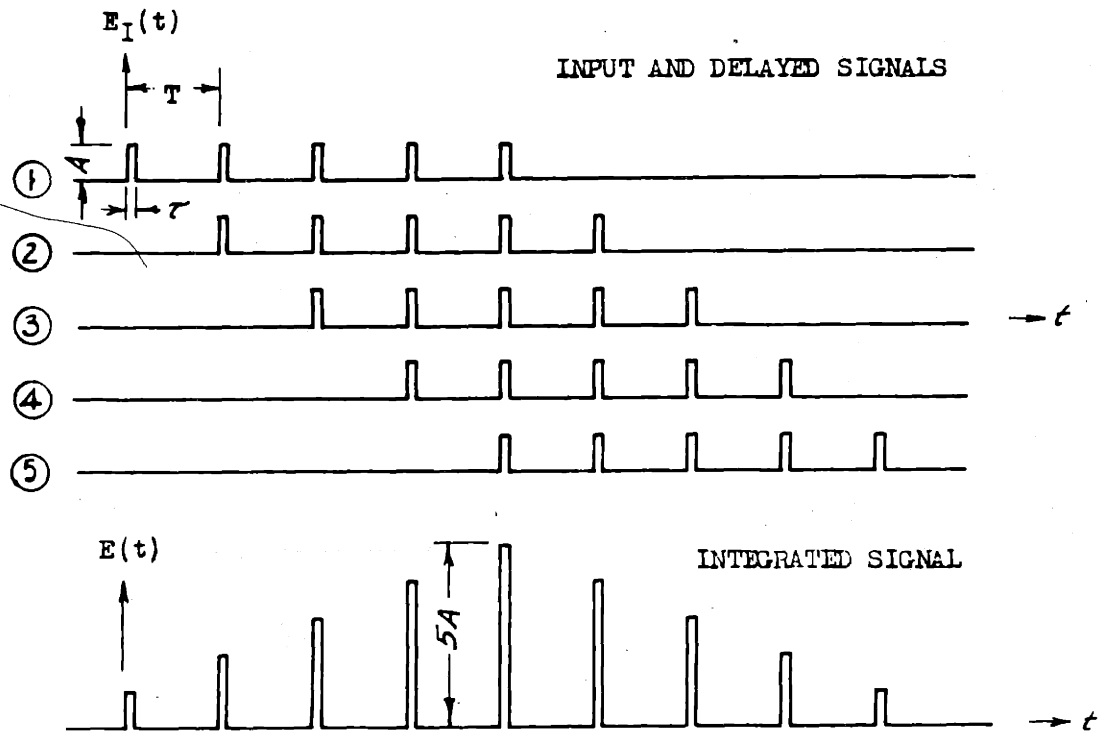
Hereafter, an integrator which performs this simple addition process is called a Type-I or ideal integrator. It will be seen, however, that this ideal integrator, while the simplest from the analytical standpoint, is realized physically only with considerable difficulty. An integrator making use of the delayed-feedback signal integration technique discussed in Chapter I is more easily constructed, but is only an approximation of the ideal integrator. Both types are analyzed in this chapter.

3.1 Type-I Integrator. Before proceeding with the analysis for this case, it is well to restate exactly what is required of the integrator. In the first place, the addition process must be performed independently at each range position. Furthermore, the integrator output at any range position at a particular instant must be the sum of incoming signal and the signals that occurred at that same range position during the preceding (N-1) repetition periods. It is readily appreciated that this operation requires a signal storage or delay process as well as the addition itself. There are, of course, many known techniques for storing signals and reproducing them at a later time. It will suffice to consider one method which readily suggests itself. This method is shown in Fig. 3.1 in the form of a block diagram. Here the signal progresses from the input terminals on the left through each signal-delay unit in turn. The signals present at the junction points marked (1), (2) . . . (N) are added in the signal-adder unit. This composite signal appears at the output terminals on the right. It will be assumed that the signal-delay units do not attenuate or otherwise distort the signals. The difficulty of constructing a Type-I integrator in the form illustrated in Fig. 3.1 can readily be seen. In a practical radar system, N is seldom less than 10 and is often nearer to 100. The delay units must each produce a delay of the order of a millisecond and transmit signals uniformly over a bandwidth of the order of a megacycle per second. Although this



TYPE-I INTEGRATOR

Fig. 3.1



RESPONSE OF A TYPE-I INTEGRATOR TO A SUCCESSION OF PULSES

Fig. 3.2

delay and bandwidth are obtainable with the ultrasonic delay line, (23-26) the multiplicity of such lines required for the system under consideration is prohibitive. An analysis of the performance characteristics of this ideal system is of interest, however, because the results will serve as a reference with which other systems can be compared.

The operation of this ideal integrator can best be understood by considering first the manner in which it responds to a succession of short-duration pulses, N in number and of amplitude A , each separated in time from the preceding by an interval T . Such an input signal and the resultant signals at each of the junction points are shown in Fig. 3.2. In this example, N has been made equal to five for simplicity. It is seen that there is a linear build-up of the integrated signal followed by a linear decay.

When the input signal is a fluctuating voltage $E_I(t)$, the output signal $E(t)$ is given by

$$E(t) = \sum_{k=1}^N E_I \{t - (k-1)T\} \quad (3.1)$$

An approximate analysis for this case was given by North.⁽²⁾ In his analysis, attention was confined to the case of signals small compared to noise, i.e., $\lambda \ll 1$, and N , the number of pulses integrated, large. In the ensuing analysis, values of N as low as three are of interest. It has already been shown that for $N = 1$, high detection probabilities are obtained only for $\lambda > 1$. Hence approximations based upon an assumption that

$\lambda \ll 1$ can no longer be used. Another approximation employed by North was that of representing the probability density function for the integrated signal by a normal density function having the same mean value and variance as the actual distribution. His justification for this assumption was the Laplace - Liapounoff theorem, more commonly called the central limit theorem. This theorem states that, subject to certain conditions, the probability density function of the sum of a number of statistically independent random variables approaches a normal distribution as the number of constituent variables increases without limit, regardless of the character of their individual probability density functions. The conditions to be satisfied amount to requiring that the ratio of the variance of any one of the constituent variables to the variance of the sum vanishes as the number of constituents increases without limit. (22)

When only a small number of samples is added in the integration process, the accuracy of the results obtained by the use of the normal approximation is open to question. The maximum error would occur if the normal approximation were used to analyze the case of detection with no integration. Since an exact analysis of this case has already been given in Chapter II, a comparison with those results of similar results based upon the use of the normal approximation will give an indication of the magnitude of the errors to be expected for small N.

The normal approximation to the probability density function of the variable Z is

$$p(z; \lambda) = \frac{1}{\sqrt{2\pi} \sigma_Z(\lambda)} \exp - \frac{[z - \bar{z}(\lambda)]^2}{2 \sigma_Z^2(\lambda)} \quad (3.2)$$

where $\bar{z}(\lambda)$ and $\sigma_Z(\lambda)$ are respectively the mean value and standard deviation of Z and both depend upon λ . An expression for the mean value of Z given by North⁽²⁾ becomes in our notation

$$\bar{z}(\lambda) = \sqrt{\frac{\pi}{2}} \epsilon^{-\lambda/2} [(1+\lambda) I_0(\lambda/2) + \lambda I_1(\lambda/2)] \quad (3.3)$$

Here $I_0(x)$ and $I_1(x)$ are modified Bessel functions of the first kind and of orders zero and one respectively.*

A graph of $\bar{z}(\lambda)$ as a function of λ , with λ expressed in decibels, appears in Fig. 3.3. A second curve in which the independent variable is $\sqrt{2\lambda}$ is also shown, $\sqrt{2\lambda}$ being the signal amplitude in units of rms noise amplitude.

This latter curve shows the dependence of the average value of the envelope upon the signal amplitude. The departure from linearity at small signal amplitudes is apparent. In this region the curve becomes parabolic, and hence equal increments of the input signal at two different signal levels do not produce equal increments in the output amplitude. This is the signal-suppression effect frequently referred to in the literature.**

*Tables of the functions $\epsilon^{-x} I_0(x)$ and $\epsilon^{-x} I_1(x)$ are given in reference (27).

**See for example reference (6).

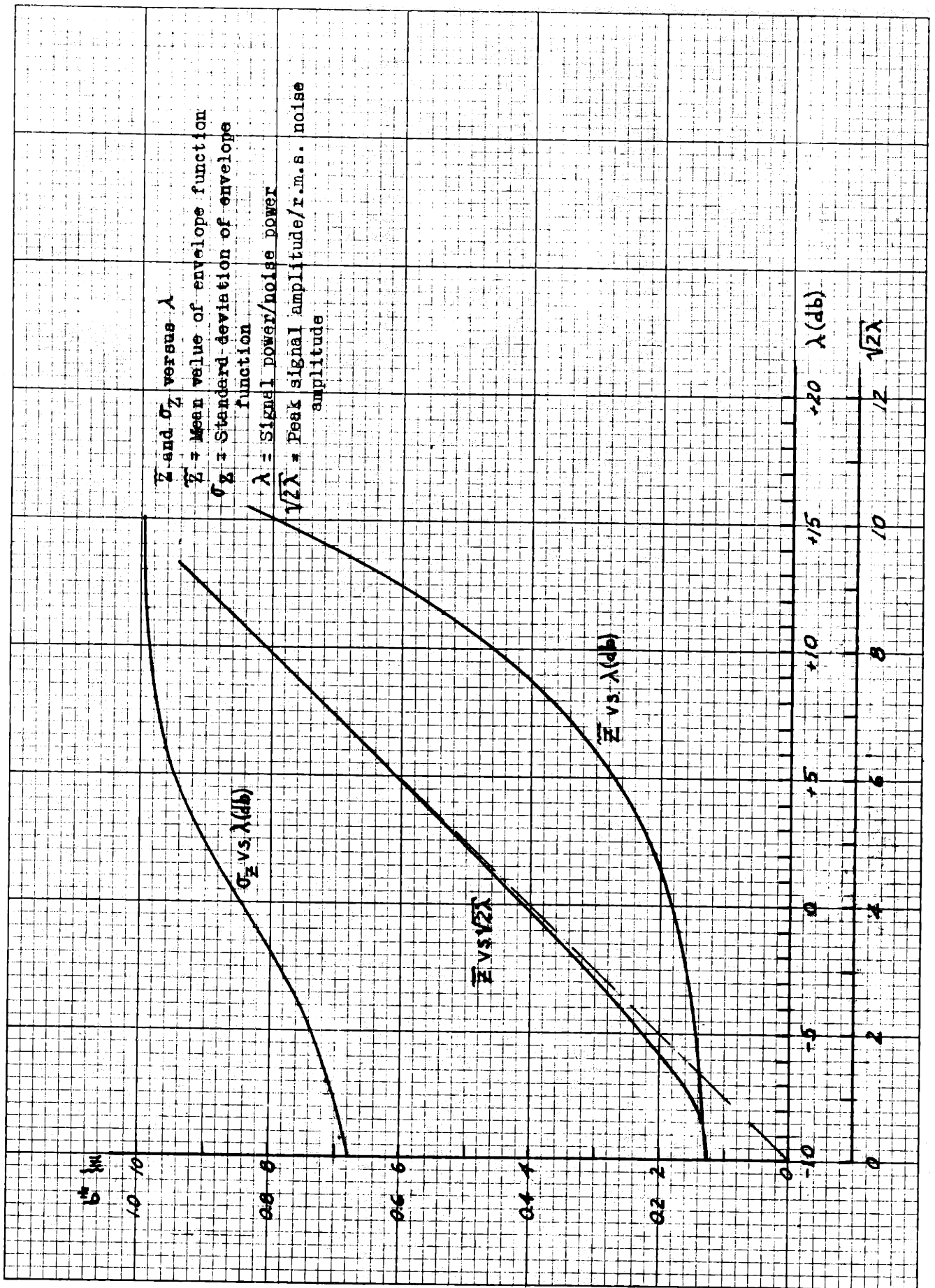


Fig. 3.3

The standard deviation $\sigma_Z(\lambda)$ can be calculated by use of the formula

$$[\sigma_Z(\lambda)]^2 = \overline{Z^2}(\lambda) - [\overline{Z}(\lambda)]^2. \quad (3.4)$$

where $\overline{Z^2}(\lambda)$ is the mean square value of Z, and is given by⁽¹⁴⁾

$$\overline{Z^2}(\lambda) = 2 + 2\lambda \quad (3.5)$$

A graph of $\sigma_Z(\lambda)$ calculated from these equations appears in Fig. 3.3.

Noise and signal probabilities can now be calculated with the help of the curves of Fig. 2.9. In this case, the variable u is defined by

$$u = \frac{Z - \overline{Z}(\lambda)}{\sigma_Z(\lambda)} \quad (3.6)$$

Values of λ_p , the signal strength required for p-percent detectability, were calculated by this method for values of p of 50, 90 and 99 percent, and for values of P_N from 10^{-2} to 10^{-8} . These values of λ_p were found to be consistently smaller than the correct values and differed by an amount ranging from 0.4 to 1.7 decibels. The largest errors occurred with the smallest values of P_N and P_D .

Although the accuracy of this approximation is lower than might be desired, it can be expected to improve rapidly as the number of samples added is increased. The extent to which this is the case will be apparent from a comparison of the results of the analysis with the experimental results.

The normal approximation will now be used to calculate the performance of a detection system based upon the use of

the Type-I integrator of Fig. 3.1 when the input signal is the envelope function $Z(t)$. According to Eq. 3.1, the integrator output signal is given by

$$E(t) = \sum_{k=1}^N Z\{t-(k-1)T\} \quad (3.7)$$

Using a simpler notation, this can be written as

$$E(t) = \sum_{k=1}^N Z_k(t) \quad (3.8)$$

where

$$Z_k(t) = Z\{t-(k-1)T\} \quad (3.9)$$

The normal approximation to the probability density function for E is

$$p(E; \lambda) = \frac{1}{\sqrt{2\pi} \sigma_E(\lambda)} \exp - \frac{[E - \bar{E}(\lambda)]^2}{2 \sigma_E^2(\lambda)} \quad (3.10)$$

where $\bar{E}(\lambda)$ and $\sigma_E(\lambda)$ are related to the corresponding statistics of the variable Z as follows:

$$\left. \begin{aligned} \bar{E}(\lambda) &= N\bar{Z}(\lambda) \\ \sigma_E(\lambda) &= \sqrt{N} \sigma_Z(\lambda) \end{aligned} \right\} \quad (3.11)$$

The following analysis will be simplified by making the change of variable

$$e = [E - \bar{E}(0)] / \sqrt{N} \quad (3.12)$$

Then

$$\bar{e}(\lambda) = \sqrt{N} [\bar{Z}(\lambda) - \bar{Z}(0)] \quad (3.13)$$

and

$$\sigma_e(\lambda) = \sigma_Z(\lambda) \quad (3.14)$$

From which it follows that $\bar{e}(0) = 0$.

The noise and signal probabilities are now expressed by the equations

$$P_N = P\left\{\frac{e_o - \bar{e}(0)}{\sigma_e(0)}\right\} = P\left\{\frac{e_o}{\sigma_e(0)}\right\} \quad (3.15)$$

and

$$P_D = P\left\{\frac{e_o - \bar{e}(\lambda)}{\sigma_e(\lambda)}\right\} = P\left\{\frac{e_o - \sqrt{N}[\bar{Z}(\lambda) - \bar{Z}(0)]}{\sigma_Z(\lambda)}\right\} \quad (3.16)$$

Here $P(u)$ is the function plotted in Fig. 2.9, and e_o is the critical level to be determined by use of Eq. 3.15, in which $\sigma_Z(0) = (2 - \pi/2)^{\frac{1}{2}}$ as may be verified by use of Eqs. 3.3, 3.4 and 3.5.

For the determination of P_D with λ specified, the function $[\bar{Z}(\lambda) - \bar{Z}(0)]$ which appears in Eq. 3.16 is graphed in Fig. 3.4 with λ expressed in decibels as the abscissa. Values of $\sigma_Z(\lambda)$ can be obtained from Fig. 3.3.

More frequently it is of interest to determine the value of λ required to give a specified value of P_D when e_o is such as to give a specified value of P_N . In this case, Eqs. 3.15 and 3.16 can be inverted, and e_o eliminated between the resulting equations to give

$$[\bar{Z}(\lambda) - \bar{Z}(0)] = [\sigma_Z(0)P^{-1}(P_N) - \sigma_Z(\lambda)P^{-1}(P_D)]/\sqrt{N} \quad (3.17)$$

where $P^{-1}(x)$ is the inverse, not the reciprocal, of $P(x)$. Because λ appears in both members of this equation, and because Eq. 3.3 for $\bar{Z}(\lambda)$ is a transcendental equation, Eq. 3.17 is most easily solved by a method of successive approximations as follows: Choose a trial value of λ , read $\sigma_Z(\lambda)$ from Fig. 3.3 and calculate the value of the right-hand member

ENGINEERING DESIGN CO.

