A STUDY OF THE SECOND-ORDER STATISTICAL PROPERTIES OF THUNDERSTORM NOISE

By

RICHARD LEON JOHNSON
II

Bachelor of Science
University of Texas at Arlington
Arlington, Texas
1963

Master of Science
Southern Methodist University
Dallas, Texas
1966

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1971

A STUDY OF THE SECOND-ORDER STATISTICAL PROPERTIES OF THUNDERSTORM NOISE

Thesis Approved:

Bennett Basore
Thesis Adviser

Paul A. McCollum

Lyle Broemeling

Wm. L. Hughes

D. D. Dunham
Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express appreciation to my thesis adviser, Professor B. L. Basore, for his continued encouragement throughout the course of this investigation and for his invaluable assistance in editing the manuscript of this study. I also want to thank the other members of my committee for their stimulating conversations and contributions at critical points in the investigation.

The financial support from the Department of the Army Electronics Command (Contract No. DAAB07-68-C-0083) is gratefully acknowledged.

Finally, I wish to thank my wife, Ethel, and my sons, Andy and Blake, for their patience and sacrifice during the course of my graduate studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Statement of the Problem. This investigation was undertaken to describe noise interference due to the radiated field of a thunderstorm in terms of its second-order statistics. To illustrate the need for a second-order statistical description, consider the problem of designing a smoothing filter to extract a wanted signal from a mixture of signal and noise. Let the corrupted signal $x(t)$ be denoted by the sum of a wanted signal $s(t)$ and unwanted noise $n(t)$,

$$x(t) = s(t) + n(t) \quad .$$

There are several possibilities concerning the nature of both inputs $s(t)$ and $n(t)$. For example, if $s(t)$ is known exactly, there is no problem; if $n(t)$ is known exactly, it is trivial to obtain $s(t) = x(t) - n(t)$. At the other extreme, if there is no a priori information at all about $s(t)$ and $n(t)$, there is no hope of extracting $s(t)$. Clearly, the interesting cases are those in which there exists some uncertainty about the inputs but not too much. For inputs described by this last case, then, $n(t)$ and $s(t)$ are assumed to be member functions of a stochastic process. Let the output signal of the filter be denoted by $\hat{s}(t)$. The optimality criteria for the design of the filter is to minimize

$$E\{[s(t) - \hat{s}(t)]^2\} \quad .$$

Let h(t) denote the impulse response of the filter and write the mean-squared-error as

$$E\{[s(t) - \int_{-\infty}^{\infty} x(t - \alpha)\ h(\alpha)\ d\alpha]^2\} \quad .$$

Using the orthogonality principal, it can be shown (17) that a minimum results if

$$R_{sx}(\tau) = \int_{-\infty}^{\infty} R_{xx}(\tau - \alpha)\ h(\alpha)\ d\alpha$$

or if

$$S_{sx}(f) = S_{xx}(f) \cdot H(j2\pi f) \quad .$$

Assuming n(t) and s(t) are uncorrelated, then[1]

$$H(j2\pi f) = \frac{S_{sx}(f)}{S_{xx}(f)} = \frac{S_{ss}(f)}{S_{ss}(f) + S_{nn}(f)} \quad .$$

A solution to the problem described above, then, does rely upon one's knowledge of the power spectral density of the noise, namely $S_{nn}(f)$. Since the power spectral density makes use only of second-order statistics, the filtering problem motivated a study of the second-order statistical properties of thunderstorm noise.

---

[1]It is pointed out that $H(j2\pi f)$ can be made physically realizable by adding constraints upon the solution. However, for the purposes of illustration, this seems unnecessary.

1.2 Existing Solutions. There have been many studies of the spectra of atmospherics and the results are surveyed by Horner (12) and Uman (24). These studies were conducted in two different ways: (1) the Fourier transform was computed for short segments of storm data, i.e. one return stroke, a stepped-leader, etc.; and (2) the electromagnetic signature was measured with a number of narrow band receivers and the spectrum was plotted directly as a function of time. The concensus of the investigators seems to be represented by the study of Arnold and Pierce (2). They found peak amplitudes of the frequency spectra for stepped-leader radiation, K-change radiation and return-stroke radiation at 20, 8 and 5 KHz, respectively, with corresponding relative peak amplitudes of 1:2:10. The spectral analysis reported above has been done using either deterministic or semitheoretical techniques. Arnold and Pierce develop the latter approach by considering an empirical spectrum resulting from (1) above, and then modify the spectrum using a statistical smoothing factor obtained theoretically.

There have been numerous investigations of the first-order probability density function of thunderstorm noise and its envelope. Horner (12) reports that the first-order probability density function of the amplitude of broadband storm noise is log-normally distributed. Horner (12) also reports that the density function of narrow band envelopes is either log-normal or Rayleigh distributed depending on center frequency and bandwidth. To the author's knowledge, statistical descriptions of the amplitude of thunderstorm noise have been confined to first-order probability density functions.

1.3 Present Contribution. The contribution offered in this study is a statistical description of the thunderstorm, as a source of noise interference, during periods of intense sferic activity. The data segments under study exhibit the type of sferic activity existing in the presence of an intense build-up preceding a discharge, or occurring while a discharge is in progress. The statistical characterization is formulated through the estimation of time averaged autocorrelation and power spectral density functions. It is believed that this description will be a useful contribution toward the design of filters to minimize thunderstorm interference in a mean-square-sense.

A brief outline of the study begins with the discussion in Chapter II concerning the calibration of the data gathering instrumentation and illuminating some of the limitations placed on the study. In Chapter III, the phenomenon of the lightning discharge is discussed, and an approximate model is developed to illustrate the effects of propagation and the frequency response of the measurement system. The essence of the study is contained in Chapters IV and V. In Chapter IV, three segments of 0.75 second duration are investigated to determine whether or not the assumption of wide sense stationarity is valid. Proceeding from the assumption of wide sense stationarity, the text of Chapter V investigates the credibility of estimating the time averaged autocorrelation and power spectral density functions by time averaging over each data segment. For the interested reader, a mathematical treatment of this technique of estimation is offered in Appendix A along with the computerized implementation given in Appendix C.

CHAPTER II

INSTRUMENTATION

2.1 Introduction. This chapter discusses the technique of data gathering and subsequent processing of the electromagnetic field signature. The sensing apparatus is mounted on board a D-18 twin engine Beechcraft airplane. A description is offered of the measurement electronics, the flight pattern used by the aircraft while gathering data and the technique of analog to digital conversion.

2.2 Data Gathering. The vertically polarized electromagnetic field is sensed through the use of a five foot vertical whip antenna mounted atop the fuselage of the airplane. The signal flow path is shown by the block diagram illustrated in Figure 1. The signal at the antenna terminals is first fed through a cathode follower which acts as a buffer between the antenna and the preamplifier. The output of the preamplifier is, then, fed through a second cathode follower stage which acts as a buffer between the preamp and the main amplifier. After the signal has been amplified by the main amplifier, it is recorded on magnetic tape. For the interested reader, a detailed circuit description of the measurement electronics is offered in Appendix D.

Once the measurement system had been assembled, the next logical consideration was the calibration of the system. Because this measurement apparatus was expected to record field changes in the LF(0-300 KHz) band, the problem of calibration was an extremely difficult one to