FIG. 3.4

2 CYCLES X 20 DIVISIONS PER INCH

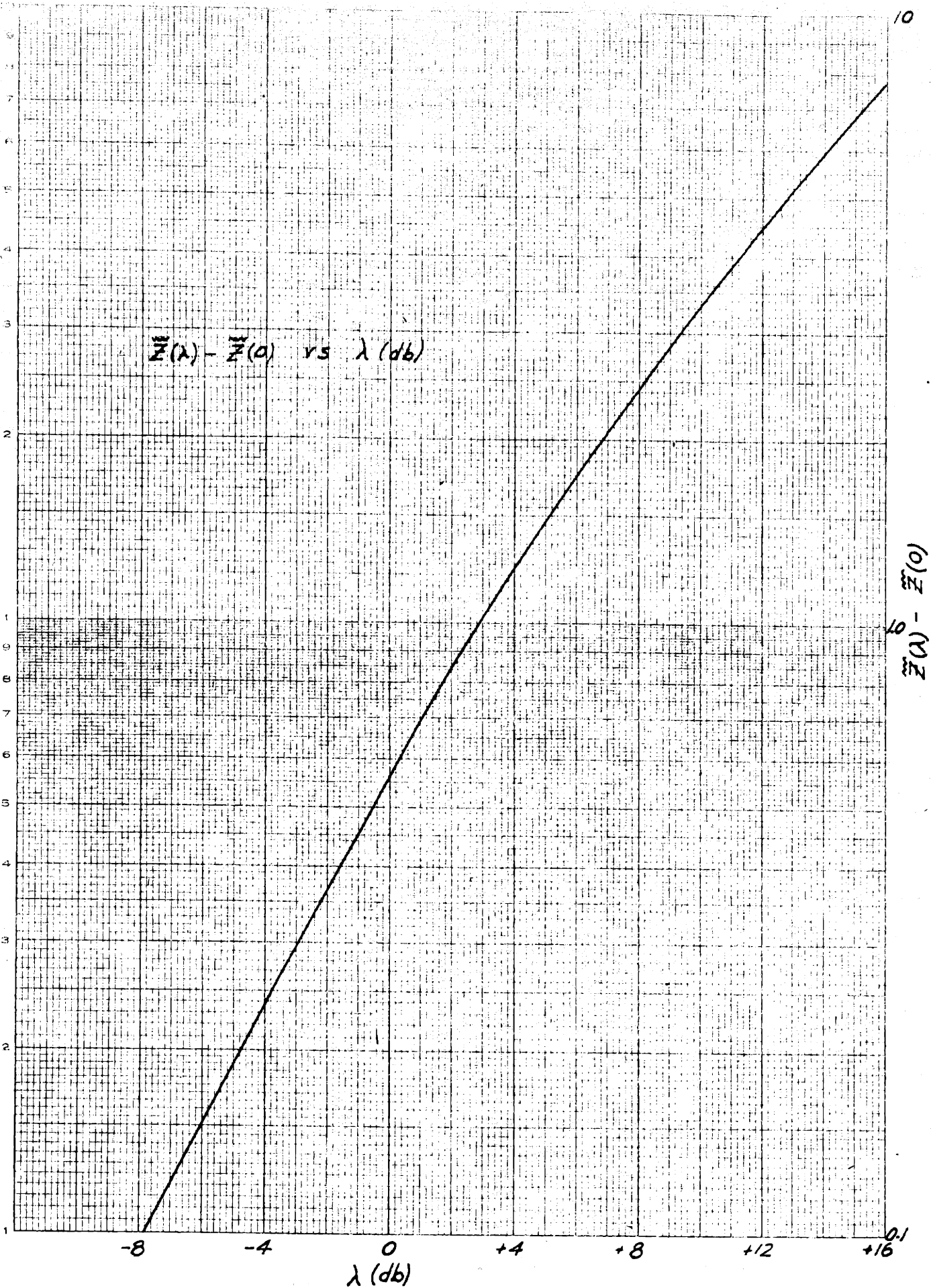


Fig. 3.4

of Eq. 3.17. Assigning this value to the function $[Z(\lambda)-Z(0)]$, read from Fig. 3.4 the corresponding value of λ . If this new value of λ differs from the trial value, use the new value as a trial value and repeat the process until agreement is obtained. Because $\sigma_z(\lambda)$ varies much more slowly with λ than does $[Z(\lambda)-Z(0)]$, this process converges very rapidly, one or two trials being sufficient in most cases.

The curves of Figs. 3.5, 3.6 and 3.7 were computed by this method. In each of these figures the signal strength, λ_p , required for p-percent signal detectability for values of p of 10, 50, 90 and 99 percent is plotted as a function of N, the number of pulses integrated. Each figure corresponds to a different value of P_N , the values being 10^{-4} , 10^{-6} , and 10^{-8} respectively. For the points at $N = 1$, the accurate λ_p values calculated in Chapter II were used in preference to approximate values that would be obtained by use of the normal approximation. Also, the correction for the effect of the radio-frequency filter on pulse amplitude, as discussed in Chapter II, has been applied.

A comparison of the three sets of curves shows that the only significant difference between them is a progressive shift in the direction of increasing λ_p as P_N is decreased, as is to be expected. However, this shift is of the order of only two decibels for a change of P_N from 10^{-4} to 10^{-8} .

The circled points in Figs. 3.4 and 3.5 show the results of the experimental study of the Type-I integrator. It is

EUGENE DIETZGEN CO.
MADE IN U. S. A.

NO. 300-110 METAL-DIEZGEN DRAFT PAPER
SEMI-LOGARITHMIC
4 CYCLES X 10 DIVISIONS PER INCH **Fig. 3.5**

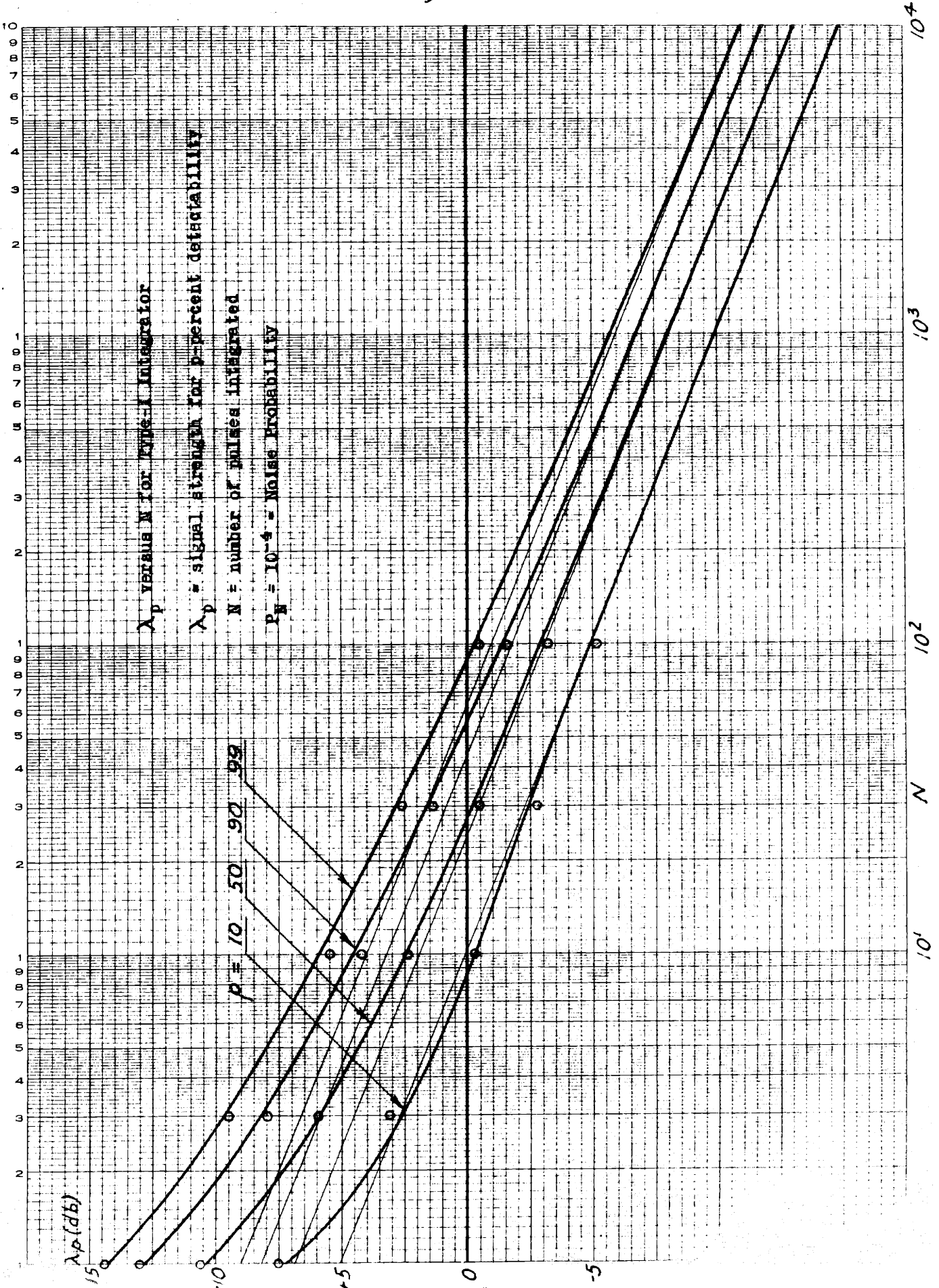


Fig. 3.5

Fig. 3.6

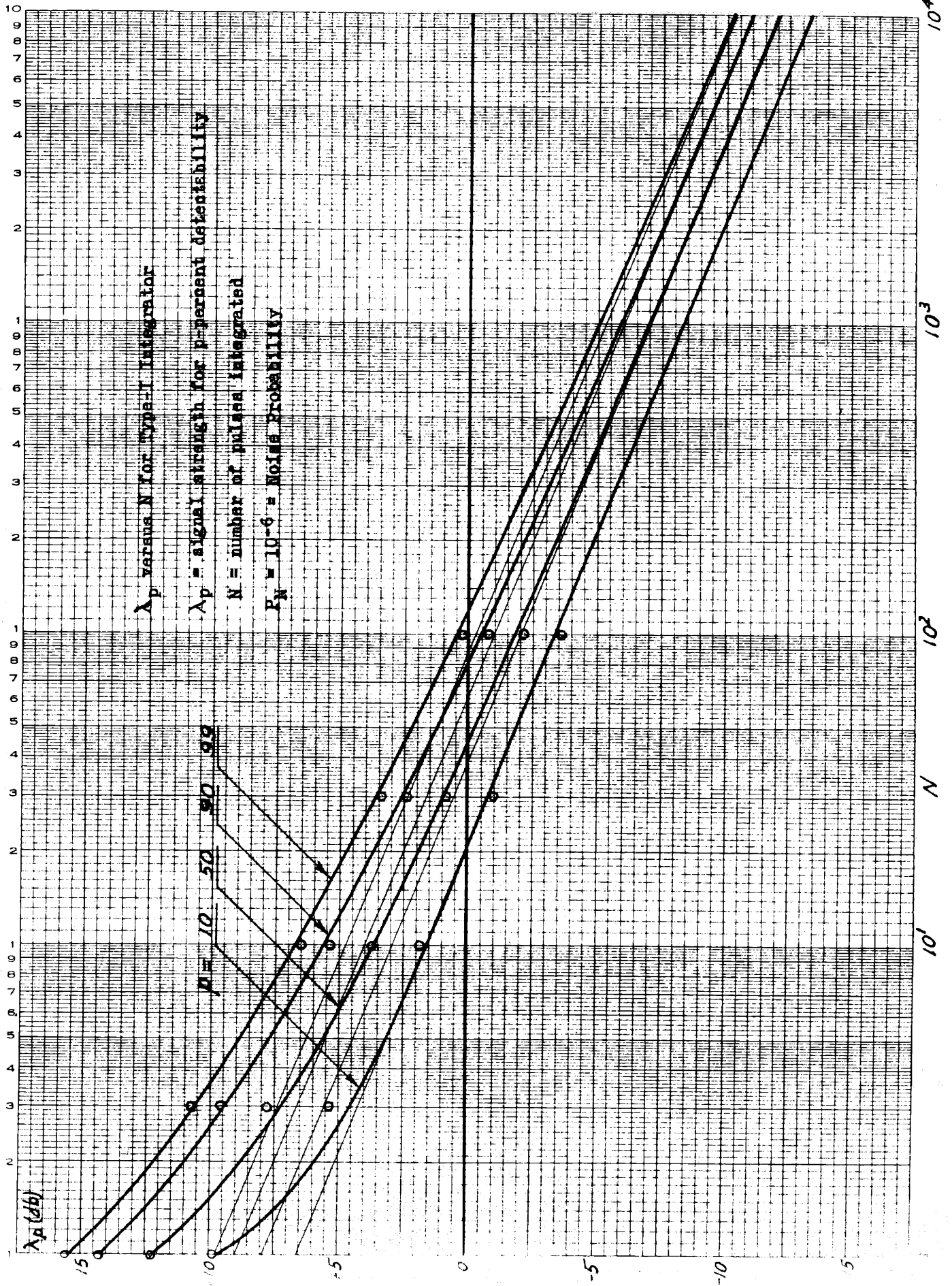


Fig. 3.6

EUGENE DIETZGEN CO.
MADE IN U. S. A.

TYPE SEMI-LOGARITHMIC GRAPH PAPER
SEMI-LOGARITHMIC
4 CYCLES X 10 DIVISIONS PER INCH
Fig. 3.7

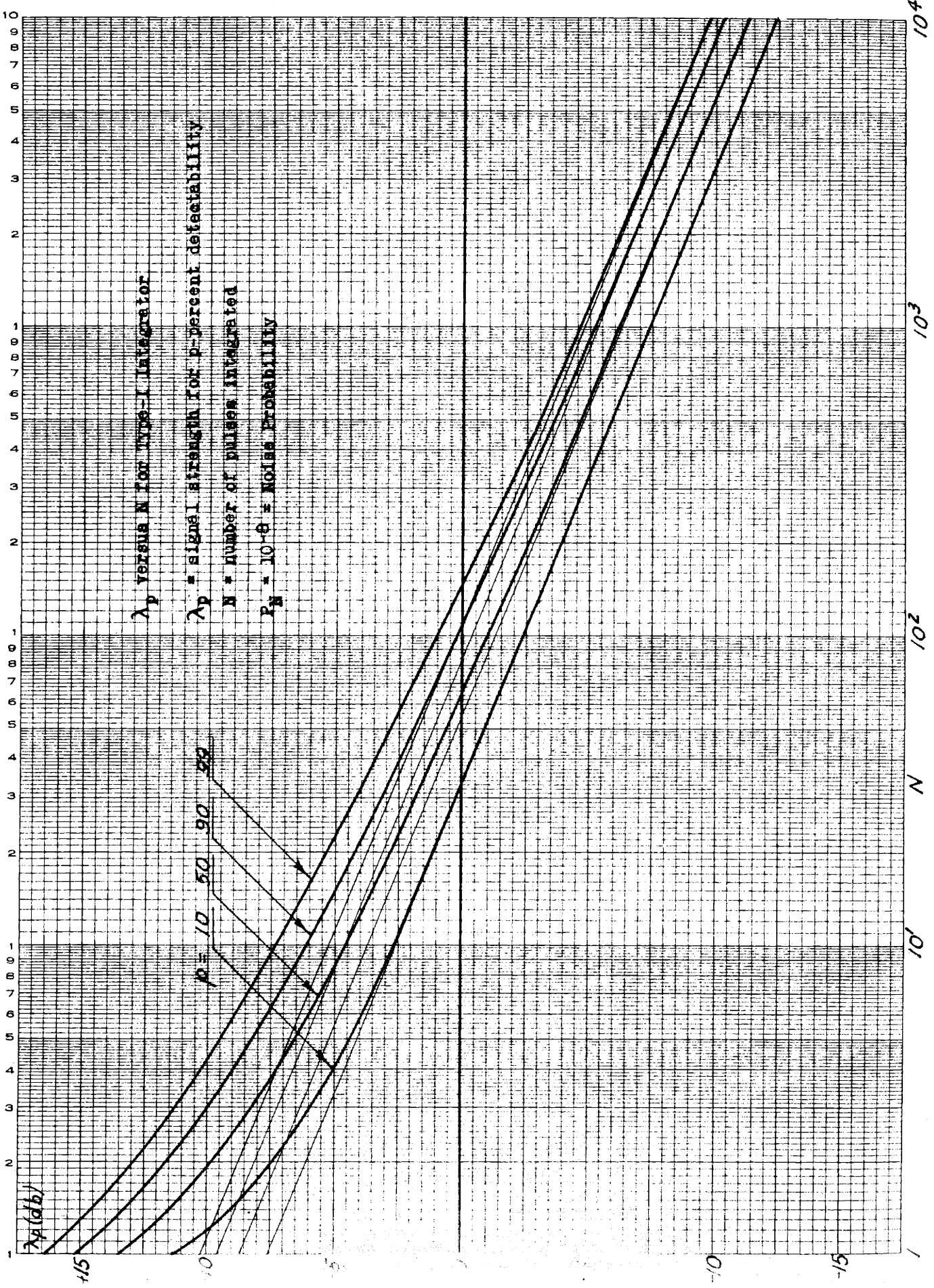


Fig. 3.7

seen that agreement to within one-half decibel is obtained at all points except for the points at $N = 3$, where the discrepancy approaches one decibel for the smallest values of p . It is believed that this discrepancy is a result of the error inherent in the use of the normal approximation with small values of N . However, the close agreement between theory and experiment obtained for larger values of N indicates that the normal approximation is sufficiently accurate for all practical purposes.

It is of interest to investigate the dependence of λ_p upon N for N very large. In this region, $\lambda_p \ll 1$. Following North,⁽²⁾ $\tilde{Z}(\lambda)$ and $\sigma_Z(\lambda)$ are accurately represented when λ is sufficiently small by

$$\tilde{Z}(\lambda) = \tilde{Z}(0) + \lambda \left. \frac{d\tilde{Z}(\lambda)}{d\lambda} \right|_{\lambda=0} \quad (3.18)$$

$$\sigma_Z(\lambda) = \sigma_Z(0) + \lambda \left. \frac{d\sigma_Z(\lambda)}{d\lambda} \right|_{\lambda=0} \quad (3.19)$$

Differentiation of Eq. 3.3 with respect to λ gives

$$\begin{aligned} \frac{d\tilde{Z}(\lambda)}{d\lambda} = & -\frac{1}{2}\tilde{Z}(\lambda) + \sqrt{\frac{\pi}{2}} e^{-\lambda/2} \left\{ -\frac{1}{2} I_1(\lambda/2) + I_0(\lambda/2) - \right. \\ & \left. -\frac{\lambda}{2} I_1(\lambda/2) + I_1(\lambda/2) + \frac{\lambda}{2} [I_0(\lambda/2) - \frac{2}{\lambda} I_1(\lambda/2)] \right\} \end{aligned} \quad (3.20)$$

Placing $\lambda = 0$ and noting that $\tilde{Z}(0) = \sqrt{\frac{\pi}{2}}$, $I_0(0) = 1$ and $I_1(0) = 0$ gives

$$\left. \frac{d\tilde{Z}(\lambda)}{d\lambda} \right|_{\lambda=0} = \frac{1}{2} \sqrt{\frac{\pi}{2}} \quad (3.21)$$

Differentiation of Eq. 3.4 with respect to λ gives

$$2\sigma_Z(\lambda) \frac{d\sigma_Z(\lambda)}{d\lambda} = \frac{d}{d\lambda} \tilde{Z}^2(\lambda) - 2\tilde{Z}(\lambda) \frac{d\tilde{Z}(\lambda)}{d\lambda} \quad (3.22)$$

From Eq. 3.5 there is obtained $\frac{d}{d\lambda} \tilde{z}^2(\lambda) = 2$. Substituting this into Eq. 3.22 and setting $\lambda = 0$ gives

$$\left. \frac{d}{d\lambda} \sigma_z(\lambda) \right|_{\lambda=0} = \frac{1}{2} \sqrt{2 - \pi/2} \quad (3.23)$$

Hence Eqs. 3.18 and 3.19 become

$$\tilde{z}(\lambda) = \frac{1}{2} \sqrt{\frac{\pi}{2}} (1 + \lambda/2) \quad (3.24)$$

$$\sigma_z(\lambda) = \frac{1}{2} \sqrt{2 - \pi/2} (1 + \lambda/2) \quad (3.25)$$

Substitution of these expressions in Eq. 3.17 gives

$$\frac{\lambda}{2} \sqrt{\frac{\pi}{2}} = \sqrt{2 - \pi/2} \left[P^{-1}(P_N) - (1 + \lambda/2) P^{-1}(P_D) \right] / \sqrt{N}$$

Neglecting the quantity $\lambda/2$ appearing within the brackets in comparison to unity, this equation becomes

$$\lambda_p = \frac{2\sqrt{4/\pi - 1}}{\sqrt{N}} \left[P^{-1}(P_N) - P^{-1}(P_D) \right] \quad (3.26)$$

Therefore, for large N the threshold signal power varies inversely as the square root of N. The asymptotes appearing in Figs. 3.5, 3.6 and 3.7 are plotted in accordance with Eq. 3.26. For values of N approaching unity, the departure from the asymptote is large and the dependence of λ_p upon N becomes more nearly an inverse first-power law.

3.2 Type-II Integrator. This designation is used for the class of integrators characterized in Chapter I by an exponential signal decay process. In analyzing this type of integrator, it is well to have in mind a specific physical system. For this purpose, the delayed-feedback integrator is considered since this was the type used in the experiments

to be reported in a later chapter. A block diagram of this system is shown in Fig. 3.8.

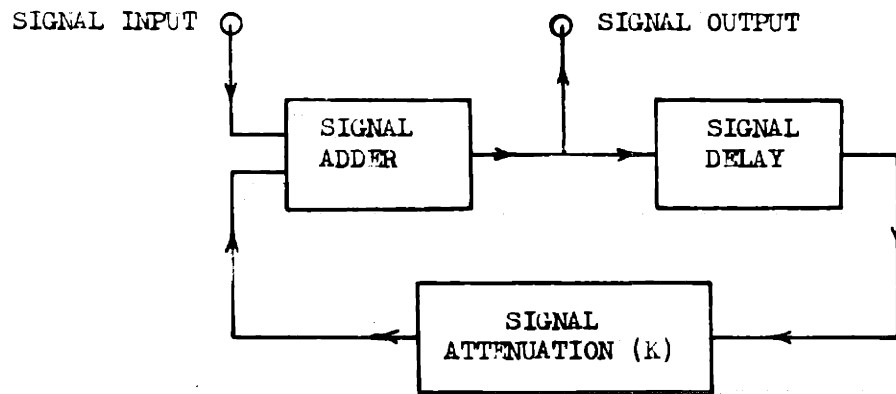
In the following analysis, it is assumed that the signal-delay unit does not attenuate or otherwise distort the signal and also that the feedback attenuator and adder units introduce no distortion.

The operation of this system can be illustrated by considering its response to a succession of rectangular pulses of constant amplitude. This response is illustrated in Fig. 3.9. For this example, an attenuation factor of $K = 0.8$ is used. The upper trace represents the input signal. The next five numbered traces each show the response that would result if the correspondingly numbered pulse in the input signal occurred alone. The bottom trace is the sum of the five traces immediately above, and represents the output signal. The exponential nature of the signal build-up and decay process is evident.

In view of this exponential behavior, it is appropriate to define an integrator time constant, N_1 , as the number of repetition periods in which a single signal pulse will be attenuated by a factor of e^{-1} or one napier. This time constant is related to the attenuation factor, K , by the relation

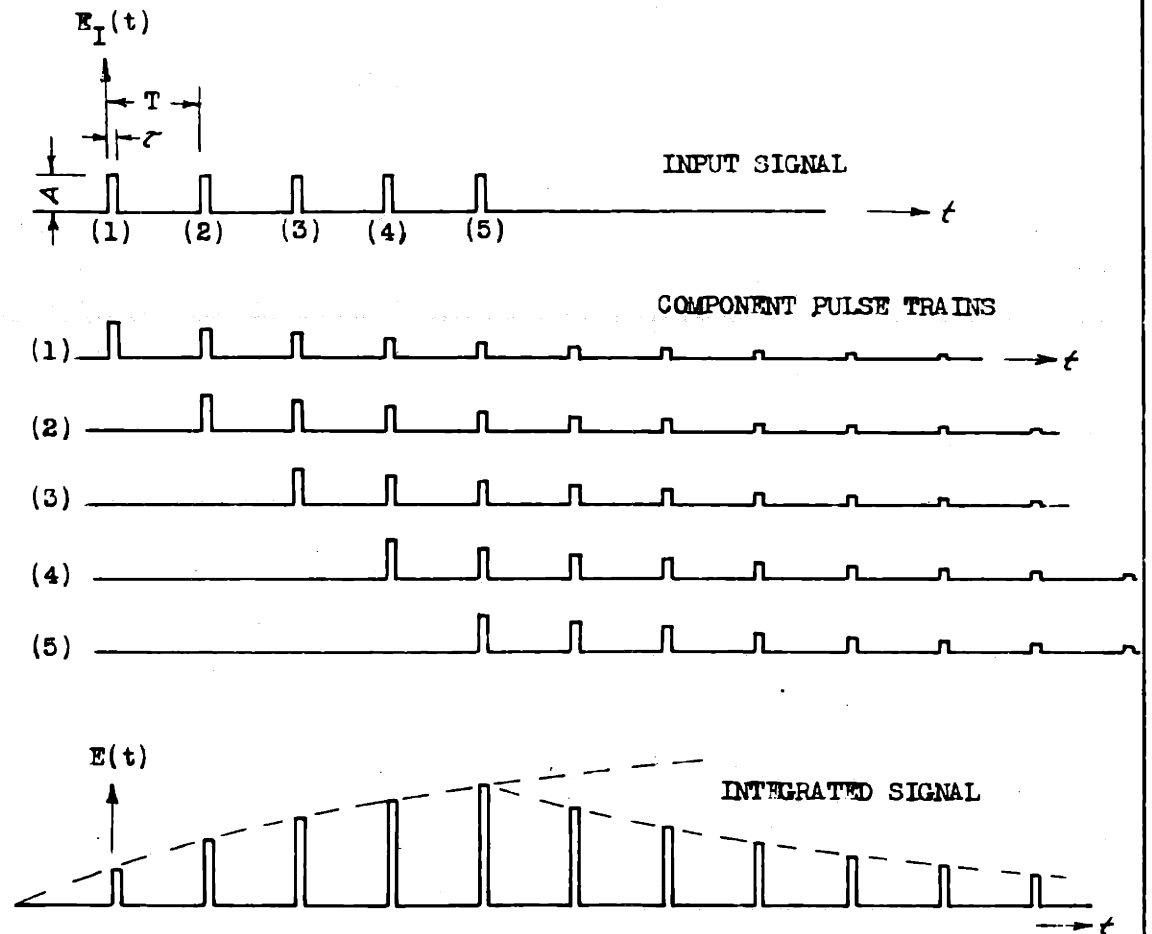
$$K^{N_1} = e^{-1}$$

or
$$N_1 = -1/\log_e K. \quad (3.27)$$



TYPE-II INTEGRATOR

Fig. 3.8



RESPONSE OF A TYPE-II INTEGRATOR TO A SUCCESSION OF PULSES

Fig. 3.9

The response of this system to an arbitrary input function $E_I(t)$ is seen to be

$$E(t) = E_I(t) + KE_I(t-T) + K^2E_I(t-2T) + \dots \quad (3.28)$$

$$= \sum_{k=0}^{\infty} E_I\{t-KT\} \quad (3.29)$$

In the following statistical analysis, the normal approximation for the probability density function of $E(t)$ will again be used. Hence the mean value and standard deviation of $E(t)$ must be found, both for the case of a noise-only range position and for a signal position. Actually, it is sufficient to consider the latter case since it includes the former as a special case. For this latter case, the first N terms of the series of Eq. 2.29 will be influenced by the presence of the signal while the remaining terms will not. Hence it is appropriate to rewrite Eq. 2.29 as follows:

$$E(t;\lambda) = \sum_{k=0}^{N-1} K^k E_I\{(t-kT);\lambda\} + \sum_{k=N}^{\infty} K^k E_I\{(t-kT);0\}$$

The signal strength λ has been explicitly included in the argument since the mean and standard deviation of E is a function of λ . All of the terms of the last summation in this equation contain a common factor K^N . Hence the equation can be rewritten as follows:

$$E(t;\lambda) = \sum_{k=0}^{N-1} K^k E_I\{(t-kT);\lambda\} + K^N \sum_{k=0}^{\infty} K^k E_I\{[t-(k+N)T];0\} \quad (3.30)$$

The mean value and standard deviation of $E(t;\lambda)$ will now be calculated with the help of the following elementary propositions from statistical theory:

(1) The mean value of a sum of random variables is equal to the sum of the means of the individual variables.

(2) The variance of the sum of statistically independent random variables is equal to the sum of the variances of the individual variables.

If $E(t;\lambda)$ is the envelope function, $Z(t)$, the condition of independence is satisfied. Hence it follows from Eq. 3.30 that

$$\overline{E}(\lambda) = \frac{1-K^N}{1-K} \overline{Z}(\lambda) + \frac{K^N}{1-K} \overline{Z}(0) \quad (3.31)$$

$$\sigma_E^2(\lambda) = \frac{1-K^{2N}}{1-K^2} \sigma_Z^2(\lambda) + \frac{K^{2N}}{1-K^2} \sigma_Z^2(0) \quad (3.32)$$

where use has been made of the following formulas for the sum of a geometric series:

$$1 + K + K^2 + \dots + K^{N-1} = \frac{1-K^N}{1-K} ; |K| < 1$$

and $\lim_{K \rightarrow \infty} \frac{1-K^N}{1-K} = \frac{1}{1-K}$

From this point on, the analysis would follow exactly the same lines as for the case of the Type-I integrator except that the attenuation factor K now appears as an additional parameter whose effect upon the performance of the system requires investigation. This was done experimentally with the objective of establishing an optimum value for K . These experiments will be described in a later chapter. However, the conclusions with regard to the optimum value of K can be used to simplify considerably the ensuing analysis. It is as follows: K should be such that the integrator time constant N_1 defined

by Eq. 3.27 is equal to N, the number of pulses in the pulse train, that is, $K = e^{-1/N}$

Substituting this expression for K into Eqs. 3.31 and 3.32, there results

$$\bar{E}(\lambda) = \frac{1}{1 - e^{-1/N}} \left[(1 - e^{-1}) \bar{Z}(\lambda) + e^{-1} \bar{Z}(0) \right] \quad (3.33)$$

and

$$\sigma_E(\lambda) = \frac{1}{(1 - e^{-2/N})^{1/2}} \left[(1 - e^{-2}) \sigma_Z^2(\lambda) + e^{-2} \sigma_Z^2(0) \right]^{1/2} \quad (3.34)$$

To simplify the analysis, it is expedient to introduce a new variable, e, defined as follows:

$$e = (E - \bar{E}(0)) (1 - e^{-2/N})^{1/2} \quad (3.35)$$

Then

$$\bar{e}(\lambda) = (\bar{E}(\lambda) - \bar{E}(0)) (1 - e^{-2/N})^{1/2} \quad (3.36)$$

and

$$\sigma_e(\lambda) = \sigma_E(\lambda) (1 - e^{-2/N})^{1/2} \quad (3.37)$$

Using Eqs. 3.33 and 3.34, there is obtained

$$\bar{e}(\lambda) = \left(\frac{1 + e^{-1/N}}{1 - e^{-1/N}} \right)^{1/2} (1 - e^{-1}) \left[\bar{Z}(\lambda) - \bar{Z}(0) \right] \quad (3.38)$$

and

$$\sigma_e(\lambda) = \left[(1 - \bar{e}^2) \sigma_Z^2(\lambda) + e^{-2} \sigma_Z^2(0) \right]^{1/2} \quad (3.39)$$

The radical in Eq. 3.38 can be rewritten as $[\text{Coth}(1/2N)]^{1/2}$, and for N even as small as 3 this can be replaced by $\sqrt{2N}$ with an error of only one-quarter percent. Hence for all practical purposes,

$$\bar{e}(\lambda) = \sqrt{2N} (1 - e^{-1}) \left[\bar{Z}(\lambda) - \bar{Z}(0) \right] \quad (3.40)$$

which can be evaluated by use of Fig. 3.4 for the factor in

brackets. Eq. 3.39 for $\sigma_e(\lambda)$ is plotted in Fig. 3.10.

The noise and signal probabilities can now be expressed as follows:

$$P_N = P \left\{ \frac{e_0 - \bar{e}(0)}{\sigma_e(0)} \right\} = P \left\{ \frac{e_0}{\sigma_e(0)} \right\} \quad (3.41)$$

$$P_D = P \left\{ \frac{e_0 - \bar{e}(\lambda)}{\sigma_e(\lambda)} \right\} \quad (3.42)$$

where $\bar{e}(\lambda)$ and $\sigma_e(\lambda)$ are given by Eqs. 3.38 and 3.39 and e_0 is the critical level. These equations can be used to obtain curves of λ_p versus N in the same manner as that used in the preceding example. Also, the method used in the preceding example to obtain the equation of the asymptote can be applied here, and the result is the equation

$$\lambda_p = \frac{2 \sqrt{(4/\pi) - 1}}{\sqrt{2N} (1 - \epsilon^{-1})} [P^{-1}\{P_N\} - P^{-1}\{P_D\}] \quad (3.43)$$

It is seen that this equation differs from Eq. 3.26, the equation of the asymptote in the previous example, only in the factor $1/\sqrt{2}(1 - \epsilon^{-1}) = 1.119$. Hence for given values of N , P_N and P_D , the asymptote for the Type-II integrator in a plot as used in Figs. 3.5, 3.6 and 3.7 will lie above that for the Type-I integrator by $10 \log 1.119 = 0.487$ decibel.

Because the difference in performance of these two types of integrators is so slight in the limiting case of large N , it is hardly to be expected that there will be any significant difference for smaller values of N . To check this point, curves of λ_p versus N were calculated for $P_N = 10^{-6}$ and $p = 10, 50, 90$ and 99. These curves, together with their asymptotes, are

PRINTED IN U.S.A.



Fig. 3.10

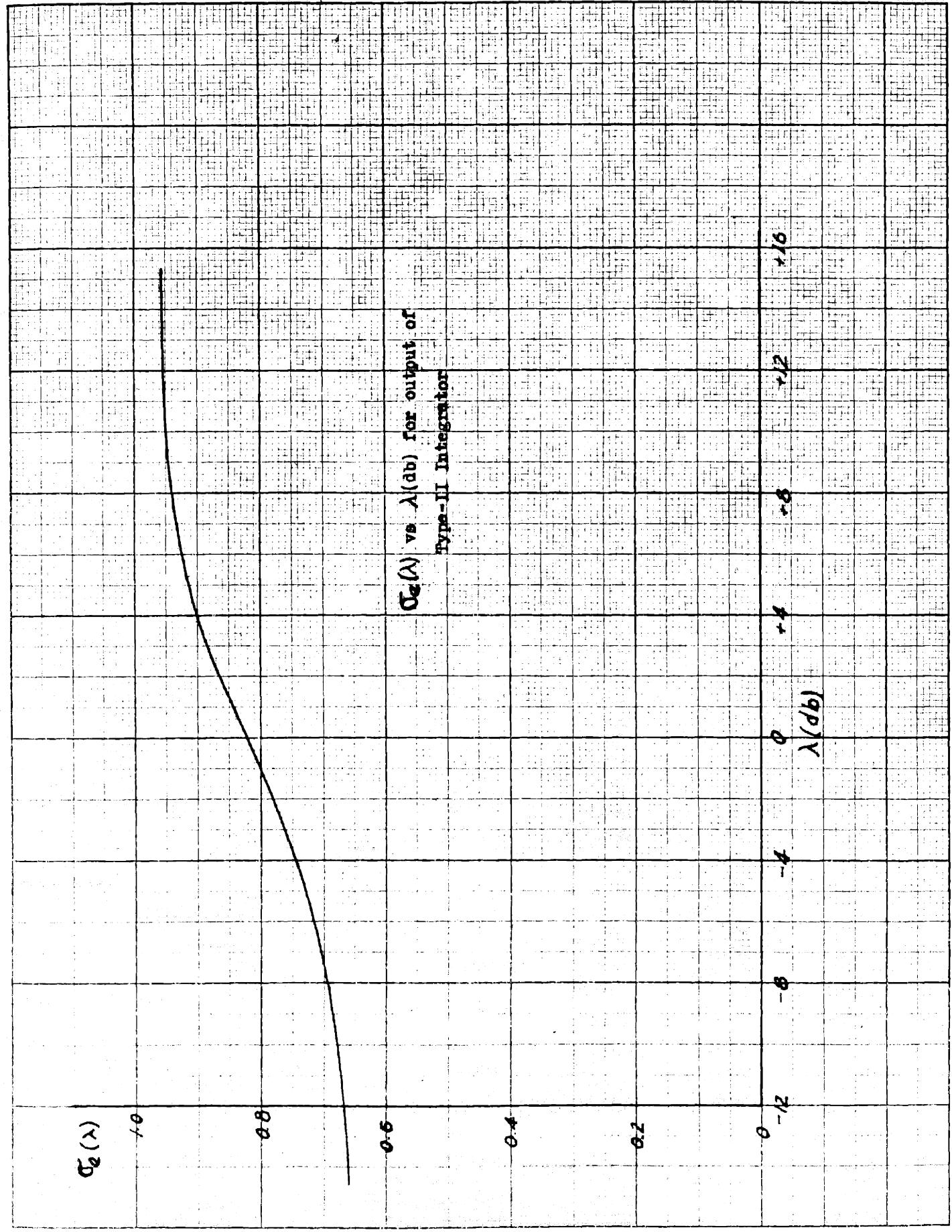


Fig. 3.10

plotted in Fig. 3.11. The corresponding results for the Type-I integrator for $p = 50$ are shown by the dashed-line curve in this figure. It is seen that a difference of approximately one-half decibel is maintained over the entire range of N except for N less than three. The circled points are again the results of the experimental study and are seen to be in good agreement with the theoretical results with the exception, as in the previous example, of the points for $N = 3$ and the smallest value of p .