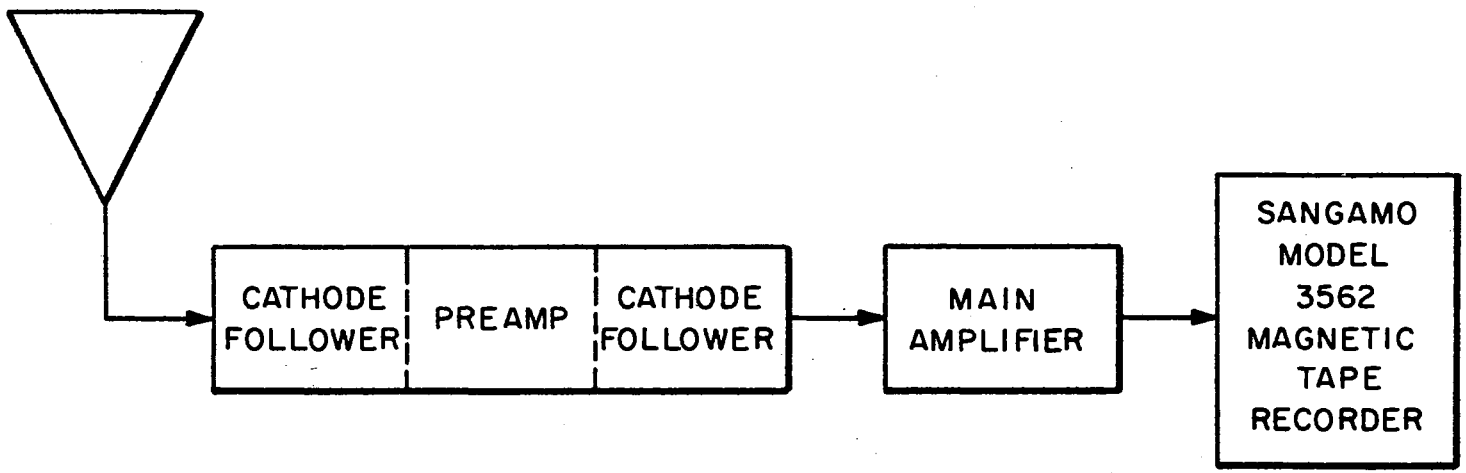


Figure 1. Block Diagram of the Data Gathering Instrumentation

solve. The essence of the problem was that of calibrating an antenna in this frequency range. The problem was further complicated by the fact that the calibration had to be done with the antenna afixed to the aircraft, thereby taking into account the perturbing influence of the airframe geometry. Ideally, one would want to measure, while in flight, the system response, as a function of frequency, to a known field generated by a previously calibrated antenna. Before developing an elaborate calibration capability, the question arose, could one settle for a relative measurement as opposed to an absolute measurement? It was concluded that for the purposes of this study a relative measurement would be more desirable than an absolute measurement. Specifically, since this study is concerned more with the structure of the second-order statistics as a function of the various data segments than it is with the absolute amplitude on any particular segment, the task of calibrating in an absolute sense presented an unnecessary complication. A calibration curve was obtained by applying a constant amplitude sinusoidal signal across the first cathode follower input terminals of Figure 1. The system response was measured as the frequency of the sinusoid was varied. Hence, there resulted a plot of the absolute value of the Fourier transform of the system impulse response as a function of frequency. This curve is shown in Figure 2. The frequency response plot describes the instrumentation response in a composite sense. Specifically, it takes into account the measurement system characteristics and the response of the replay electronics with a 32:1 time base expansion of the recorded storm signature. Because it was thought that this calibration technique may have altered, in a significant way, the capacity coupling between the antenna and the airframe, and presented an invalid frequency

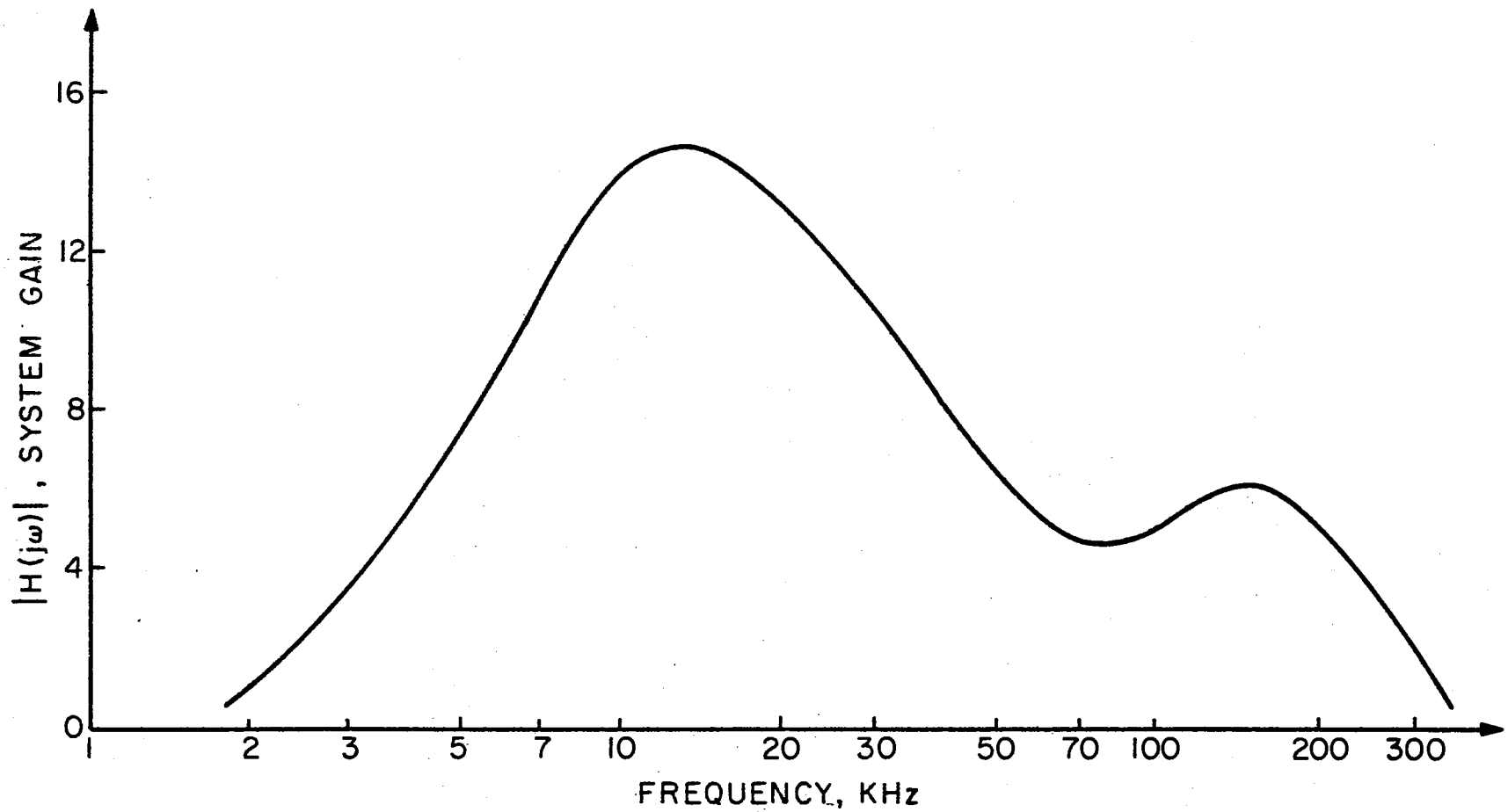


Figure 2. Frequency Response of the Measurement System

response curve, a second calibration technique was implemented. This technique involved no physical attachment to the input of the measurement system. It consisted of operating a spark generator in the vicinity of the aircraft. The spectrum of the radiated field of the spark was known a priori to be flat over the range 1-100 KHz. Because the current in successive sparks varied significantly, the calibration instruments displayed fairly erratic behavior; however, by smoothing, a curve almost identical to that of Figure 2 was obtained. The exception was a slightly more rapid fall-off in frequency above 10 KC. It is believed that the curve of Figure 2 is the more accurate of the two because of the inconsistency of the source in the spark generator method; however, because of the near agreement it is concluded that the antenna characteristic is flat over the spectrum of interest.

The data gathering is done by flying the airplane in an elliptical path in the horizontal plane at an altitude of five to six kilometers centered some thirty kilometers from the storm cell. Generally the flight path is executed by flying in one direction for ten minutes, making a 180^{o} turn and flying in the opposite direction for ten minutes. This then modifies the elliptical path proportional to wind velocity and the resulting drift of the storm cell itself, the objective here being one of maintaining the path as near constant as possible with respect to the storm cell.

2.3 Analog to Digital Conversion. Because the recording bandwidth capability increases proportional to tape speed, the data recording is done at a tape speed of 60 inches per second for a duration of one minute every five minutes. To convert the data to digital form for computerized reduction, a data processing assembly has been designed. The

processing assembly consists principally of a Sangamo Model 4744 magnetic tape replay unit, an EECO 761A analog-to-digital (A/D) converter, an HP-2115A computer and an HP-3030 9 track digital tape transport. The analog data tape is recovered by the replay unit and is fed into the A/D converter. The digitized data from the A/D converter is then transmitted to a memory buffer in the HP-2115A computer. Concurrently, another memory buffer in the computer is transmitting the previous block of digital data to the HP-3030 tape transport. The internal timing logic of the computer interchanges the functions performed by the two memory buffers so that the analog data can be digitized continuously. Because the highest allowable sampling rate is 25 KHz, without skipping analog data, the analog data tape is replayed at a speed of 1 7/8 inches per second. This results in a 32:1 time base expansion and, correspondingly, the Nyquist sampling rate is reduced by a factor of 32. Because the computational time required for data reduction is a function of the number of samples characterizing a particular segment of data, it was deemed necessary to sample at as slow a rate as possible without producing excessive aliasing errors. Using this criteria, a sampling rate of 20 KHz was chosen, implying an unaliased bandwidth capability of 320 KHz in the original data. In view of Figure 2 this sampling rate seemed reasonable. In addition, memory buffer blocks of 2000 sample points each were used, each block containing 1/32 millisecond of real time data. To offer some insight into the quantity of data being handled, typical records of 3/4 second require 480,000 sample points to characterize the segment.

The word format used by the HP-2115A is 16 bits, where bits 0-14 represent the number in integer binary form and bit 15 is the sign bit.

A program has been developed for use on the IBM 360/50 to convert the sample points into a 32 bit format where bits 0-23 represent the number in integer binary form, bits 24-30 are the base ten exponent in decimal and bit 31 is the sign bit. Hence, there results a real, floating point format. This final step then generates a third data tape from the one produced by the HP-3030. Once the third tape is generated, the data are ready for the reduction process described in Appendix A, using the program of Appendix C.

In summary, two areas of special significance discussed in this chapter are reemphasized. The first is the overall frequency response of the measurement system. Because the passband, or 3 db points, is from 6.5 to 30 KHz, it will be this portion of the thunderstorm noise spectrum which will receive greatest attention in the discussions of later chapters. It should be pointed out that this restriction is not as severe as it may seem at the onset. As reported in Chapter I, it is generally accepted as fact that the spectral content of a lightning discharge peaks between 5-10 KHz, and the electromagnetic field spectrum is reduced by a factor of approximately 10 at 20 KHz (2). Hence, in spite of the fact that there currently exists limited communication activity in this frequency band, the study concentrates on that portion of the spectrum where the amplitude of the measured thunderstorm signature is large with respect to the internally generated noise in the sensing instrumentation.

A second area of particular significance is the quantity of samples which must be used to characterize the data segment. For example, to replay 3/4 second of data requires 24 seconds of real time, and as mentioned previously 480,000 sample points are recorded on the HP-3030

tape reel during the replay. Following conversion of this tape to a tape with 32 bit words, the program described in Appendix C computes the second-order density function. The result is printed in thirty 9 x 9 matrices, one matrix for each value of τ. Additionally, the program computes the autocorrelation function for thirty different values of τ and uses this result to compute the power spectral density function in 64 increments of 1 KHz. The actual CPU time involved is, typically, 1 hour and 30 minutes. Hence, the data handling portion of this study presented a problem of tying up the IBM 360/50 for sustained computing periods.

CHAPTER III

LIGHTNING AND PROPAGATION MODEL

3.1 Introduction. Since lightning discharges produce the most dramatic features of the electromagnetic signature recorded during a thunderstorm, an account is given of the sequence of events occurring during a discharge. A model describing the propagation of the changing electric field is then derived, and incorporating empirical data, the noise power spectral density is predicted.

3.2 The Lightning Process. The probable distribution of thundercloud charges as reported by Uman (24) is shown in Figure 3. The solid black circles indicate locations of effective point charges, typically P = +40 coulombs, N = -40 coulombs and p = +10 coulombs. These charged regions are created in a manner not completely understood at this time by the turmoil of wind, water and ice in the presence of a gravitational field and a temperature gradient. One way in which these charges are neutralized is the cloud-to-ground lightning discharge, or flash, which consists of one or more strokes, each stroke being preceded by a leader streamer.[1] The leader process propagates from the cloud-to-ground and is followed immediately by the luminous return stroke propagating back up the ionized channel. There are generally three to four strokes per

[1]The sequence of events and parameters quoted in this text regarding the lightning flash are those reported by Uman (24).

flash. The streamer preceding the first return stroke is termed the stepped-leader. The stepped-leader process begins with a local electrical breakdown between the N and p regions of the thundercloud. The resulting strong concentration of negative charge at the cloud base produces an electric field which causes a negatively charged column to descend toward the earth. The streamer is called a stepped-leader because it descends in a typical step of 50 meters and then pauses for approximately 50 μsec before proceeding downward again. Its average velocity is about 1.5×10^5 m/sec and it carries approximately 100 amperes near ground.

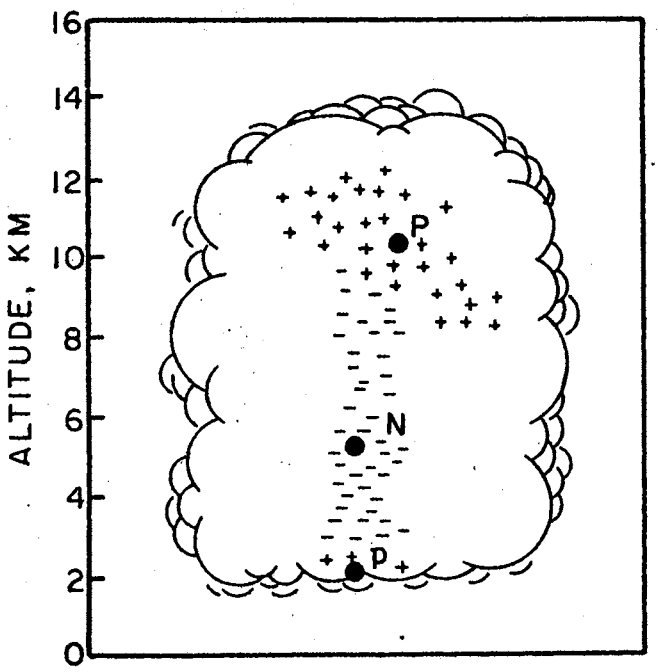


Figure 3. Probable Distribution of Thundercloud Charges

When the negative column is near the earth's surface, the resulting high electric field causes upward-moving discharges from the ground toward the leader tip. As one of these discharges contacts the streamer, the tip becomes grounded while the remainder of the leader has a distributed negative potential back up the leader channel. The return-stroke process carries the ground potential back up the leader channel, propagating up the ionized channel at a typical velocity of 5×10^7 km/sec and carrying currents in the order of 20 KA.

Upon the completion of the first return stroke, the top of the channel is left positively charged with a strong electric field existing between the top of the channel and the remaining negative charge in the cloud. Positive streamers, called J-streamers, advance upwards and outwards reaching pockets of negative charge and establish negative recoil streamers within the cloud. The field changes due to these negative recoil streamers are called K-changes. When the positive charge has been neutralized, a negative charge may collect at the top of the channel with sufficient density to start a second leader toward the ground. This second leader is called a dart leader and differs from the stepped-leader in that it propagates toward the earth in one continuous movement. A typical velocity for the dart leader is 2×10^6 m/sec. The second return stroke then propagates back up the ionized channel established by the dart leader. The dart leader-return stroke process continues until the negative charge can no longer sustain a new leader.

Intracloud discharges generally occur between an upper positive charge center and a lower negative charge center. A stepped-leader of the type occurring in ground discharges propagates between the two charge centers. There appear to be relatively weak return strokes

occurring as the leader contacts charge pockets of polarity opposite to that of the leader.

An approximate model will now be developed, showing how these strokes are reflected in the data recordings.

3.3 Electric Field and Propagation Model. Consider the current filament shown in Figure 4.

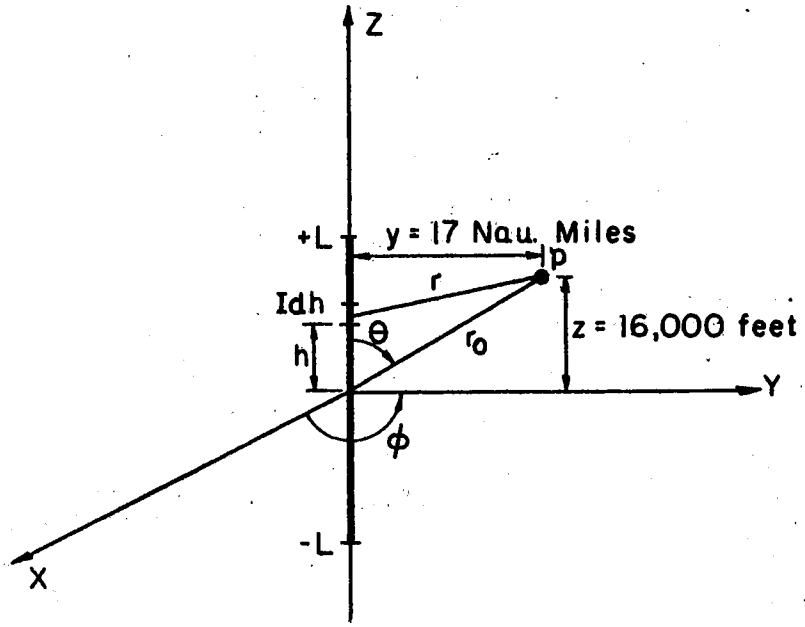


Figure 4. Coordinate System for the Propagation Model

Assume that the filament carries a current given by

$$I = I_0(e^{-\alpha t} - e^{-\beta t}) + I_1 e^{-\gamma t}$$

where

$$\alpha = 2.0 \times 10^4/\text{sec}$$

$$\beta = 2.0 \times 10^5/\text{sec}$$

$$\gamma = 1.0 \times 10^3/\text{sec}$$

$$I_0 = 30 \text{ KA}$$

$$I_1 = 2.5 \text{ KA} \quad .$$

Further, the length of the filament is assumed to grow at a rate given by

$$v = v_0 \, e^{-\nu t}$$

where

$$v_0 = 8 \times 10^4 \text{ km/sec}$$

$$\nu = 3 \times 10^4/\text{sec} \quad .$$

Using these parameters, the magnetic vector potential $\vec{A}$ can be written as

$$A_Z = \frac{\mu}{4\pi} \int_{-L}^{L} \frac{I(\frac{t-r}{c}) \, dh}{r} \quad ,$$

$$A_Z = \frac{\mu}{4\pi} \int_{-L}^{L} \{I_0[e^{-\alpha(\frac{t-r}{c})} - e^{-\beta(\frac{t-r}{c})}] + I_1 \, e^{-\gamma(\frac{t-r}{c})}\} \frac{dh}{r} \quad ,$$

where L is the length of the channel at time t-r/c and the lower limit introduces the effect of terrain reflection. Since $r = \sqrt{(z-h)^2 + y^2}$, the integral can be broken up into three separate integrals of the same form given by

$$\int_{-L}^{L} \frac{\mathrm{Exp}[\frac{\alpha}{c}\sqrt{(z-h)^2+y^2}]}{\sqrt{(z-h)^2+y^2}}\,dh \quad .$$