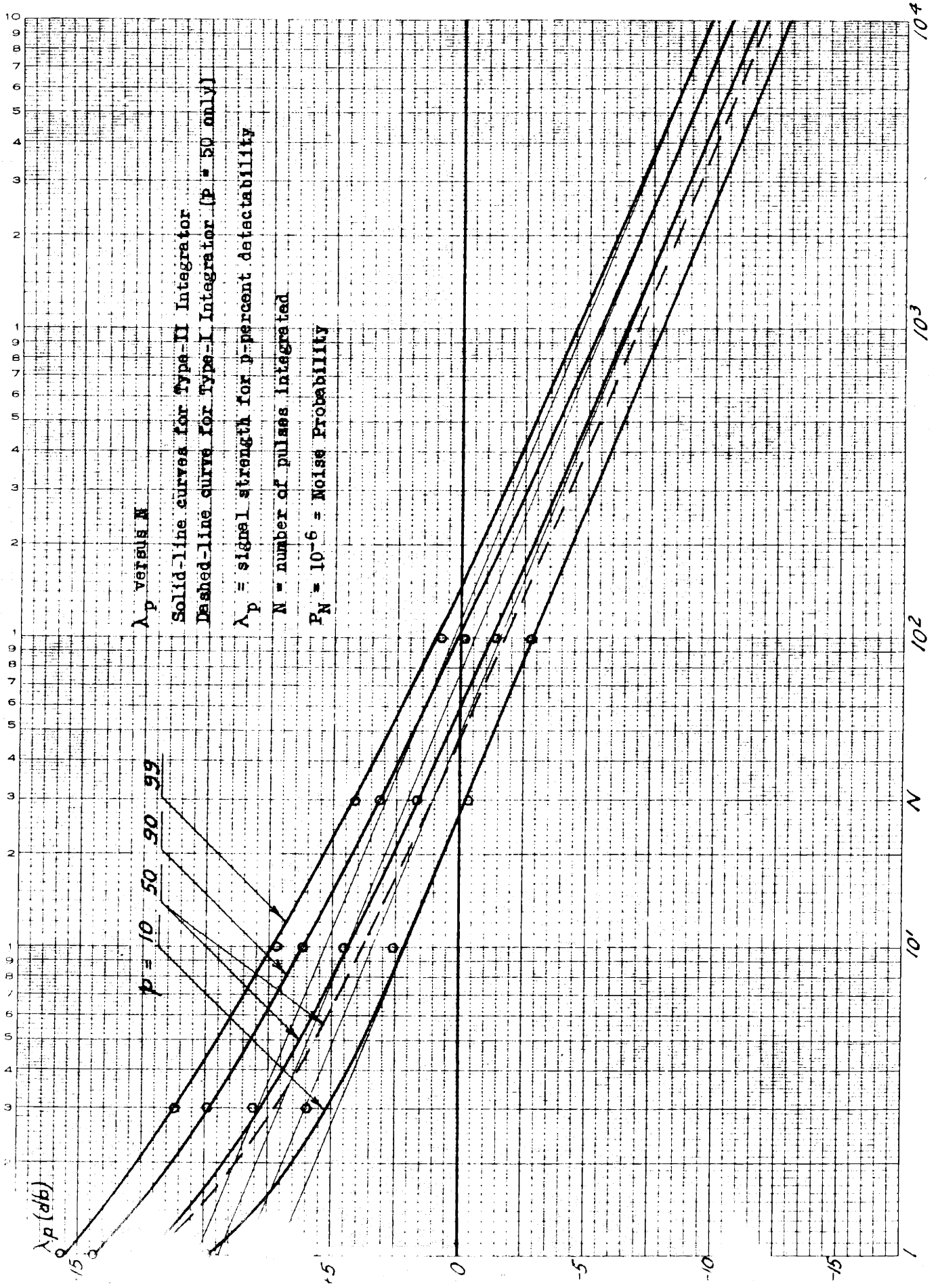


Fig. 3.11

CHAPTER IV

DESCRIPTION OF THE EXPERIMENTAL SYSTEM

The description of the experimental system is divided into two principal parts: (1) a general description of the complete system, and (2) a more detailed discussion of the performance characteristics and special features of the individual units of the system.

4.1 General Description. The principal components of the experimental system are indicated by the block diagram of Fig. 4.1. Waveforms representative of the functions of these units are shown in Fig. 4.2. Fig. 4.3 is a general view of the experimental system. All of the functions of this system are timed with reference to the master trigger pulses generated in the signal integration unit. These trigger pulses are shown as trace (1) in Fig. 4.2. In this experimental system, the repetition frequency of these master trigger pulses is approximately 2,000 per second. This frequency is fixed by the length of the signal-delay unit in the signal integrator.

The trigger and gate generator appearing in the upper left-hand corner of Fig. 4.1 receives the master trigger pulses. By means of a series of flip-flop counters, trigger pulses are generated at a repetition frequency equal to a sub-multiple of the master trigger frequency. These trigger pulses are shown in Fig. 4.2 as trace (2) and are called pulse train triggers. They are applied to the pulse train

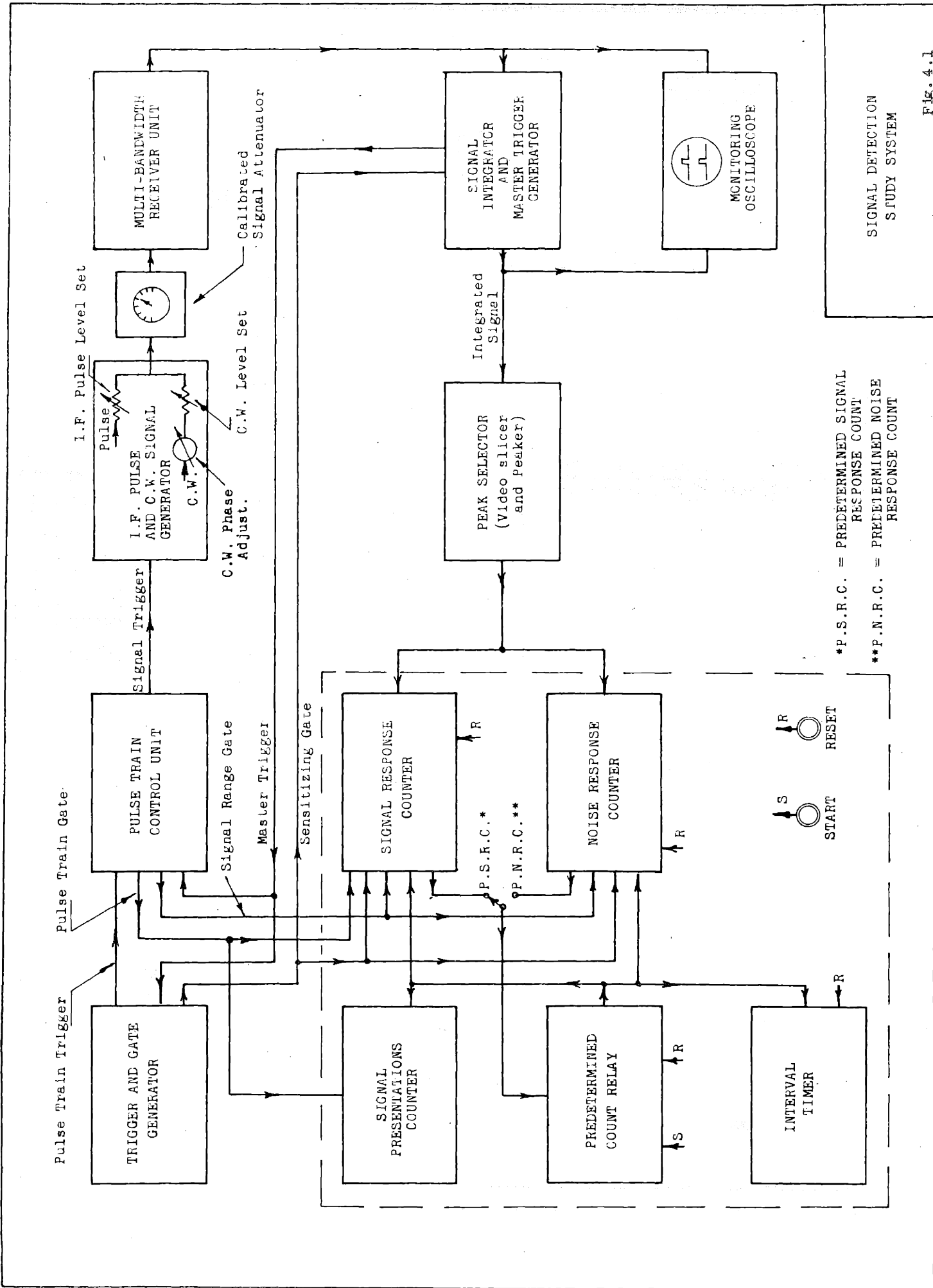
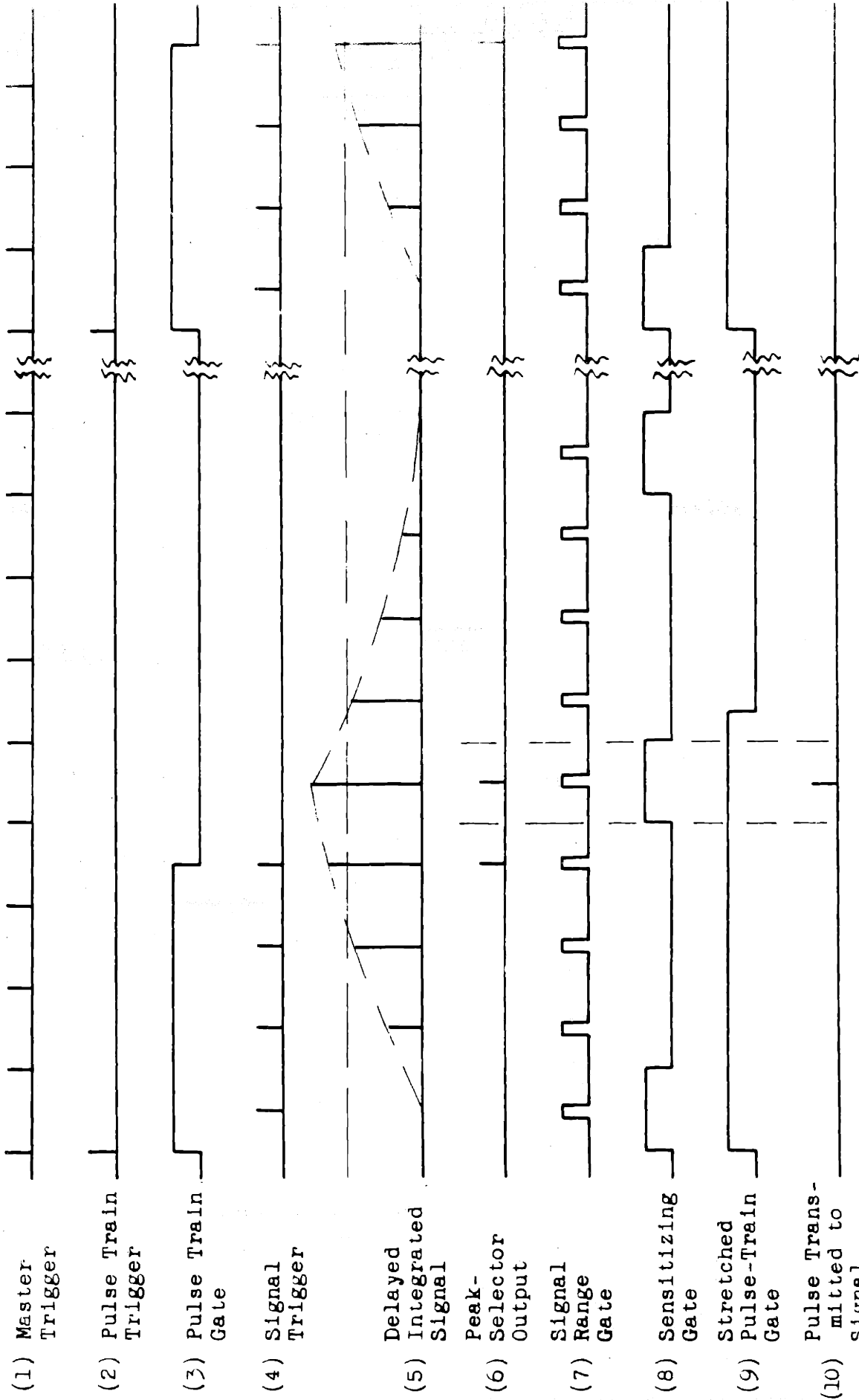
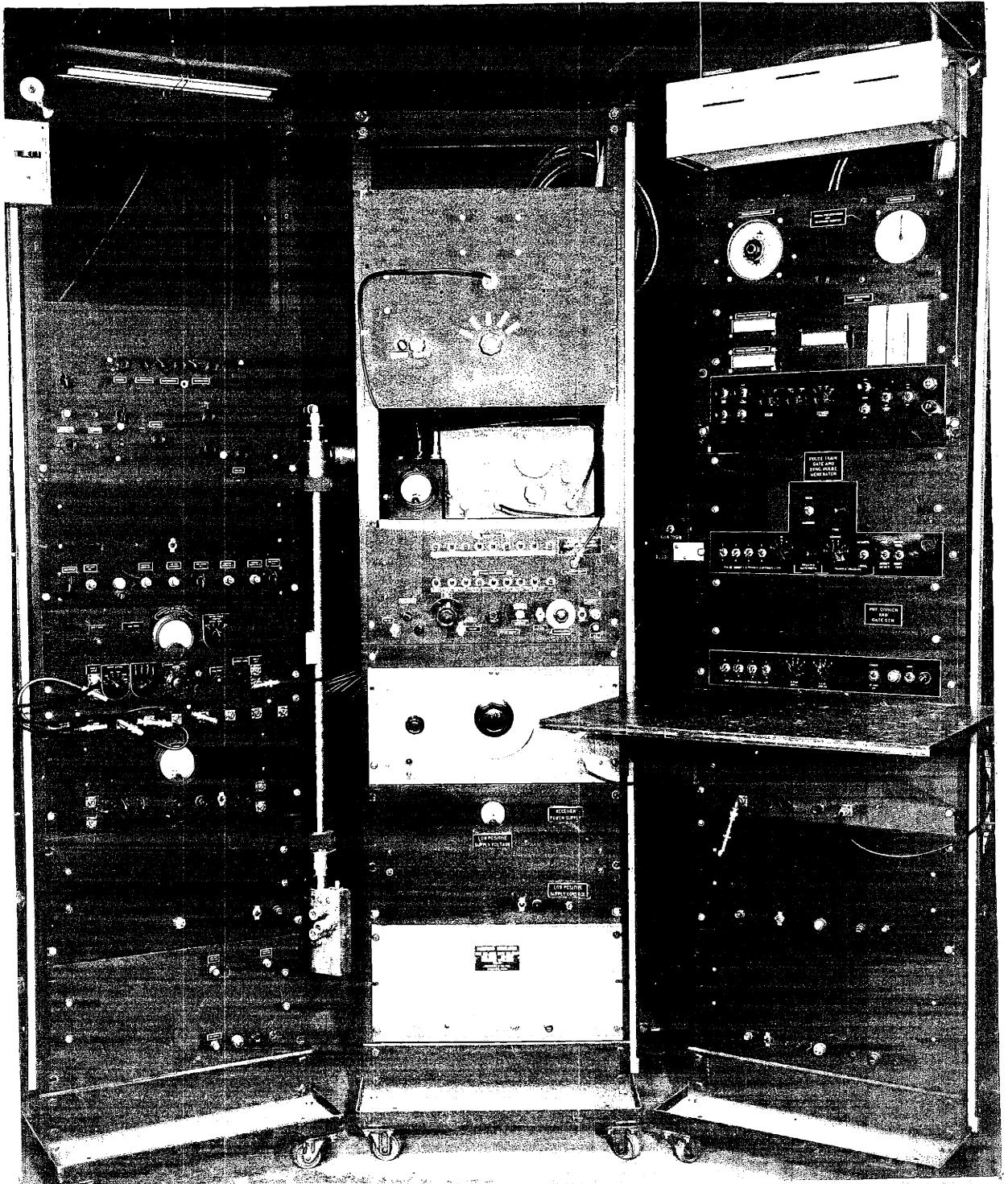


Fig. 4.1

SIGNAL DETECTION STUDY SYSTEM



TIMING DIAGRAM FOR SIGNAL
DETECTION STUDY SYSTEM



GENERAL VIEW OF EXPERIMENTAL SYSTEM

Fig. 4.3

control unit where they initiate the pulse train gate as shown in trace (3). Also generated within the pulse train control unit are trigger pulses delayed in time relative to the master trigger pulses, the delay being adjustable according to the desired range of the target signal. These delayed pulses are gated by the pulse train gate to produce the signal trigger, shown in trace (4). These gated trigger pulses are transmitted to the intermediate frequency (i.f.) pulse and c.w. signal generator and are also applied to a pulse counter circuit in the pulse train control unit. At the end of a predetermined count, this counter produces an impulse which terminates the pulse train gate, thus terminating the signal trigger pulse train. This sequence of events is repeated following each pulse train trigger pulse.

The i.f. signal pulse trains thus generated are transmitted to the receiver via the calibrated signal attenuator which sets the level of the signals applied to the receiver. The receiver is typical of pulse-type receivers used in practice. Sufficient amplification is available to raise the input circuit noise to the maximum level usable at its output terminals. The input signal level is calibrated in terms of the receiver noise level by a method to be described later. A choice of i.f. bandwidths from 0.1 megacycle to 10 megacycles is available in this receiver, although the smallest bandwidth was used in most of the experiments. This was done in order to minimize the effect of bandwidth limitations in other parts of the system.

The signal from the receiver is applied to the signal integrator unit where a choice of two modes of operation is available, one for the study of Type-I integration and the other for the study of Type-II integration. The output signal for the latter mode is represented by trace (5) in Fig. 4.2, where the envelope of the integrated signal pulse train is indicated by the dashed curve. The integrated signal is applied to the peak selector unit which produces an impulse whenever the instantaneous amplitude of the integrated signal passes through a predetermined voltage level, or slicing level, with positive slope. This slicing level is indicated in trace (5) by the horizontal dashed line. The resulting peak selector output impulses are shown in trace (6). Such impulses can result both from signal pulse trains and from chance occurrences of large noise pulses. It is desired to count separately the signal responses and noise responses. Therefore two sets of counters are provided, a signal response counter and a noise response counter. Because these counters can not in themselves distinguish between signal and noise responses, they are gated by means of a signal range gate pulse which straddles the signal pulse, as indicated by trace (7) of Fig. 4.2. The signal response counter is sensitized only during this pulse, while the noise response counter is simultaneously disabled. This signal range gate pulse is generated in the pulse train control unit.

Because it is desired to observe the integrated signal only at the time that it reaches its maximum value, the signal

response counter is also gated by sensitizing gate pulses which are represented by trace (8) in Fig. 4.2. These gate pulses are generated in the trigger and gate generator. Their spacing is made equal to the pulse train duration by appropriately setting the division ratio of an adjustable pulse-repetition-frequency divider. The divider output pulse defines the leading edge of the sensitizing gate, and the next master trigger pulse defines the trailing edge. It is noted that the sensitizing gate occurs one repetition period after the end of the signal pulse train. Hence, in order to obtain coincidence between the maximum-amplitude portion of the pulse train and the sensitizing gate, the signal at the output of the integrator delay line is used, this being delayed by one repetition period relative to the pulse train at the input. This same sensitizing pulse is applied to the noise response counter to prevent noise pulse counting except during the repetition period when signal responses can be counted. In order to prevent noise pulses occurring within the signal range gate from producing spurious signal response counts, the signal response counter is sensitized only during the pulse train gate. For this purpose, the pulse train gate of trace (3) is stretched so as to overlap the sensitizing gate, as shown in trace (9). The resulting coincidence between the peak selector output, the signal range gate, the sensitizing gate and the stretched pulse-train gate is indicated by the pulse

appearing in trace (10). This pulse is transmitted to the signal response counter. The stretched pulse-train gate is not applied to the gating circuit which controls the signals supplied to the noise response counter because, in practice, noise responses occur independently of whether signal pulse trains are present or not and must be counted accordingly.

The initiation of the counting process is controlled manually by a count start push-button switch, while its termination is controlled by the predetermined count relay which receives impulses from the signal response counter. The signal presentations counter indicates the total number of signal pulse trains generated within the counting interval. The signal detection probability is thus the ratio of the signal response count to the signal presentations count.