Let h be measured from Z, and expand the exponential in Taylor series form,

$$e^{\frac{\alpha}{c}\sqrt{h^2+y^2}} = 1 + \frac{\alpha}{c}\sqrt{h^2+y^2} + \frac{(\frac{\alpha}{c}\sqrt{h^2+y^2})^2}{2!} + \frac{(\frac{\alpha}{c}\sqrt{h^2+y^2})^3}{3!} + \ldots \quad .$$

The integral now appears

$$\int_{-L-z}^{L-z} \{\frac{1}{\sqrt{h^2+y^2}} + \frac{\alpha}{c} + \frac{(\frac{\alpha}{c})^2\sqrt{h^2+y^2}}{2!} + \frac{(\frac{\alpha}{c})^3(h^2+y^2)}{3!} + \frac{(\frac{\alpha}{c})^4\sqrt{(h^2+y^2)^3}}{4!} + \ldots\}dh.$$

Define the function $f(\frac{\alpha}{c})$ as follows:

$$f(\tfrac{\alpha}{c}) = \{\ell n(h + \sqrt{h^2+y^2}) + \tfrac{\alpha}{c}h + (\tfrac{\alpha}{c})^2\tfrac{1}{2!}$$

$$\cdot \tfrac{1}{2}[h\sqrt{h^2+y^2} + y^2\,\ell n(h + \sqrt{h^2+y^2})]$$

$$+ (\tfrac{\alpha}{c})^3\tfrac{1}{3!}(\tfrac{h^3}{3} + y^2h) + (\tfrac{\alpha}{c})^4\tfrac{1}{4!}$$

$$\cdot \tfrac{1}{4}[h\sqrt{(h^2+y^2)^3} + \tfrac{3y^2h}{2}\sqrt{h^2+y^2} + \tfrac{3y^4}{2}\ell n(h + \sqrt{h^2+y^2})]$$

$$+ (\tfrac{\alpha}{c})^5\tfrac{1}{5!}[\tfrac{h^5}{5} + \tfrac{2h^3y^2}{3} + y^4h] + \ldots\}\Big|_{-L-z}^{L-z} \quad .$$

Hence

$$A_z = \frac{\mu}{4\pi} \{I_0[e^{-\alpha t} f(\tfrac{\alpha}{c}) - e^{-\beta t} f(\tfrac{\beta}{c})] + I_1 e^{-\gamma t} f(\tfrac{\gamma}{c})\} \quad .$$

The magnetic field intensity can now be determined by the following relation,

$$\vec{H} = \frac{1}{\mu} (\nabla \times \vec{A}) \quad .$$

By observing that since P is in the Y-Z plane, the vector identity $1_\phi = -1_x$ can be used to obtain

$$H_\phi = -H_x = \frac{\partial A_z}{\partial y} \quad .$$

Using,

$$\frac{\partial f(\frac{\alpha}{c})}{\partial y} = \frac{1}{h + \sqrt{h^2+y^2}} \cdot \frac{y}{\sqrt{h^2+y^2}} + (\tfrac{\alpha}{c})^2 \frac{1}{2!} \frac{1}{2} [\frac{yh}{\sqrt{h^2+y^2}}$$

$$+ 2y \ln(h + \sqrt{h^2+y^2}) + y^2 \frac{1}{h + \sqrt{h^2+y^2}} \cdot \frac{y}{\sqrt{h^2+y^2}}] + (\tfrac{\alpha}{c})^3 \frac{1}{3!} 2yh$$

$$+ (\tfrac{\alpha}{c})^4 \frac{1}{4!} \frac{1}{4} [3yh \sqrt{h^2+y^2} + 3yh \sqrt{h^2+y^2} + \frac{3y^3h}{2} \frac{1}{\sqrt{h^2+y^2}}$$

$$+ \frac{12y^3}{2} \ln(h + \sqrt{h^2+y^2}) + \frac{3y^4}{2} \frac{1}{h + \sqrt{h^2+y^2}} \cdot \frac{y}{\sqrt{h^2+y^2}}]$$

$$+ (\tfrac{\alpha}{c})^5 \frac{1}{5!} [\frac{4h^3y}{3} + 4y^3h] + \ldots \Big]_{-L-z}^{L-z}$$

the magnetic field intensity can then be computed from

$$H_\phi = -\frac{1}{4\pi}\{I_0[e^{-\alpha t}\frac{\partial f(\frac{\alpha}{c})}{\partial y} - e^{-\beta t}\frac{\partial f(\frac{\beta}{c})}{\partial y}] + I_1 e^{-\gamma t}\frac{\partial f(\frac{\gamma}{c})}{\partial y}\} \quad .$$

The electric field intensity is determined from

$$\vec{E} = \frac{1}{\varepsilon_0}\int (\nabla \times \vec{H})\, dt$$

and from the definition of $(\nabla \times \vec{H})$, the E_z component is

$$E_z = \frac{1}{\varepsilon_0}\int \frac{1}{y}\frac{\partial}{\partial y}(y H_\phi)\, dt \quad .$$

Expansion of the inegrand results in

$$E_z = -\frac{1}{4\pi\varepsilon_0}\int \frac{1}{y}\{I_0[e^{-\alpha t}\frac{\partial f(\frac{\alpha}{c})}{\partial y} - e^{-\beta t}\frac{\partial f(\frac{\beta}{c})}{\partial y}] + I_1 e^{-\gamma t}\frac{\partial f(\frac{\gamma}{c})}{\partial y}\}\, dt$$

$$-\frac{1}{4\pi\varepsilon_0}\int \{I_0[e^{-\alpha t}\frac{\partial^2 f(\frac{\alpha}{c})}{\partial y^2} - e^{-\beta t}\frac{\partial^2 f(\frac{\beta}{c})}{\partial y^2}] + I_1 e^{-\gamma t}\frac{\partial^2 f(\frac{\gamma}{c})}{\partial y^2}\}\, dt \quad .$$

The expression for E_z is approximated using a formula offered by Hildebrand (9),

$$\int_{x_0}^{x_2} f(x)\, dx = \frac{h}{3}(f_0 + 4f_1 + f_2) - \frac{h^5}{90} f^{IV}(\xi) \qquad (x_0 < \xi < x_2)$$

where f(x) is an arbitrary function which is Riemann integrable with respect to x, this expression being essentially Simpson's rule. Let the sample separation be 5 μsec and $x_0 = r_0/c$, taking into account a finite propagation interval and compute,

$$\frac{\partial^2 f(\frac{\alpha}{c})}{\partial y^2} = \frac{h\sqrt{h^2+y^2} + h^2+y^2 - y\left[\frac{yh}{\sqrt{h^2+y^2}} + 2y\right]}{[h\sqrt{h^2+y^2} + h^2+y^2]^2}$$

$$+ (\tfrac{\alpha}{c})^2 \tfrac{1}{4} \left\{ \frac{h\sqrt{h^2+y^2} - \frac{y^2 h}{\sqrt{h^2+y^2}}}{h^2+y^2} + 2\,\ell n(h + \sqrt{h^2+y^2}) \right.$$

$$\left. + \frac{2y^2}{(h + \sqrt{h^2+y^2})\sqrt{h^2+y^2}} + \frac{(h\sqrt{h^2+y^2} + h^2+y^2)2y^2 - y^3\left[\frac{yh}{\sqrt{h^2+y^2}} + 2y\right]}{[h\sqrt{h^2+y^2} + h^2+y^2]^2} \right\}$$

$$+ (\tfrac{\alpha}{c})^3 \tfrac{1}{3} h + (\tfrac{\alpha}{c})^4 \tfrac{1}{32} \left\{ 2h\sqrt{h^2+y^2} + \tfrac{7}{2} y^2 h \frac{1}{\sqrt{h^2+y^2}} - \frac{y^4 h}{2} \frac{1}{\sqrt{(h^2+y^2)^3}} \right.$$

$$+ 6y^2\,\ell n(h + \sqrt{h^2+y^2}) + 2y^3 \frac{1}{h + \sqrt{h^2+y^2}} \cdot \frac{y}{\sqrt{h^2+y^2}}$$

$$\left. + \frac{1}{2} \frac{1}{h + \sqrt{h^2+y^2}} \cdot \frac{y}{\sqrt{h^2+y^2}} + \frac{(h\sqrt{h^2+y^2}+h^2+y^2)5y^4 - y^5\left[\frac{yh}{\sqrt{h^2+y^2}} + 2y\right]}{2[h\sqrt{h^2+y^2} + h^2+y^2]^2} \right\}$$

$$+ (\tfrac{\alpha}{c})^5 \tfrac{1}{30} [\tfrac{h^3}{33} + 3y^2 h] + \ldots \Big]_{-L-z}^{L-z} .$$