The period of time over which noise pulses are counted in any one run is indicated by the interval timer. The ratio of the noise response count to the elapsed time is the noise triggering rate.

The reset push-button switch, shown in Fig. 4.1, restores the electronic counters, the predetermined count relay and the interval timer to their "zero" condition. Resetting of the electromechanical counters is done manually.

4.2 Equipment Details. In this discussion, attention is confined primarily to the parts of the system which are not of

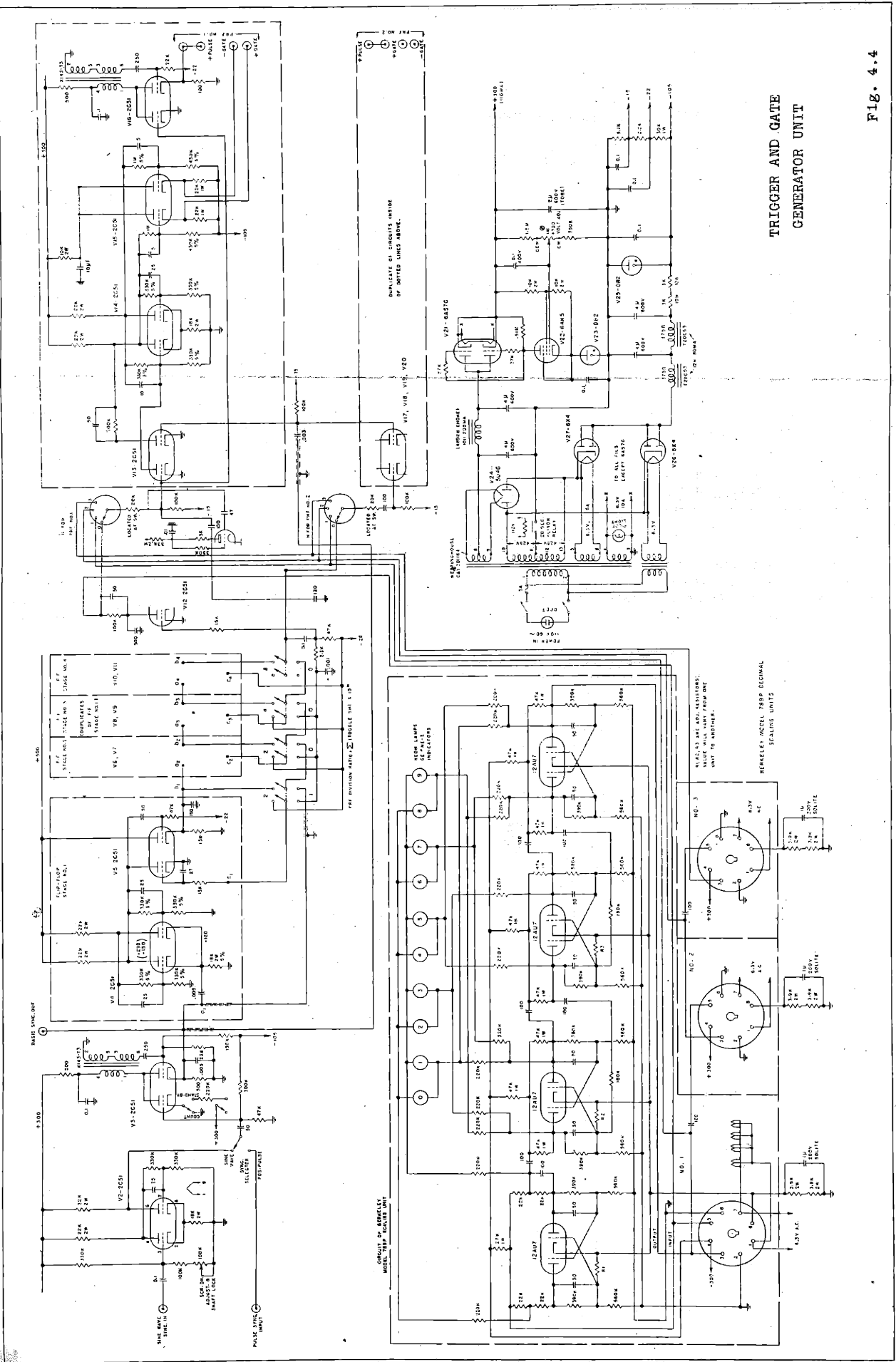
a conventional nature, although the more conventional portions will be described briefly for the sake of completeness.

4.2.1 Trigger and Gate Generator. A schematic circuit diagram of this unit appears in Fig. 4.4. It is primarily an adjustable-ratio pulse-scaling unit and supplies a low-repetition-rate trigger pulse (10 p.p.s. or less) to the pulse train control unit. The first portion of the scaling chain (V4-V11) is adjustable in scaling ratio from 1:1 to 16:1 in unit steps. Following the adjustable scaling circuit is a series of three scale-of-ten counters. By means of selector switches, any scaling ratio R satisfying the following relation can be obtained:

$$R = A 10^B, \quad A = 1, 2, 3 \dots 16, \quad B = 0, 1, 2, 3.$$

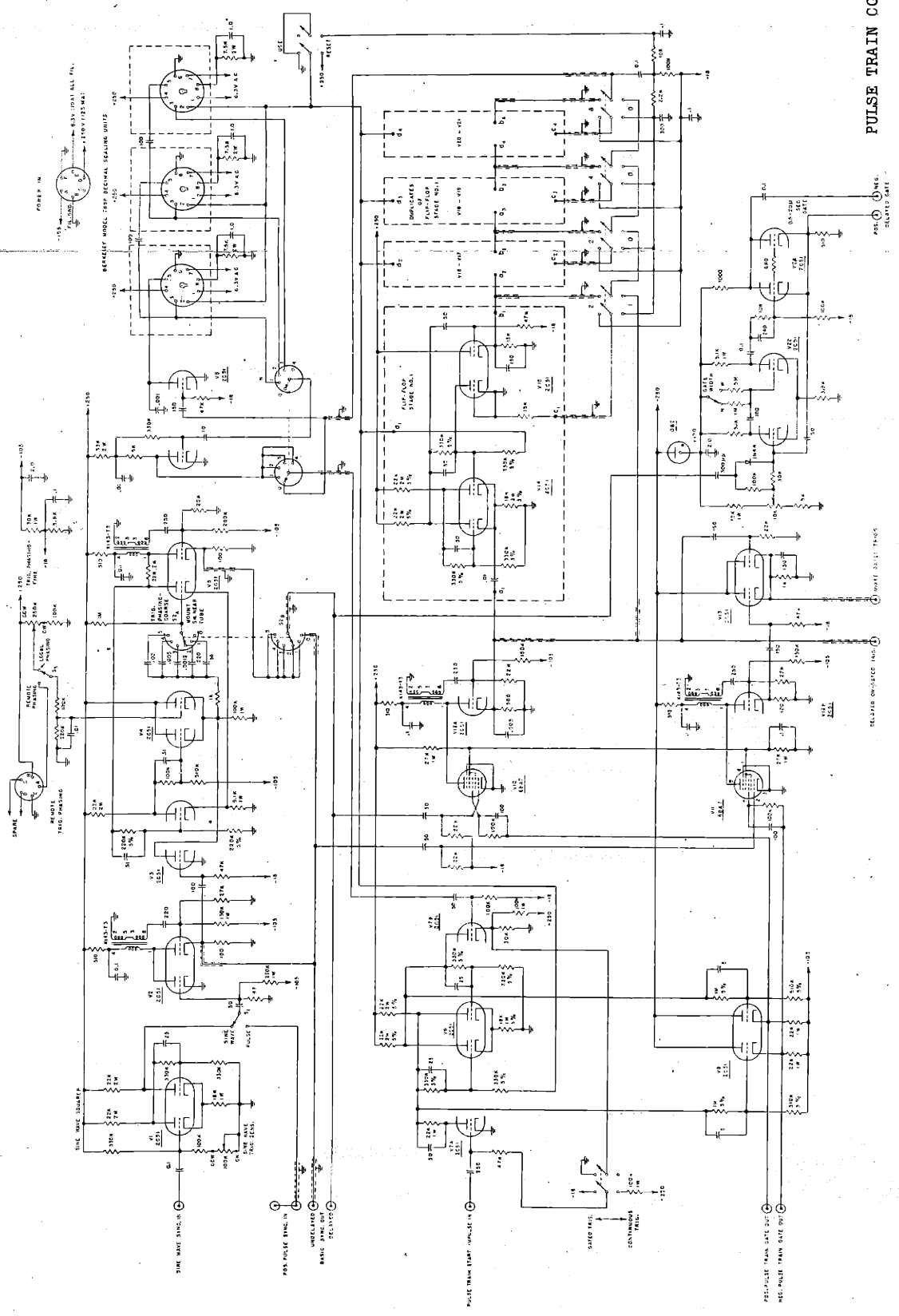
Sensitizing gate pulses beginning with each pulse train trigger pulse and ending with the next input pulse are generated with both positive and negative polarity (V13-V16). Trace (8) of Fig. 4.2 is representative of these gate pulses.

4.2.2 Pulse Train Control Unit. The schematic circuit diagram of this unit appears in Fig. 4.5. A trigger pulse delayed with respect to the master trigger pulse is generated by a mono-stable multivibrator and blocking oscillator circuit (V3-V5). This delayed trigger pulse is gated in V10 by the pulse train gate which is generated by V7-V9 and associated circuits. The pulse train gate is initiated by the pulse train trigger. This gate allows trigger pulses to be fed into a binary counter chain (V14-V21), which is followed by a



TRIGGER AND GATE GENERATOR UNIT

FIG. 4.4



PULSE TRAIN CONTROL UNIT

FIG. 4.5

series of three scale-of-ten scaling units. A pulse is derived from the scaling circuits after a predetermined count has been reached, this count being controlled by selector switches arranged as in the trigger and gate generator just described. This predetermined count setting is the pulse train number N. The counter output pulse terminates the pulse train gate, thus ending the pulse train. The delayed trigger pulse also triggers a gate generator (V22, V23) which produces the signal range gate shown as trace (7) in Fig. 4.2.

4.2.3 I.F. Pulse and C.W. Signal Generator. Fig. 4.6 is a schematic circuit diagram of this unit. A 15-mc oscillator stage (V1) feeds two frequency-doubling stages, V2 and V3. The 30-mc output signal from each is fed via separate level-setting attenuators to a common output circuit. Amplifier stage V3 is normally biased beyond cut-off and is brought into conduction by a rectangular pulse from the pulse generator circuit (V5-V8). The duration of this pulse is adjustable over the range 0.5 to 50 microseconds. Timing markers of 0.2-microsecond width and 1.0-microsecond spacing are generated by a delay-line controlled blocking oscillator circuit (V9). A sufficient amount of coupling exists between the marker generator and the multivibrator via the cathode-grid capacity of V5 to cause the multivibrator pulse to terminate at one of the marker pulses. Hence, as the pulse width control is varied, the pulse width varies in steps, which facilitates the adjustment of pulse width to a specified value.

Amplifier stage V2, when in operation, supplies a c.w. signal whose phase relative to that of the pulsed signal is adjusted by the tuning of the resonant circuit in the plate of V1. This c.w. signal is required in the procedure used to calibrate the signal power level in terms of the receiver noise power level, a variation of a method described by Lawson.⁽⁴⁾ As a first step, the receiver amplification is adjusted to give the desired receiver output noise level. The c.w. signal power level is then made equal to the receiver power level by observing the increase in the average rectified d-c output current of the receiver video detector. For the linear detector, this factor is known from theoretical considerations to be 1.45 when the c.w. power level is equal to the noise power level at the input to the detector.⁽⁴⁾ This condition is obtained by use of the c.w. level-setting attenuator, the calibrated signal attenuator being at a convenient setting. For the comparison of the pulse signal with the c.w. signal, the attenuation setting of the calibrated attenuator is reduced by 30 to 40 decibels, thus raising the signal level. The receiver amplification is reduced accordingly, and hence the comparison is made in the absence of any appreciable amount of residual noise. This comparison is based upon the fact that the resultant of signal and c.w. will be of the same amplitude as the c.w. alone if the pulse signal is of exactly twice the amplitude of the c.w. signal and 180 degrees out of phase; but, of course, the phase of

the resultant will differ from that of the c.w. by 180 degrees. Therefore, except for transient disturbances excited by the phase reversal at the beginning and end of the pulse, the receiver output signal will appear on an oscilloscope as a straight line. In order to assure that a steady-state condition is reached in the receiver circuits following the first phase reversal, it is necessary that the receiver bandwidth be at least three times the reciprocal of the pulse length. This requirement can be met by increasing either the receiver bandwidth or the pulse length at the time the comparison is made. Because of the possibility that the phasing adjustment will affect the amplitude of the c.w. signal, a preliminary adjustment of the phase is advisable. The adjustment of pulse-signal amplitude and c.w. phase to obtain the desired condition is very easily accomplished by observation of the oscilloscope pattern. It is apparent that when this condition is satisfied the initial setting of the calibrated attenuator will result in the pulse signal power being 6 decibels greater than the receiver noise power. Hence it is convenient to use an initial attenuator setting of +6 decibels relative to the desired zero decibel setting.

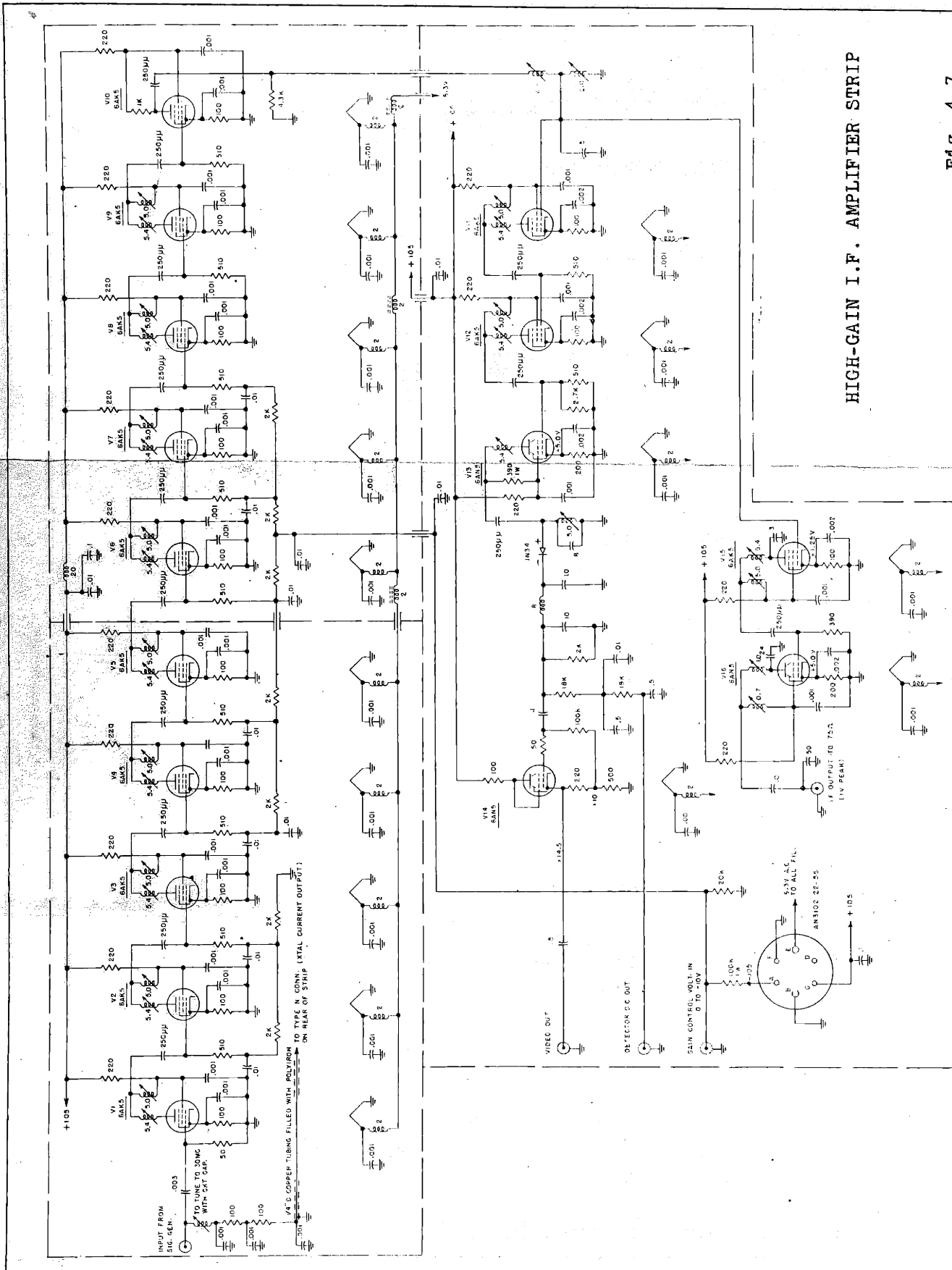
4.2.4 Multi-Bandwidth Receiver Unit. The receiver consists of two main sections: a high-gain, wide-band amplifier strip, followed by a multi-bandwidth unit which contains high-level i.f. amplifier stages, video detectors and an output

cathode follower. The schematic circuit diagram of the high-gain unit is shown in Fig. 4.7. The first ten stages of this unit (V1-V10) make use of identical double-tuned coupling networks having a single-stage bandwidth of 20 megacycles centered at 30 megacycles. Following the tenth stage, the i.f. amplifier splits into two channels, one of which feeds into a diode video detector and cathode follower for monitoring purposes. The other channel supplies i.f. signals to the multi-bandwidth unit via a terminated 75-ohm coaxial cable.

Fig. 4.8 is a schematic circuit diagram of the multi-bandwidth unit. The widest bandwidth channel is comprised of V1 and V2 and associated circuits. Again, double-tuned circuits are used. The overall receiver i.f. bandwidth for this channel is 10 megacycles (to the half-power points) and is largely determined by the coupling circuit between V2 and the video detector. The successively narrower channels are as follows:

3.0	megacycles bandwidth	-	V4, V5
1.0	"	"	- V6, V7
0.30	"	"	- V8-V11, V14
0.10	"	"	- V8, V11-V14

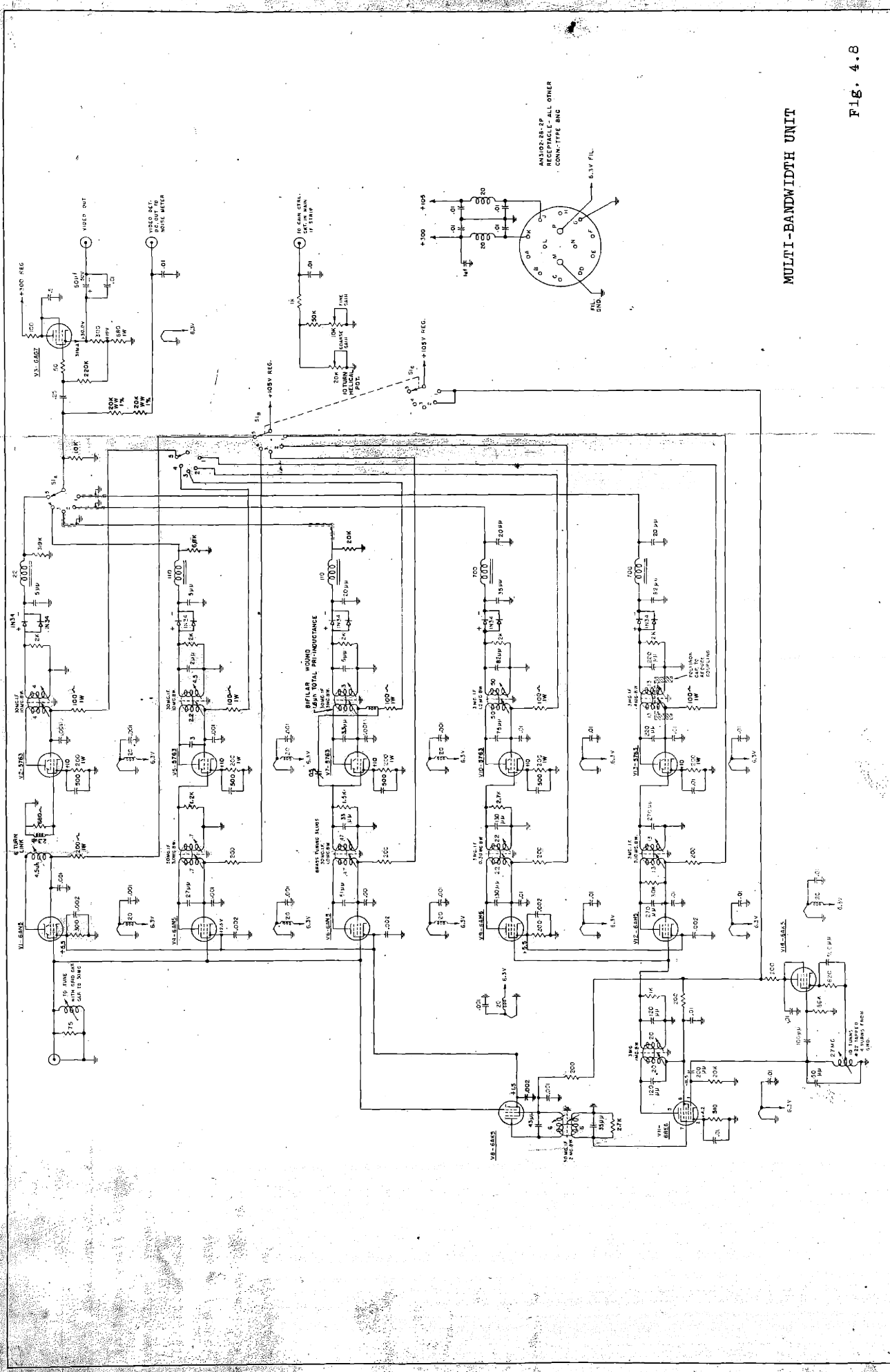
In all of these latter channels, the bandwidth is determined by the circuit immediately preceding the output i.f. amplifier stage. Initially, it was intended that the output amplifier - detector coupling circuit be the band-narrowing circuit because the higher transfer impedance that could



HIGH-GAIN I.F. AMPLIFIER STRIP

Fig. 4.7

PHOTO BY ELECTRONICALLY PRINTED FROM THE RECORD OF 1952

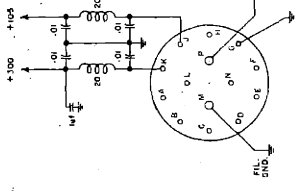


MULTI-BANDWIDTH UNIT

FIG. 4.8

AN502-28-2P
RECEPACLE - ALL OTHER
- CONN: TYPE BNC

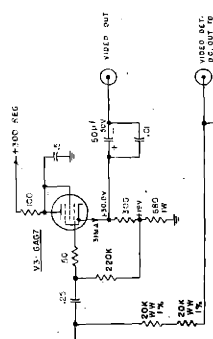
6.3V FIL



100V REC.

100V REC.

10 CAR CTRL.
DET. IN MAIN
AT 100V



thereby be obtained would produce higher peak output voltage. This, in turn, would minimize the effects of diode detector non-linearities present at low signal levels. However, the variation in detector impedance with signal level caused objectionable variations in the bandpass characteristic. Hence wide-band, low-impedance output coupling circuits were used to minimize this effect, and the band-narrowing circuit was moved ahead of the output stage as shown.

The two narrowest bandwidths are obtained with coupling circuits having reasonable Q's by using a center frequency of 3.0 megacycles. The 30-megacycle i.f. signal is heterodyned to 3.0 megacycles in the pentagrid mixer stage (V11) by a 27-megacycle local oscillator signal (V14).

The selectivity curve of the 0.1-megacycle bandwidth channel is shown in Fig. 4.9. This curve is typical of the other channels and because the 0.1-megacycle bandwidth channel was used in the majority of the experiments, only this one curve is shown. The noise bandwidth of this channel, as defined by Eq. 2.3, was found by graphical integration of the curve of Fig. 4.9 to be 0.105 megacycles.

The over-all amplitude response characteristic of each channel of the multi-bandwidth unit is shown in Fig. 4.10. The inset at the right shows the response at the lower signal levels and indicates that the linearity of the two narrow-bandwidth channels is slightly better than for the other channels. At higher levels, these channels show markedly superior

KRUFFEL & ESSER CO., N. Y. NO. 289-80
1 1/2" X 3 1/2" to the Inch.
MADE IN U. S. A.

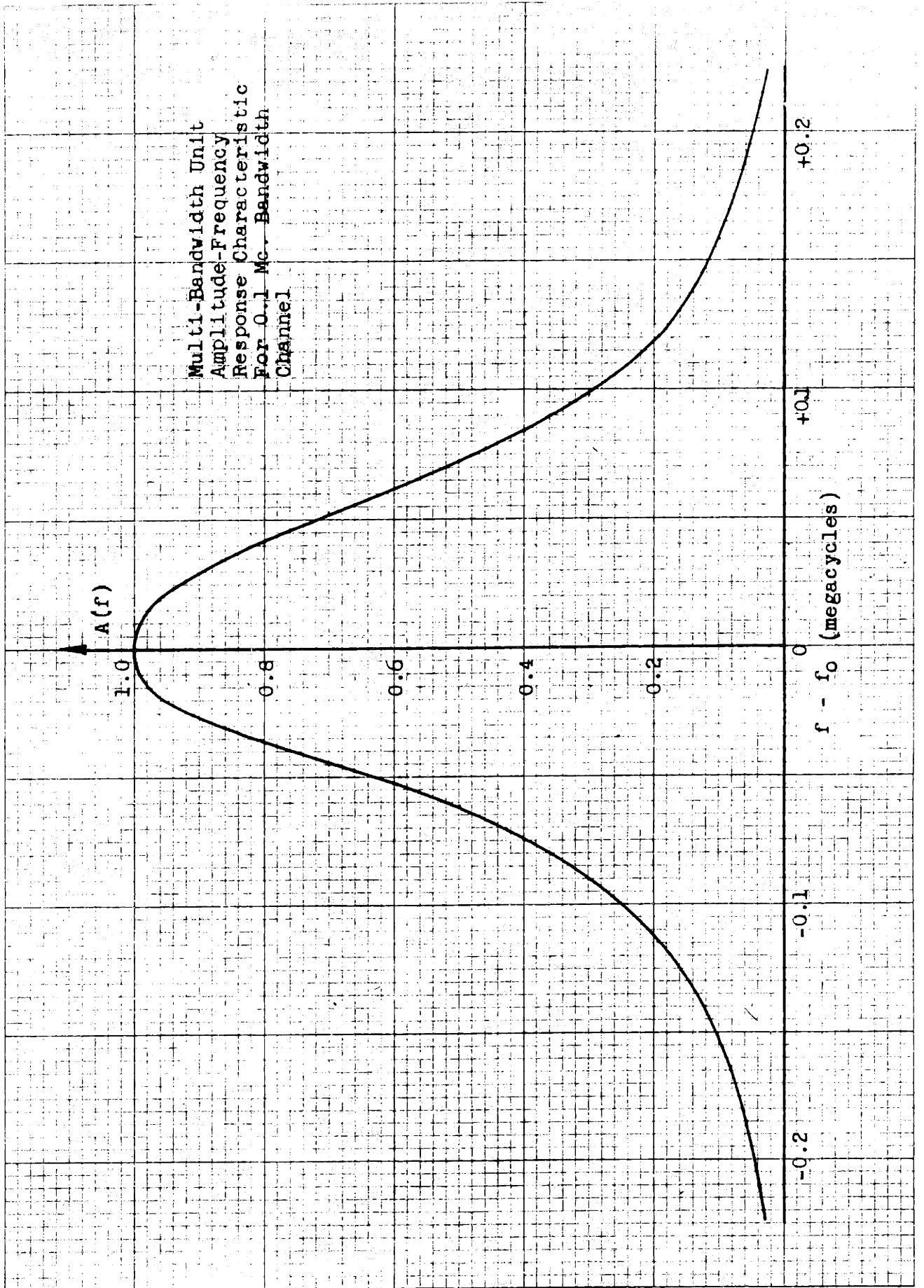


Fig. 4.9

KEUFFEL & ESSER CO., N. Y., NO. 359-14
Millbrook, New York
MADE IN U. S. A.

**Multi-Bandwidth Receiver
Overall Amplitude Response
Characteristic for Pulsed
I.F. Signals**

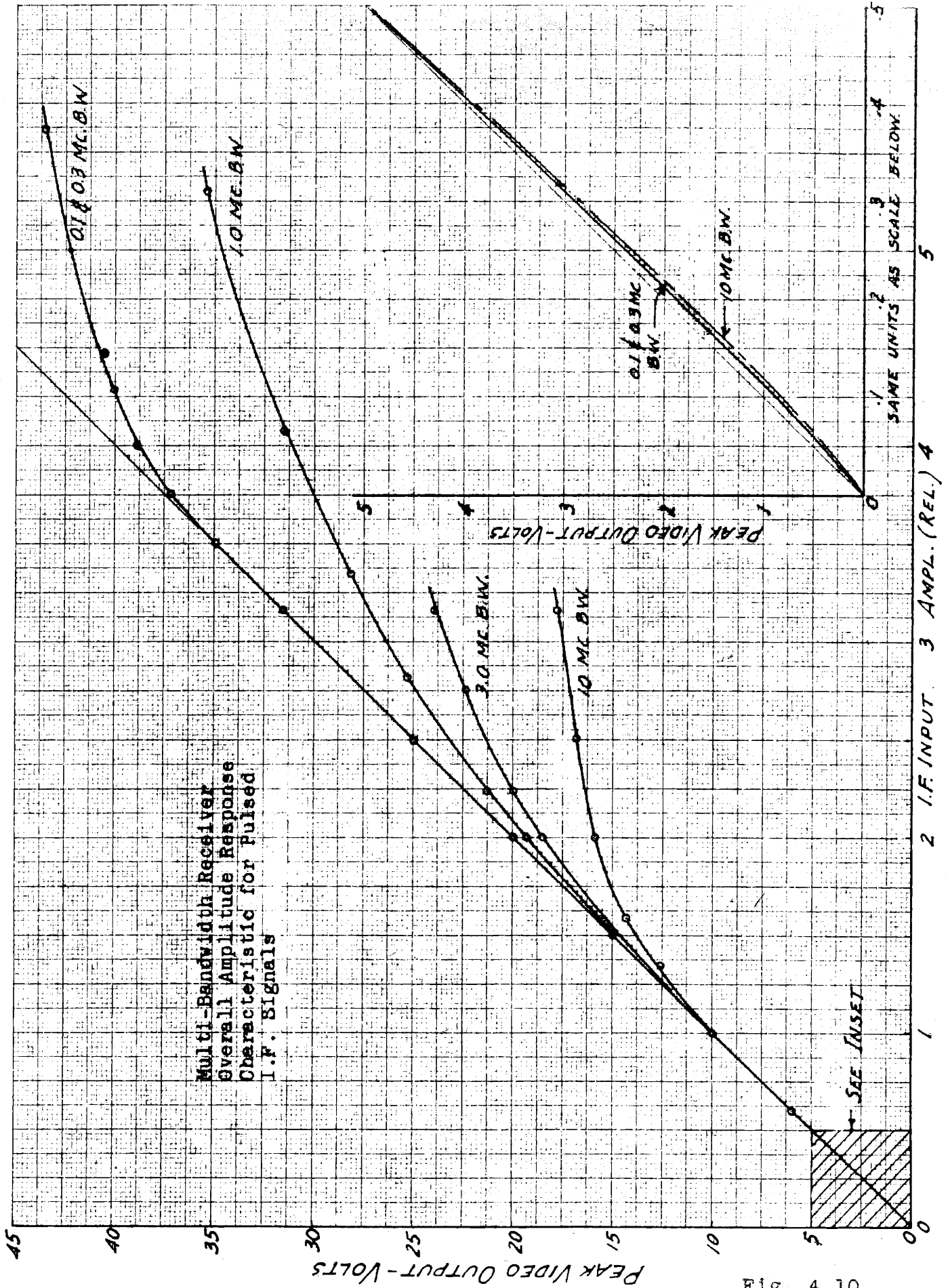
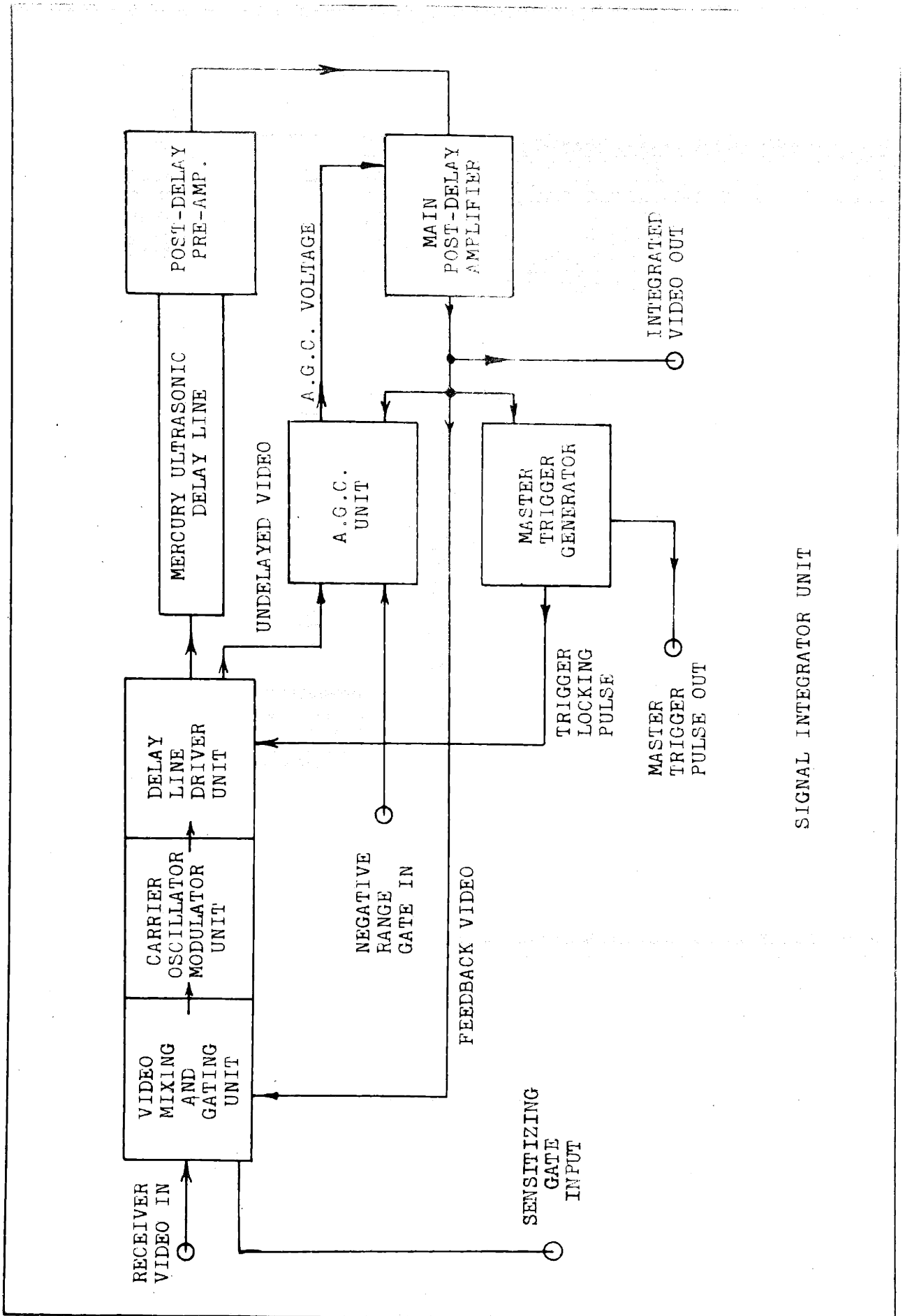


Fig. 4.10

characteristics. This superiority is a result of the higher value of diode load resistance used in these channels than for the wider-band channels. In the experiments, the average level of the receiver output noise was 5 volts. In this region, the amplitude response characteristic is seen to be very linear. Hence, the theoretical relation between detector d-c output and c.w. signal level used in the calibration of the signal generator can be applied with confidence.

4.2.5 Signal Integrator Unit. A block diagram of the equipment used in the study of the Type-I and Type-II signal integration systems is shown in Fig. 4.11. With but one exception, the experimental system used in both cases was identical. The system as used for the Type-I integrator will be described first. This system is not identical to the basic Type-I integrator described in Chapter III and shown in Fig. 3.1. The prohibitive number of signal delay units that would be required in such a system has already been noted. For the purposes of this investigation, equivalent performance is obtained by the use of one delay unit and a signal recirculation technique similar to that described in Chapter III in the discussion of the Type-II integrator and shown in the block diagram of Fig. 3.8.* It is recalled from the

*The technique used was suggested by Dr. R. M. Ashby, who was the author's supervisor while employed by the Naval Research Laboratory Field Station, Boston, Massachusetts.



SIGNAL INTEGRATOR UNIT

discussion of the Type-I integrator in Chapter III that its output at any instant is the sum of N input signals all added with equal weights. To accomplish this with a regenerative system, the loop gain must be unity. In order to satisfy this loop-gain requirement and at the same time preserve stability, the feedback path is opened periodically for one complete repetition period. Thus, the integration process becomes cyclic in nature with the signal and noise building up in the integrator loop from zero during the time the feedback path is closed. The number of basic repetition periods in the integration cycle is denoted by N_1 , called the integration number. At the end of an integration cycle, the total signal is therefore the sum of N_1 input signals, just as in the case of the basic Type-I integrator. It is readily appreciated, however, that a pulse train of N_1 pulses will be fully integrated only if it coincides exactly with the integration cycle. Subject to this condition, the gated integrator is equivalent to the basic Type-I integrator. Although this condition could not be met in practical systems because of the uncertainty as to the time of occurrence of any particular signal pulse train, it is easily satisfied in the laboratory because the signal pulse trains can be timed as required.