A computer program was written to carry out this simulation. The

power spectral density is shown in Figure 5 before the effects of atmospheric attenuation are taken into account, and after these effects are considered. The propagation characteristic of the atmosphere used here, is reported by Horner (12) and was obtained empirically by studying the spectra of the same atmospheric recorded at widely spaced stations. Theoretically, the phenomena of low frequency propagation have been explained in terms of waveguide modes of propagation in the space between the earth and the ionosphere. Propagation is dependent upon the height of reflection in the ionosphere and on the electrical properties of the ionosphere and the ground. Models of the waveguide mode of propagation are generally based on the assumption that the ionosphere and the surface of the earth are smooth, sharp boundaries and that the earth's magnetic field can be neglected. A survey of these studies is offered by Horner (12), pp. 148-149. Finally in Figure 6 the effects of the instrumentation are accounted for. These two figures represent the essence of the problem in studying a phenomenon such as thunderstorm noise, namely that of discerning the field generated by the storm cell from the signal which is recorded by the measurement apparatus. Clearly the electromagnetic field experiences a number of perturbing influences before reaching the sensing instrumentation. After reaching the measurement system, the signal is further modified by the system response. Ideally, if one could characterize all these influences, the process could be reversed, and there would result a pure and unperturbed description of the thunderstorm electromagnetic field.

In summary, the central thought in this chapter has been that of demonstrating how the recorded data is related to the actual magnetic signature of a thunderstorm. Modifications taken into account were

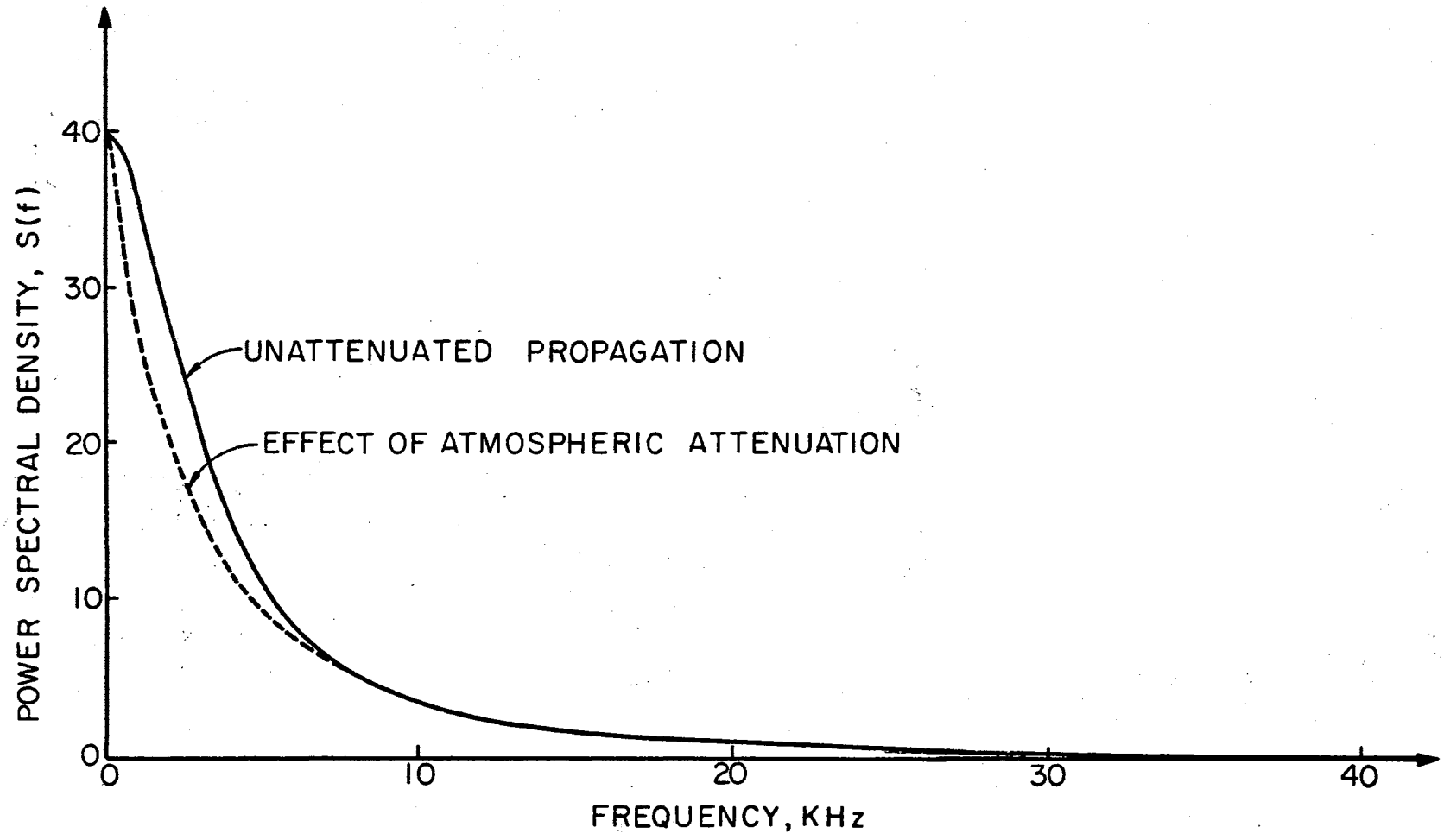


Figure 5. Effects of Atmospheric Attenuation

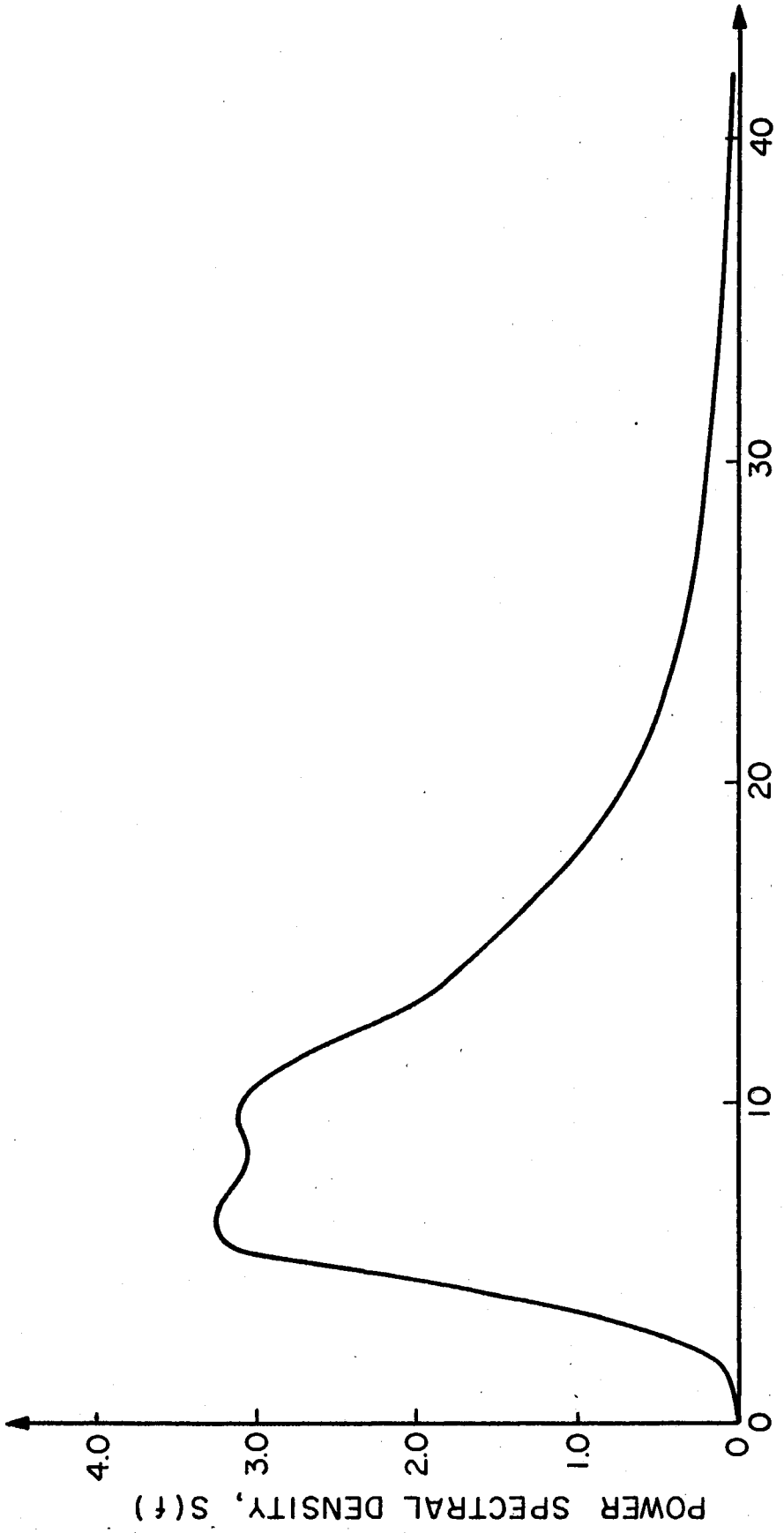


Figure 6. Effects of Instrumentation

first of all, a reflection from the earth's surface using image theory in the field equation for magnetic vector potential. Secondly, the signature was altered in accordance with the effects of atmospheric attenuation as reported by Horner (12), and it appears that this effect is negligible above 7 KHz. Finally, the field is modified by taking into account the calibration curve of the data gathering instrumentation. The thoughts developed above, then, will be reconsidered in the discussion of the results of data reduction.

CHAPTER IV

AN INVESTIGATION OF SFERIC STATIONARITY

4.1 Introduction. It would appear that the study discussed in this chapter provides an a priori justification for the stationary approach used in the data reduction of Chapter V. Actually, the data were first reduced, and the results indicated that the stationary assumption might be justified. The three data segments which are investigated in this chapter, then, were selected through a cut-and-try application of the data reduction technique; however, the natural order of logic, as in a mathematical proof, dictates that the stationary argument precede the discussion of stationary data reduction. The study of sferic stationarity is based on the assumption of a shot noise model. The data are analyzed to determine whether or not this model reasonably describes the storm signature.

4.2 Statement of the Stationary Hypothesis. Stated in its simplest terms, the hypothesis being tested is that the sferic trigger mechanism can be modeled by a Poisson impulse process, whose parameter is independent of time. The investigation is separated into a two step inquiry: first, is the Poisson impulse model a reasonable description; and secondly, if the conclusion to the first part is in the affirmative, is the Poisson parameter constant over the interval of interest? Before embarking upon the study of the data themselves, a model will be developed which illustrates the hypothesis in greater detail. Then, it will

be demonstrated how the investigation of the data will be carried out.

Consider the block diagram shown in Figure 7. The signal z(t) is the sferic triggering mechanism and s(t) is the signal recorded by the instrumentation discussed in Chapter II. The network h(t), then, represents the waveform of the sferic, its modification by propagation, and the effects of the measurement system. The problem is one of relating the statistics of s(t) with those of z(t), and as a preliminary step in the study, this will be done on a theoretical basis. In this way, the full extent of the hypothesis under test will be illustrated.

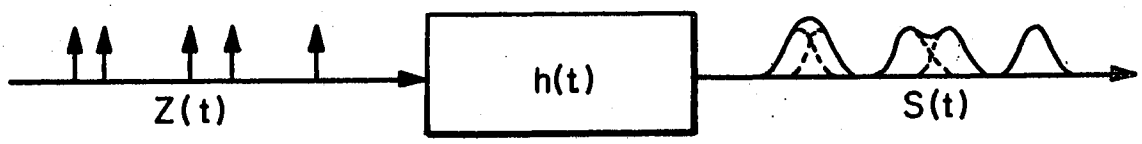


Figure 7. Simplified Block Diagram

First, the statistics of z(t) are formulated. Let z(t) be a sequence of impulses which is characterized mathematically as follows,

$$z(t) = \sum_i \delta(t - t_i) \quad .$$

Further, it is hypothesized that for any interval (t_1,t_2), the probability of finding k impulses in that interval is

$$P\{k \text{ in } (t_1,t_2)\} = e^{\int_{t_1}^{t_2} \lambda(t)dt} \frac{\left[\int_{t_1}^{t_2} \lambda(t)dt\right]^k}{k!} \qquad k = 0,1,2,\ldots \quad .$$

This is the Poisson probability density function with parameter $\lambda(t)$. If the average incidence of impulses varies as a function of time, then $\lambda(t)$ will be a function of time and the process, $z(t)$, is called a Poisson impulse process with a nonuniform density in the rate of occurrence. If the average incidence of impulses is the same for every (t_1,t_2) in the time interval of interest, then $\lambda(t)$ is time invariant and the process $z(t)$, exhibits wide sense stationarity. Specifically, a process is said to be wide sense stationary if its mean value is a constant and its autocorrelation, $R(t_1,t_2)$, depends only on $t_1 - t_2$, i.e.

$$E\{z(t)\} = \mu = \text{constant}$$

$$E\{z(t_1)\ z(t_2)\} = R(t_1 - t_2) \quad .$$

To show that $z(t)$ is wide sense stationary for constant λ on the interval (t_1,t_2), consider its mean and autocorrelation as given by Papoulis (17).

$$E\{z(t)\} = \lambda(t)$$

$$R_{zz}(t_1,t_2) = \lambda(t_1)\ \lambda(t_2) + \lambda(t_2)\ \delta(t_1 - t_2) \quad .$$

If λ is time invariant, then

$$E\{z(t)\} = \lambda$$

$$R_{zz}(t_1,t_2) = \lambda^2 + \lambda\delta(t_1 - t_2) \quad .$$

These statistics will be used to determine the statistics of $s(t)$.

$s(t)$ is related to $z(t)$ by $h(t)$, and it is characterized mathematically by

$$s(t) = \sum_i h(t - t_i) \quad .$$

To formulate the statistics of s(t), assume that h(t) is an ideal bandpass filter whose impulse response is,

$$h(t) = \frac{\sin at}{t} \cos\omega_0 t \qquad a << \omega_0 \quad .$$

The Fourier transform of h(t) is as shown in Figure 8. The bandpass assumption seems reasonable in view of the instrumentation characteristic described in Chapter II, i.e. no DC component. Thus, s(t) is called shot noise, and for nonuniform $\lambda(t)$, its mean and autocorrelation are given by Papoulis (17),

$$E\{s(t)\} = \lambda(t) * h(t) = \int_{-\infty}^{\infty} \lambda(t - \tau)\, h(\tau)\, d\tau$$

$$R_{ss}(t_1,t_2) = [\lambda(t_1) * h(t_1)][\lambda(t_2) * h(t_2)] + [\lambda(t_2)\, h(t_1 - t_2)] * h(t_2)$$

where * denotes convolution. If λ is time invariant,

$$E\{s(t)\} = \lambda H(0)$$

$$R_{ss}(\tau) = \lambda^2 H^2(0) + \lambda \int_{-\infty}^{\infty} h(\tau + \beta)\, h(\beta)\, d\beta \quad .$$

Since H(0) = 0, then

$$E\{s(t)\} = 0$$

$$R_{ss}(\tau) = \lambda \int_{-\infty}^{\infty} h(\tau + \beta)\, h(\beta)\, d\beta \quad .$$

Thus, if it can be shown that z(t) is a Poisson impulse process with uniform density and that h(t) is a linear time invariant system, a case

is developed justifying the reduction of thunderstorm data, s(t), using procedures applicable to stationary analysis. Specifically, R_{ss} would be a function of one variable $\tau = |t_1 - t_2|$ and not two variables, (t_1, t_2).

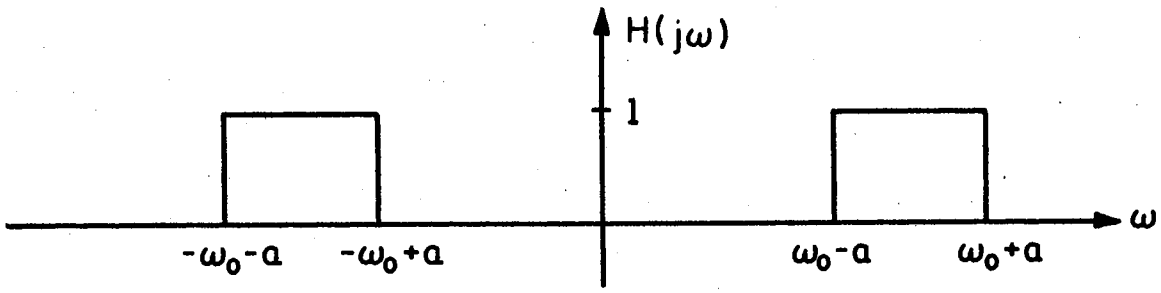


Figure 8. Idealized Bandpass Filter

An alternate statistic which can be used in testing for constant λ is the waiting time between impulses (For a detailed derivation of the following development, the interested reader is referred to Hogg and Craig (10), pp. 91-92.). Let ω denote the time between impulses as in Figure 9, such that

$$\omega_i = t_i - t_{i-1} \quad .$$

If the ω_i's are mutually independent, with uniform density, then

$$f_w(\omega) = \lambda e^{-\lambda\omega} U(\omega) \quad .$$

Thus,

$$F_w(\omega) = \begin{cases} 0 & \omega < 0 \\ 1-e^{-\lambda\omega} & \omega \geq 0 \end{cases} .$$

This implies that

$$\lambda\omega = -\ell n[1 - F_w(\omega)] \qquad \omega > 0$$

and if $-\ell n[1 - F_w(\omega)]$ is plotted as a function of ω, the result will be a straight line with a slope of λ, if λ is constant.