In the experimental system, the signal delay is obtained by use of a mercury ultrasonic delay line. The techniques for the construction and use of this type of line to obtain a signal delay of the order of a millisecond and an overall

bandwidth of the order of a few megacycles are amply described in the literature. (23-26) The mercury ultrasonic delay line used in this experimental system can be seen in Fig. 4.3 mounted on the side of the left-hand relay rack. The signal delay obtained with this line is approximately 500 microseconds. The signal to be transmitted is modulated onto a radio-frequency carrier (in this case 23 megacycles), amplified by a linear amplifier and delivered to the input terminals of the delay line as indicated in Fig. 4.11. Because of conversion loss at the transmitting and receiving crystals of the delay line and attenuation in the delay medium itself, the level of the output signal is approximately 60 decibels below that of the input signal. This loss is overcome by the post-delay amplifiers shown in the block diagram. The amplified signal is rectified to recover the video modulation. This delayed signal is fed back to the video mixing and gating unit as shown, the level of the feedback signal being adjusted to give unity loop gain. The gating of the feedback signal is accomplished before it is mixed with the input signal. The video gating circuit is controlled by the sensitizing gate pulse as represented by trace (9) of Fig. 4.2.

The master trigger pulses are generated in the following manner: A blocking oscillator pulse originating in the Master Trigger Generator unit is applied to the control grid circuit of the final amplifier stage of the delay line driver unit and cuts it off. The synchronizing pulse thus appears

at the input to the delay line as a 100 percent downward modulation of the carrier. The corresponding video pulse appearing at the output of the main post-delay amplifier is fed to the blocking oscillator in the master trigger generator to initiate the next trigger pulse. A clipping circuit removes the signal modulation from the synchronizing waveform fed to the blocking oscillator. This blocking oscillator pulse constitutes the master trigger pulse.

In order to obtain exact superpositioning of the pulse signals in the integration process, the spacing between the master trigger pulses, as determined by the total delay around the trigger loop, must be equal to the delay around the signal loop. It is seen from Fig. 4.11 that these two loops have certain parts in common. Differences in delay can arise only in the parts of the two loops that are not common. The principal source of additional delay in the trigger generator loop is that inherent in the blocking oscillator in the Master Trigger Generator. This delay is compensated by that inherent in the part of the video loop between the feedback input to the video mixer and gating unit and the point in the line driver unit where the synchronizing trigger is applied. A precise adjustment of trigger timing is obtained by varying the bias of the blocking oscillator circuit.

Two forms of Automatic Gain Control (A.G.C.) action are used. The first is for the purpose of compensating for gross variations in delay-line attenuation and post-delay amplifier

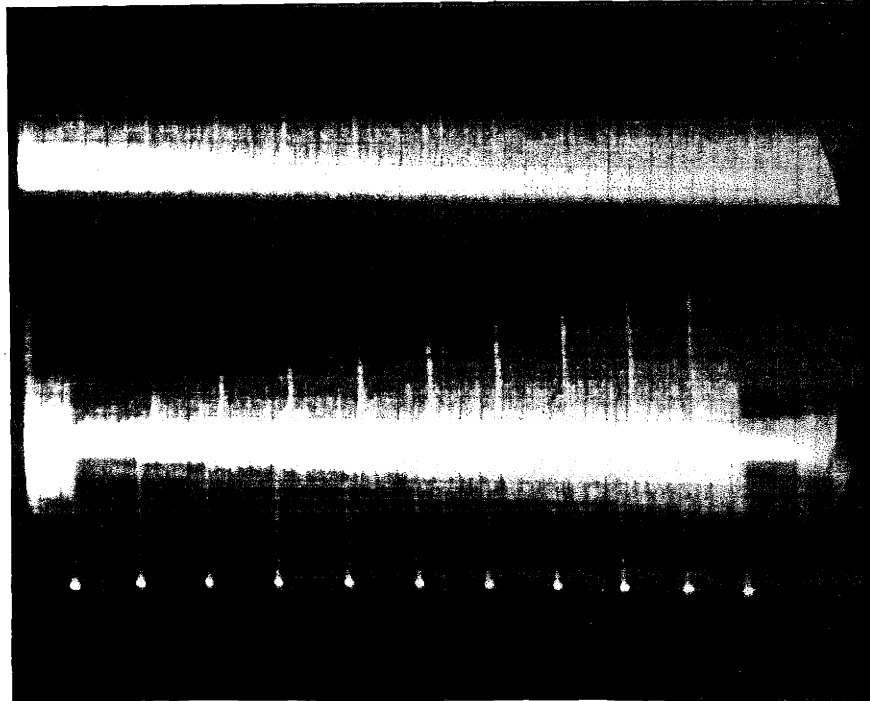
gain. In this portion of the A.G.C. circuit the carrier amplitude at the output of the main post-delay amplifier is compared with that at the input to the delay line. The error signal derived from this comparison is amplified in a two-stage direct-coupled amplifier and applied as an A.G.C. voltage to the first stage of the main post-delay amplifier. This portion of the A.G.C. system is called carrier A.G.C.

Although the major sources of variation of loop gain are compensated by the carrier A.G.C. system, a portion of the feedback loop is not thus compensated and was found to be a source of objectionable long-time variations in loop gain. Hence a supplementary A.G.C. system sensitive to loop gain variations was added. This system makes use of the fact that the overall amplification factor for the integrator is increasingly sensitive to loop gain variations as the loop gain approaches unity. Hence, with a constant-level input noise signal, the output noise level is also sensitive to loop gain variations and can be measured in a video rectifier to provide an A.G.C. voltage. In order to prevent the A.G.C. rectifier from responding to the intermittent signal pulse trains, it is desensitized by the signal range gate, as indicated in Fig. 4.11. This stabilization method is called noise A.G.C. It is apparent that this method of stabilizing the loop gain is satisfactory only for a laboratory experiment in which the input noise level is stabilized and the range position of the

signals in the pulse train is known. As these conditions were satisfied, this noise A.G.C. system proved to be both simple and effective.

Another matter of considerable importance in any practical system but of little consequence in this experimental investigation is the effect upon pulse signals of limited video bandwidth in the integration loop. If the pulses are of insufficient duration, they become seriously attenuated after repeated recirculation through the loop. For this reason, the bandwidth of the signals as supplied to the integrator was kept as small as possible by using the narrowest receiver bandwidth available (100 kc), thus permitting the use of 10-microsecond duration pulses. Under these conditions, the transient distortion of the signals caused by the integrator is negligible.

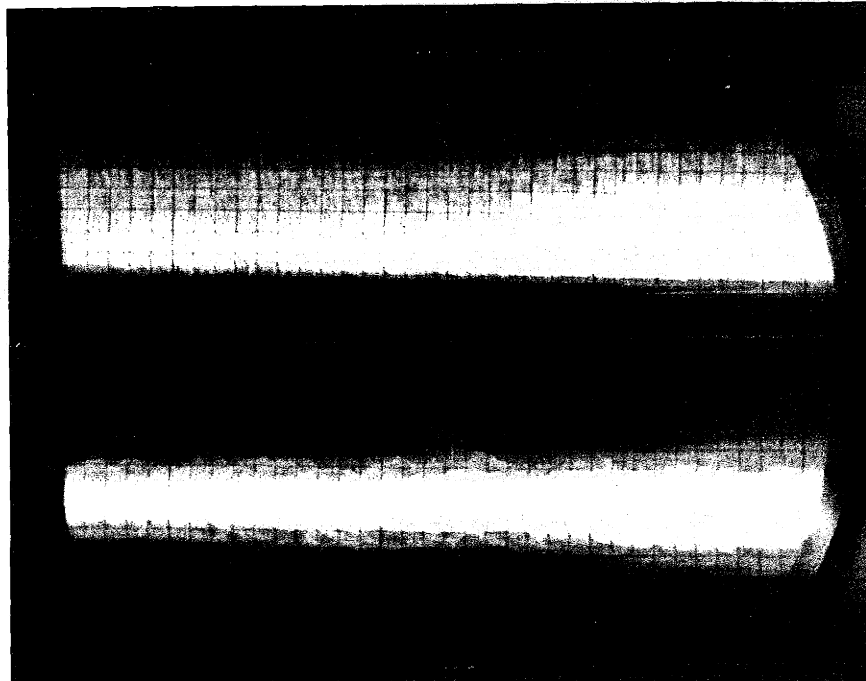
The response of this modified Type-I integrator to pulse signals accompanied by random noise is illustrated by Fig. 4.12, where one complete integrator cycle is shown by the lower trace. The bright-tipped downward deflections are the basic synchronizing pulses. The duration of the integration cycle is ten basic repetition periods. The build-up of signal pulses relative to the noise during the integration cycle is apparent. The input signal to the integrator is shown in the upper trace where the signal pulses can barely be seen.



RESPONSE OF MODIFIED TYPE-I INTEGRATOR
TO PULSE SIGNALS IN NOISE

For the investigation of Type-II integration, the only change required in the system just described is the elimination of the feedback gating action and the adjustment of the loop gain to the desired value. The elimination of the gating action requires only the removal of the sensitizing gate. The basis for choosing the loop gain setting and the method of achieving the desired value is described in the next chapter.

The response of the Type-II integrator to pulse signals accompanied by random noise is illustrated by the lower trace in Fig. 4.13. In this case, the pulse train consists of 100 pulses at a level of 6 decibels above noise, and the integrator loop gain is 0.99. Because of limited resolution of the oscilloscope from which the photograph was made, the individual pulses of the pulse train are not resolved. Hence only the envelope of the integrated pulse train is visible. The randomizing effect of noise upon the signal amplitude accounts for the jagged build-up curve, although the exponential character of the envelope is still apparent. The upper trace again shows the input signal to the integrator.



RESPONSE OF TYPE-II INTEGRATOR
TO PULSE SIGNALS IN NOISE

CHAPTER V

EXPERIMENTAL PROCEDURE AND RESULTS

5.1 Detection Without Integration. The first experiments performed in this investigation were for the determination of detection probabilities as a function of signal strength. In these experiments no form of signal integration was used. The parameters controlled were as follows:

- (1) Receiver bandwidth (Δf)
- (2) Signal pulse duration (τ)
- (3) Pulse train number (N)
- (4) Signal strength (λ)
- (5) Signal detection probability (P_D)
- (6) Noise probability (P_N)

The definitions of these parameters and the methods by which they are controlled are discussed in earlier chapters. Therefore only the choice of values used need be discussed here. For the receiver, the 300-kc bandwidth channel was chosen in preference to the wider-band channels because its amplitude response characteristic is linear over a considerably greater range, and in preference to the 100-kc bandwidth channel because tuning of the signal generator carrier frequency is less critical with the greater bandwidth. The pulse duration was made 3.0 microseconds in order to satisfy approximately the well-known optimizing relation $\tau\Delta f = 1$.^(4,6) For this relation to be satisfied exactly, the pulse duration

should have been 3.3 microseconds. However, the pulse-signal duration is adjustable only in multiples of one microsecond with the pulse-duration stabilization circuit in operation. The references just cited show that the threshold signal level is a very slowly varying function of the pulse duration in the vicinity of its optimum value. Hence this slight departure from optimum has a negligible effect upon the experimental results.

The pulse train number, N , was one of the parameters of the investigation and was varied over the range from one to one thousand. A signal response was counted if at least one of these pulses exceeded the slicing level. The sensitizing gate described in the previous chapter was on full time because there was no integration cycle involved in these experiments.

A second parameter in these experiments was the slicing level which, of course, determines the noise response rate and hence the noise probability. No attempt was made to preset the slicing level to establish specified noise response rates, but rather, convenient values of the slicing level were chosen and the corresponding noise response rate was recorded. However, the range of values of slicing level chosen was great enough to cover a range of noise probability values from 0.5×10^{-2} to 1×10^{-8} .

Before starting on experimental runs, the pulse train number, the slicing level and the signal strength were set to the

desired values. Then, with all counters re-set to zero, the run was started and allowed to continue until a predetermined number of signal response counts was obtained, at which point the run was automatically terminated in the manner described in the previous chapter. This predetermined count was generally in the neighborhood of 1000 to 2000. It was found that this number was sufficiently large to insure accuracy equivalent to approximately ± 0.1 decibel variation in signal strength. This accuracy is of the same order as the re-setting accuracy of the signal attenuator, and probably somewhat better than the reproduceability of the signal-to-noise-ratio calibration.

In any series of runs, the signal strength was varied over a range of values such that the detection probability varied from something in excess of 99 percent to below 5 percent. Fig. 5.1 shows the results of a number of such runs for several values of the slicing level, X_L . In this case a pulse train number $N = 1$ was used.

The slicing level, X_L , is expressed in units of the average noise amplitude as measured from its most negative excursions. The calibration of the video slicer in these units was accomplished as follows: With zero signal input to the slicer, the slicing-level control (a ten-turn Helipot with a 1000-division dial) was set to mid-scale (dial setting = 500), following which the slicer zero-axis adjustment was set to make the zero axis coincide with this control setting. The

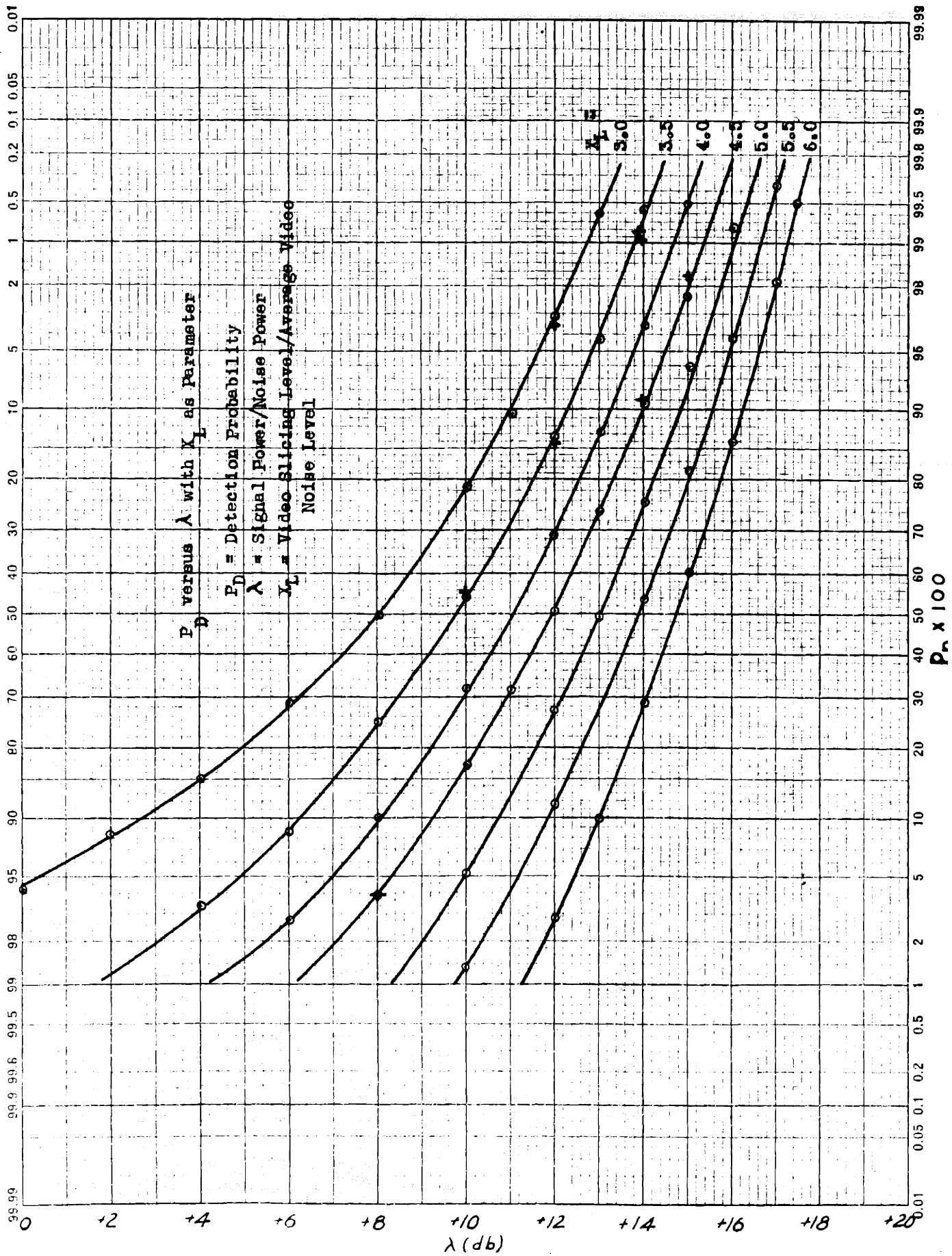


Fig. 5.1

control setting was next changed to a lower slicing level (for example to a dial setting of 400) and the input signal was restored. The amplitude of the input signal was then increased until the downward noise troughs just reached the slicing level as indicated by the noise response counter. Since the input signal is capacitively coupled to the slicer, its a-c axis coincides with the zero axis of the slicer. Hence the dial setting of 400 corresponds to the most negative excursion of the noise, while a setting of 500 corresponds to the average value measured relative to these negative peaks. Because the relationship between slicing level and Helipot setting is linear, any desired slicing level can readily be established.

During the signal detection runs, data on noise response rates were also accumulated. These measured noise-response rates, after conversion to equivalent noise probabilities by use of Eq. 2.7, are plotted as a function of the slicing level in Fig. 5.2. Using slicing levels read from Fig. 5.2 for several values of P_N , the curves of Fig. 5.1 were interpolated to obtain the points plotted in Fig. 2.13.

The results of similar experiments in which $N > 1$ are plotted in Fig. 2.14 as circled points. The method by which the data for the case of $N = 1$ can be used to calculate corresponding points was discussed in Chapter II, the results of these calculations being shown by the crosses in Fig. 2.14. It is seen that the circles and crosses coincide almost exactly, which indicates a high degree of equipment stability and calibration reproduceability.