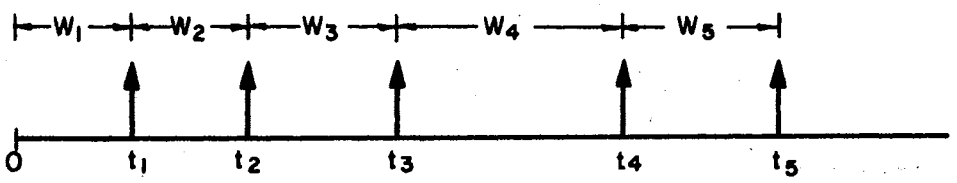


Figure 9. Waiting Time Between Impulses

To conduct the test of the stationary hypothesis, the data are subdivided into a multiplicity of segments. Each of the three thunderstorm records are 0.75 seconds in duration, and each has been quartered to form four segments of 187.5 milliseconds. Each of these segments has been further subdivided into six strips of 31.25 milliseconds. The original reason for this particular subdivision was to reproduce the data on 8 x 10 glossy prints from 35 mm film; it resulted that this also produced convenient segments for the investigation of stationarity.

The essence of the problem in conducting this investigation is that of identifying the separate impulses in a burst of sferic activity.

For example, a typical segment of data is shown in Figure 10. Consider Pulse #1, lasting typically 250 microseconds. It seems particularly attractive to assert that this pulse resulted from one triggering impulse. Pulse #2 is typical of the record of a continual burst of activity lasting usually 2 to 3 milliseconds. It does not seem as likely as for Pulse #1 that there exists a single triggering impulse. If an arbitrary threshold were established such that for every excursion the data makes above the threshold, an impulse is assumed to have occurred, then the impulses are not too dense to be credible for a pulse of type #3; however, for Pulse #2 an unreasonable result would be obtained.

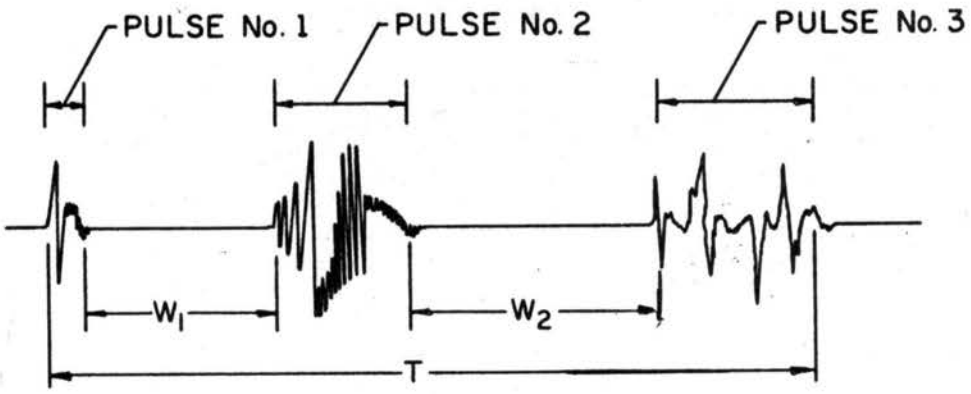


Figure 10. Typical Segment of Thunderstorm Data

It appears desirable to treat all three pulse types as having each resulted from single triggering impulses. Although some impulses will not be accounted for, (i.e. Pulse #3 may have resulted from three triggering impulses close in time) the preponderance of the data is composed of pulses like Pulse #1, and hence, the percentage error will be small.

As a check on the validity of this argument, the pulse widths will be recorded along with the estimates of λ. An implication of this viewpoint is that estimates of λ, based on a pulse count, will almost certainly reflect a somewhat lower density than that which actually exists.

One estimate of λ will be generated on each record of 0.75 seconds of storm data. This will be computed from

$$\hat{\lambda} = \frac{\text{number of pulses}}{T = 0.75 \text{ sec}} \quad .$$

Using $\hat{\lambda}$ and the Poisson density function, a theoretical plot of the pulse distribution can be made for the constant λ case. In addition, for every 187.5 millisecond segment, an empirical density function will be plotted by counting the number of pulses in each 31.25 milliseconds of the data. The empirical density function will be compared with the theoretical density function to test the Poisson hypothesis.

As a check on these two measurements, i.e. the pulse width and number of pulses, a third and final measurement is to be considered. It is the waiting time between pulses. This second estimate of λ can be obtained from an empirical distribution of the waiting time by plotting $-\ell n[1 - F_w(\omega)]$ as a function of ω. If this results in a straight line, then λ is constant and equals the slope of the line. This computation will be carried out for every 187.5 milliseconds of storm data. Hence, four curves will result for each record. It should be pointed out that the measurement of waiting time is made as illustrated in Figure 10. This is a convenient method of measurement because, first of all, the quiet periods are very well defined and, secondly, this avoids the quandry of where the impulse occurred (i.e. where within the burst) or even if a multiplicity of impulses occurred. Because the waiting

times are not referenced to points in time, rather, separated by the pulse widths, it is expected that the resulting estimate of λ will be higher in value than that which actually exists. Certainly, it will be larger than the estimate produced by counting pulses.

Summarizing, the investigation of stationarity proceeds from the premise that a large number of the pulse widths are short with respect to the measurement interval. Using these pulses, λ is estimated in two ways and an empirical density function is estimated. These estimates are then used to see if (1) the process can be modeled by a Poisson triggering mechanism and (2) if the Poisson trigger has a constant parameter.

4.3 Results of the Study. The three measurements mentioned above were taken from the analog data of Record #1 and are shown graphically in Figures 11 through 16. Overall, 164 pulses were identified and, of these pulses, approximately 75 percent had a width of one millisecond or less. Because of this relatively narrow pulse width, the 31.25 millisecond base period seemed reasonable. The pulse widths as a function of the 187.5 millisecond quartiles of Record #1 are shown in Figure 11. If the 164 pulses were uniformly spread over the record, one would expect a Poisson distribution with λ = 219/sec. For this λ and t = 31.25 milliseconds, the Poisson density function would peak at 6 pulses. Shown in Figure 12 is this plot in solid lines as compared with the dotted histogram obtained empirically. The first three segments are suggestive of a Poisson density function peaking at 7 pulses while the fourth segment indicates a peak around 4 or 5 pulses. From these plots, it is concluded that the Poisson model may be valid; however, the constancy of λ is suspect. In an effort to resolve this point, attention