Fig. 5.2

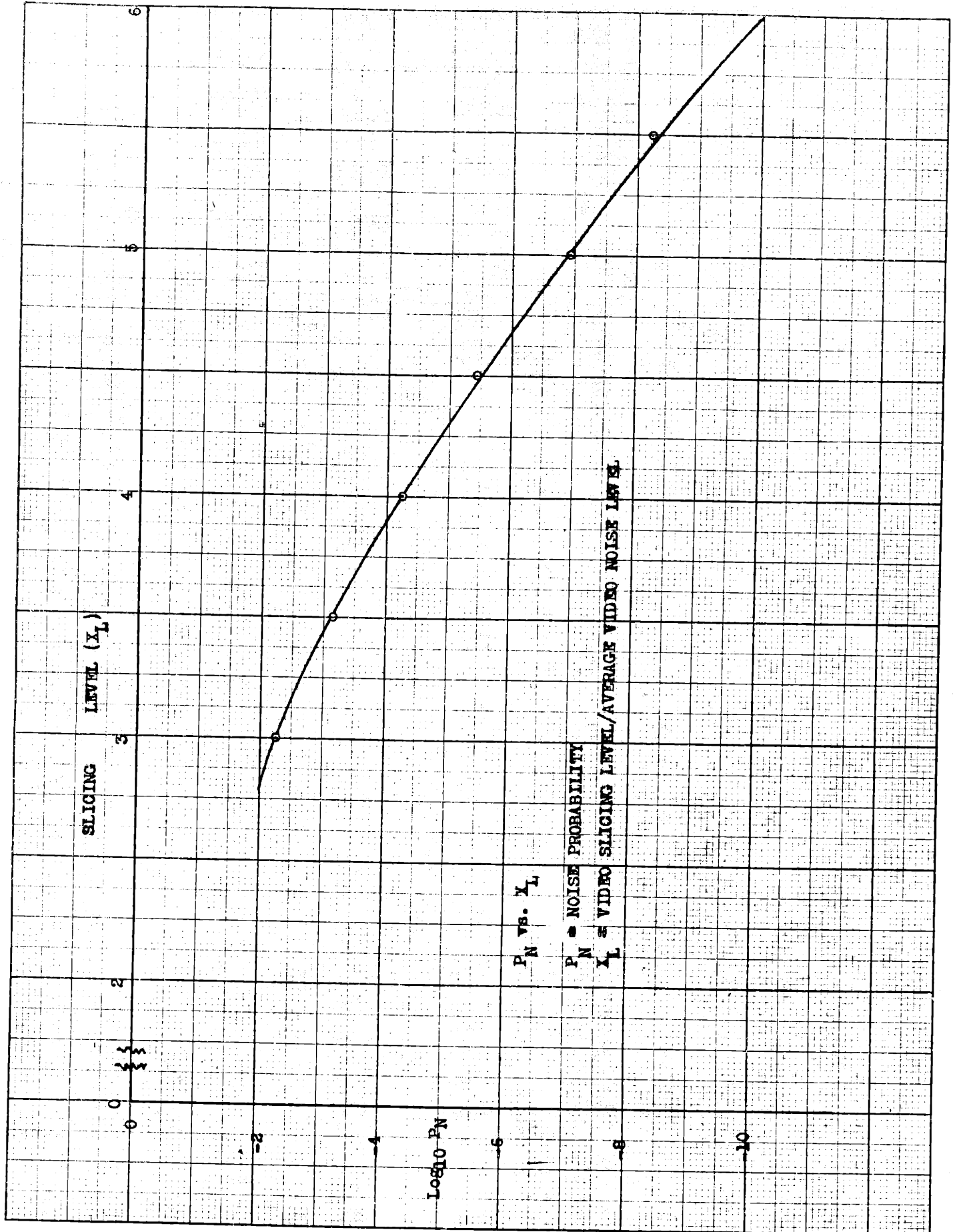


Fig. 5.2

5.2 Detection With Type-I Integrator. The properties of this type of integrator were investigated experimentally by making use of the simpler gated-feedback integrator described in Chapter IV. It was shown there that the two types are equivalent as regards the build-up of signals and noise provided that each signal pulse train occupies exactly one integration cycle. This condition was achieved by making the integration number, N_1 , as determined by the division ratio setting for the sensitizing gate generator, equal to the pulse train number, N , the pulse train generator being triggered by the pulse train trigger pulse, which pulse coincides with the leading edge of the sensitizing gate.

Although the only new parameter studied in this phase of the investigation was the integration number, N_1 , two additional parameters of the integrator required adjustment and checking. The first of these was the video gate balance adjustment. At the beginning of a run the gate was always balanced but was found to drift appreciably from balance if not occasionally readjusted. The effect of an unbalance of the gate is to introduce into the integration loop a spurious signal in the form of a square pulse of duration equal to the basic interpulse period. Because the low-frequency response of the integrator does not extend to zero frequency, this spurious signal excites a long-period transient which persists throughout the integrating period. The only objectionable effect of this transient is to shift the "zero" axis of the

waveform at the end of the integration period relative to the true a-c axis. This has the effect of shifting the slicing level and hence changes the false alarm rate. Fortunately, however, in comparison with the false alarm rate, the detection probability for a given signal strength is quite insensitive to the slicing level. For example, a change in slicing level sufficient to cause a two-to-one change in false alarm rate is compensated by a change in signal level of less than 0.2 decibel, as far as effect on signal detection probability is concerned. It was found possible to remain well within these limits by adjusting the gate balance at the beginning of each run.

The second adjustment required was the integrator loop gain. It was desired to maintain the loop gain at a value close to unity. This adjustment is most easily made by observing the response of the integrator to a single pulse which occurs at the beginning of the integration cycle. When the gain is unity, the integrator output consists of a train of pulses lasting throughout the integration cycle, all of the pulses being of equal amplitude. Following this adjustment the noise automatic gain control circuit is brought into operation, and the equipment is then ready for use.

The results of this experimental study of the modified Type-I integrator are plotted as circled points in Figs. 3.5 and 3.6 to permit comparison with the theoretical results derived in Chapter III. The largest value of N for which experimental points are shown is 100. Meaningful results could

not be obtained for appreciably larger values of N because of equipment limitations. The effects of slight variations in the loop gain become very pronounced. Even more serious, minute deviations from linearity of the amplitude response characteristic of the amplifiers in the integrator loop cause serious deviations from linearity in the amplitude response characteristic of the integrator under closed-loop conditions.

5.3 Detection With Type II Integrator. The basic integration system used in these experiments is the delayed-feedback type described in Chapters III and IV. It differs from the gated-feedback integrator used in the preceding experiments in that the feedback path is not gated but is closed continuously. The circuit is stable, therefore, only if the loop gain is kept below unity. It does not follow, however, that the loop gain should be made as close to unity as is consistent with stability considerations. In fact, it is easily seen that an optimum loop gain setting exists for each specified pulse train number, N . In Chapter III the performance of the Type-II integrator was calculated under the assumption that the loop gain, K , was such as to make the integrator time constant, N_1 , as defined by Eq. 3.27, equal to the pulse train number, N . A comparison of those results with the results for the case of no integration shows that a substantial improvement in signal detectability is obtained with the Type-II integrator for that particular condition of

adjustment. The effect of making N_1 much greater than N can be readily anticipated. Thinking of N as a measure of the number of signal or noise samples that contribute the total integrator output signal at any particular range position, it is apparent that the number of noise samples that can accumulate is much greater than the number of signal samples. Hence, under these conditions the integrator will cause a greater increase in the level of the noise than of the signal, and the signal detectability will be degraded. Therefore, lying between the two extremes just considered, there exists an optimum integrator time-constant setting.

The object of the first phase of the experiments with the Type-II integrator was to determine the optimum value for the integrator time constant. Accordingly, the variation of the threshold signal level as a function of the integrator time constant was investigated. This was done for values of the pulse train number, N , of 10 and 30. The results are shown in Fig. 5.3 where the input signal-to-noise ratio required for a detection probability of 90 percent is plotted in decibels as a function of the ratio N_1/N . It is evident that the optimizing condition is $N_1/N = 1$. In order to obtain a specified value of N_1 , the integrator loop gain, K , must be adjusted in accordance with Eq. 3.27. The method used to obtain a particular value of K is based upon the fact that the amplification of continuously repetitive pulses in the integrator increases by a factor of $1/(1-K)$ when the gain in the

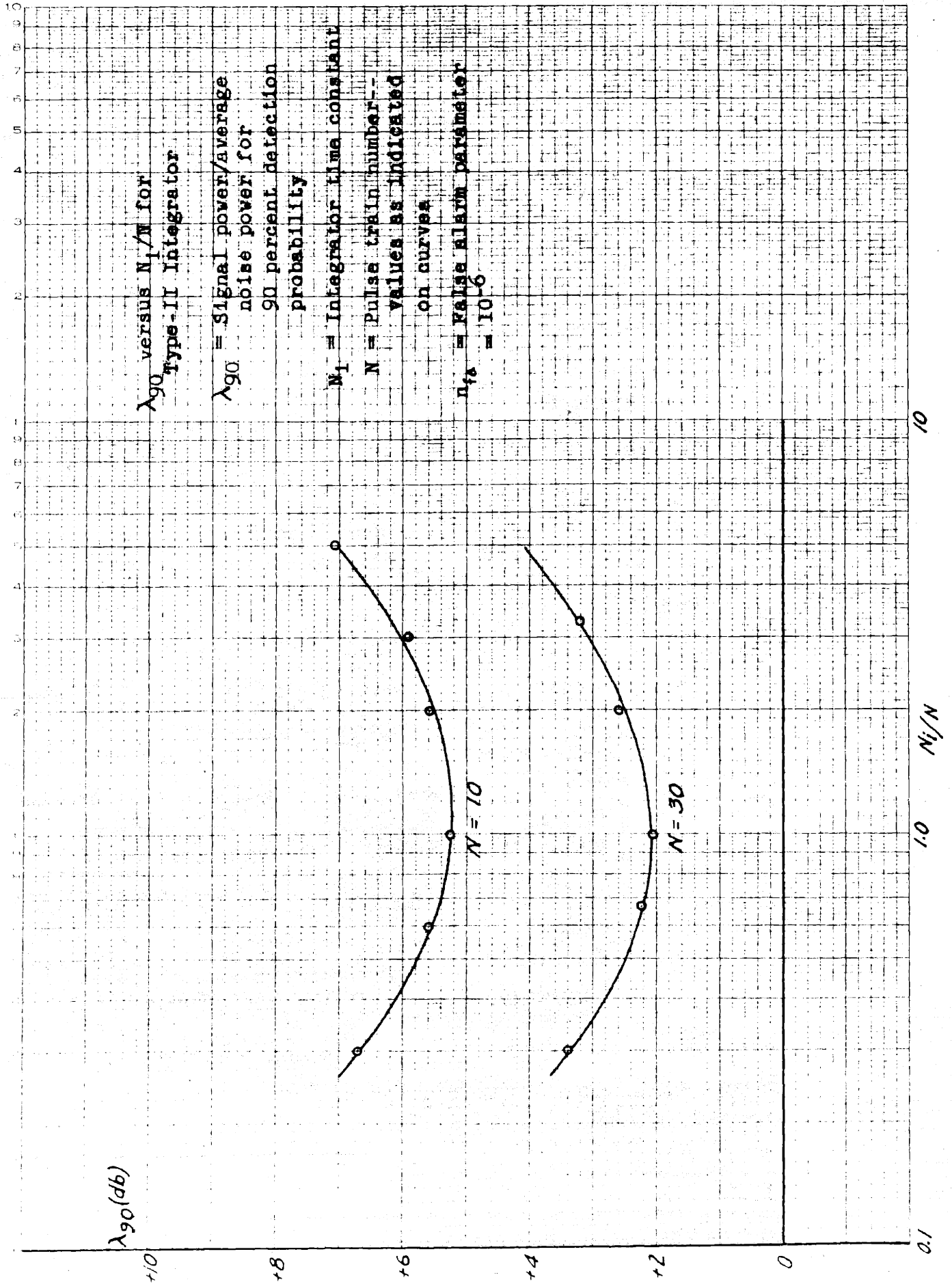


Fig. 5.3

feedback path is increased from zero to the point where the loop gain is K . The procedure used is as follows: A pulse signal is applied to the integrator input via a calibrated attenuator with the feedback gain at zero. The level of the input signal is first adjusted to produce an output signal of some fiducial amplitude. The input signal is then attenuated by a factor $(1-K)$ and the feedback gain is increased until the output amplitude is again at the fiducial value. This is the required loop gain adjustment. The receiver output signal is then substituted for the pulse signal and the noise A.G.C. circuit is brought into action, thus making the equipment ready for operation.

Using the optimizing relation, $N_1 = N$, a series of experiments was conducted. These experiments were closely analogous to those for the case of the gated-feedback integrator. Signal detection probabilities were determined as a function of signal level for a series of values of the false alarm parameter and with values of the pulse train number, N , of 3, 10, 30 and 100. The points appearing in Fig. 3.11 were obtained in these experiments, and again it is seen that the experimental data conforms closely to the theoretical curves.

As was shown theoretically in Chapter III, the difference in performance between the Type-I and Type-II integrators is almost insignificant, the former being superior by only about 0.5 decibel over most of the range of N . This slight difference is also apparent from the experimental results.

5.4 Significance of the Experimental Results. In the case of each of the basic signal detection systems studied experimentally, the results obtained were found to be in very good agreement with the corresponding theoretical results obtained in earlier chapters. These results indicate that under most conditions the use of the normal approximation to the probability density function for the integrated signal yields results which are in error by no more than one-quarter decibel, this figure being a conservative estimate of the accuracy of the experimental results with which the theoretical results are compared. Only in the case of N less than ten is the error appreciably larger, discrepancies of nearly one decibel having been noted for $N = 3$. For $N = 1$ the exact probability density function is available and yields results that agree with experimental results to within the limits of experimental error. The lack of experimental data for $N > 100$ in the integration experiments is not considered to be serious inasmuch as the asymptotic behavior of the curves of λ_p versus N is clearly established by the data plotted in Figs. 3.5, 3.6 and 3.11 for $N \leq 100$.

APPENDIX A

REFERENCES

- (1) L. N. Ridenour, "Radar System Engineering", Vol. 1, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1947).
- (2) D. O. North, "An Analysis of the Factors which Determine Signal-Noise Discrimination in Pulse Carrier Systems", Tech. Rep. No. PTR-6C, June 1943, RCA Laboratories Division, Radio Corp. of America.
- (3) A. V. Haeff, "Minimum Detectable Signal and Its Dependence upon Parameters of Radar Systems", Proc. IRE, 34 (Nov. 1946), 857-861.
- (4) J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals", Vol. 24, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1950).
- (5) R. Payne-Scott, "The Visibility of Small Echoes on Radar PPI Displays", Proc. IRE, 36 (Feb. 1948), 180-196.
- (6) J. H. Van Vleck and D. Middleton, "A Theoretical Comparison of the Visual, Aural and Meter Reception of Pulsed Signals in the Presence of Noise", Jour. of Applied Physics, 17 (Nov. 1946), 940-971.
- (7) J. V. Harrington and T. F. Rogers, "Signal-to-Noise Improvement Through Integration in a Storage Tube", Proc. IRE, 38 (1950), 1197-1203.
- (8) R. M. Ashby, F. W. Martin, and J. L. Lawson, "Modulation of Radar Signals from Airplanes", MIT Radiation Laboratory Report No. 914, Mar. 28, 1948.
- (9) T. T. Eaton and I. Wolff, "An Experimental Investigation of the Factors which Determine Signal-Noise Discrimination in Pulse Radar Systems--In Particular the Effect of Integration", Tech. Rep. No. PTR-7C, June 1943, RCA Laboratories Division, Radio Corp. of America.
- (10) T. T. Eaton, "Final Report on Noise Reduction by Delayed Feed-Back", April 15, 1943, a report of work done at RCA Laboratories Division, Radio Corp. of America, Princeton, N. J., under NDRC Project 595, Contract OEM-sr-252.

- (11) W. W. Hansen, "A Coincidence Method of Noise Reduction", MIT Radiation Laboratory Report No. 119(V-6S), 1941.
- (12) United States Patent No. 2,401,416, issued June 4, 1946.
- (13) A. S. Jensen, J. P. Smith, M. H. Mesner and L. E. Flory, "Barrier Grid Storage Tube and Its Operation", RCA Review, 9 (1948), 3.
- (14) S. O. Rice, "Mathematical Analysis of Random Noise", BSTJ, 23 (1944), 282-332, and 24 (1945), 46-156.
- (15) C. E. Shannon, "Communication in the Presence of Noise", Proc. IRE, 37 (1949), 10-21.
- (16) S. M. Kaplan and R. W. McFall, "The Statistical Properties of Noise Applied to Radar Range Performance". For summary, see Proc. IRE, 38 (1950), 195.
- (17) H. W. Leverenz, "Final Report on Research and Development Leading to New and Improved Radar Indicators", June 30, 1945, RCA Laboratories Division, Radio Corp. of America, a report of work done under NDRC Contract No. NDC-rc-150--See especially Section II-C (pp. 53-61) by D. O. North.
- (18) J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses", Phil. Trans. of the Royal Soc., Series A, 231 (1933), 289-337.
- (19) M. G. Kendall, "The Advanced Theory of Statistics, Vol. II" (Charles Griffin and Company, Ltd., London, 1948).
- (20) S. O. Rice, "Properties of a Sine Wave Plus Random Noise", B.S.T.J., 27 (1948), 109-157.
- (21) "Tables of Probability Functions, Vol. II", prepared by the National Bureau of Standards and available from the Supt. of Documents, U. S. Govt. Printing Office, Wash., D. C. (1948).
- (22) J. V. Uspensky, "Introduction to Mathematical Probability" (McGraw-Hill Book Company, Inc., New York, 1937).
- (23) H. B. Huntington, A. G. Emslie, and V. W. Hughes, "Ultrasonic Delay Lines I", J. Franklin Inst., 245 (Jan. 1948), 1-23.
- (24) A. G. Emslie, H. B. Huntington, H. Shapiro, and A. E. Benfield, "Ultrasonic Delay Lines II", J. Franklin Inst., 245 (Feb. 1948), 101-115.

- (25) B. Chance, V. Hughes, E. MacNichol, D. Sayre, F. C. Williams, "Waveforms", Vol. 19, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1949). See especially Chapter 23, pp. 751-765, "Supersonic Delay Devices".
- (26) B. Chance, R. I. Hulsizer, E. F. MacNichol, F. C. Williams, "Electronic Time Measurements", Vol. 20, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1949). See especially Chapter 12, pp. 471-525, "Delay and Cancellation of Recurrent Wave Trains".
- (27) G. N. Watson, "A Treatise on the Theory of Bessel Functions", Revised Edition (The MacMillan Company, New York, 1945).

APPENDIX B

BIOGRAPHICAL SKETCH OF AUTHOR

Harold V. Hance was born on January 9, 1918, at St. Louis, Mo. He received the degree of Bachelor of Science in Electrical Engineering from the California Institute of Technology in June, 1939. From 1939 to 1940, he studied at the Graduate School of Stanford University, specializing in electronics and radio engineering subjects. From 1941 to 1949 the author was employed by the Naval Research Laboratory, first in Washington, D. C., where he was engaged in the development of display systems for radar and electronic identification systems, and later at the Boston Field Station, Boston, Mass., where he was engaged in the development of techniques related to moving target indication systems, and where he became interested in the signal detection problem. Since 1949 he has been employed by the Naval Air Development Center, Johnsville, Pa., as an Electronics Engineer with the Radar Division of the Aeronautical Electronics and Electrical Laboratory. In June, 1946, the author entered the Massachusetts Institute of Technology as a part-time student to pursue graduate study in the Department of Electrical Engineering.