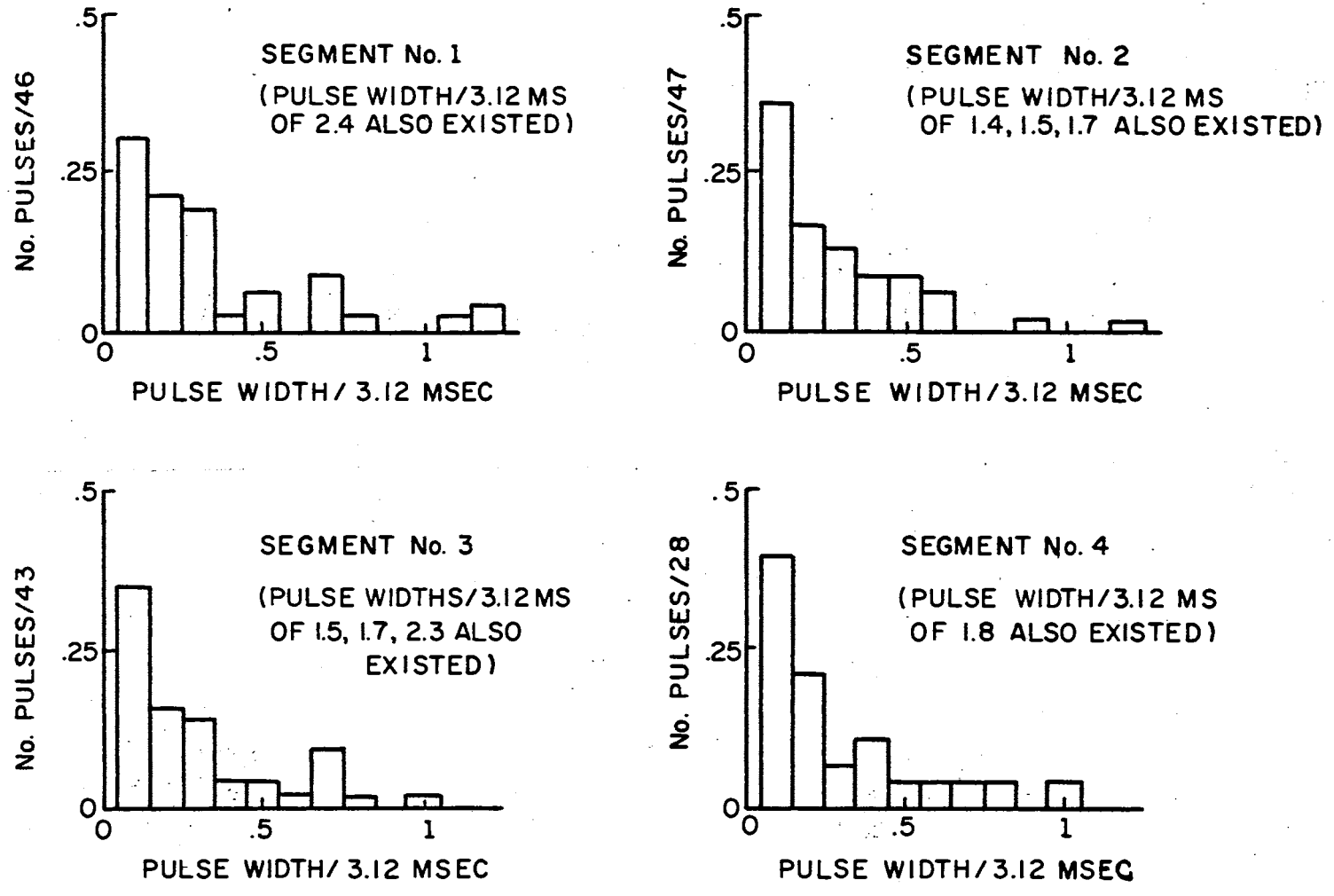


Figure 11. Distribution of Pulse Widths for Record #1

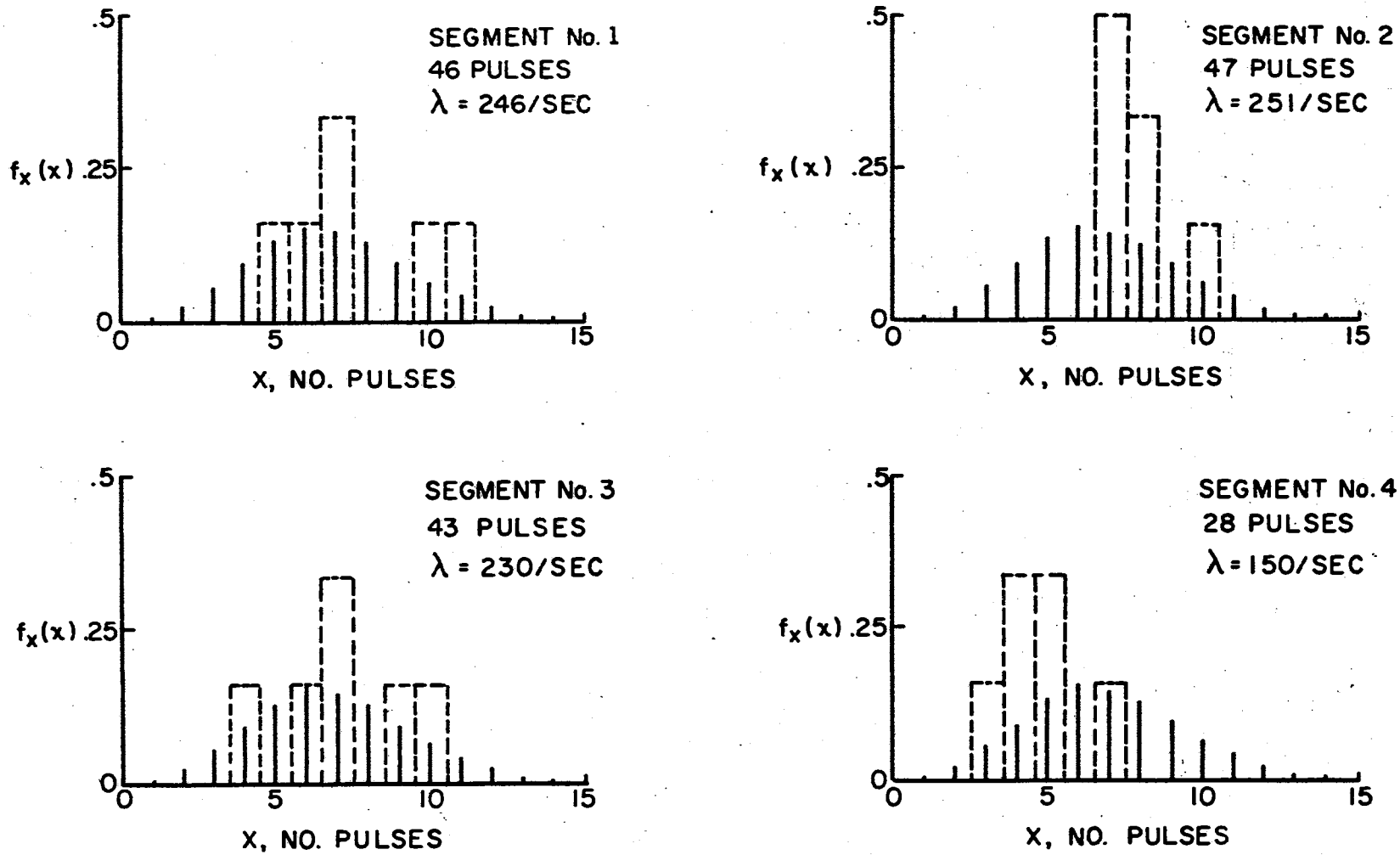


Figure 12. Distribution of Pulses for Record #1

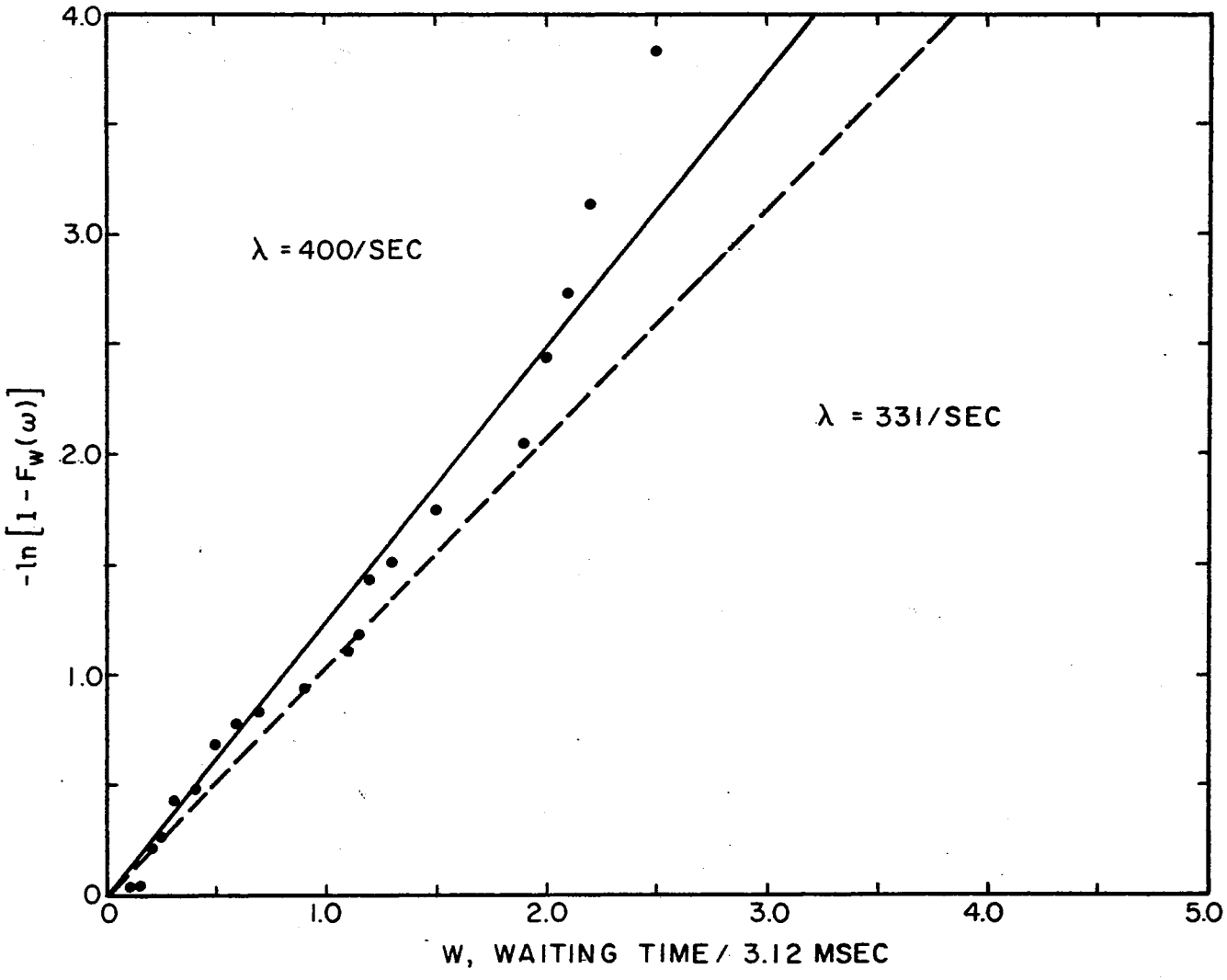


Figure 13. Distribution of Waiting Times, Segment 1, Record #1

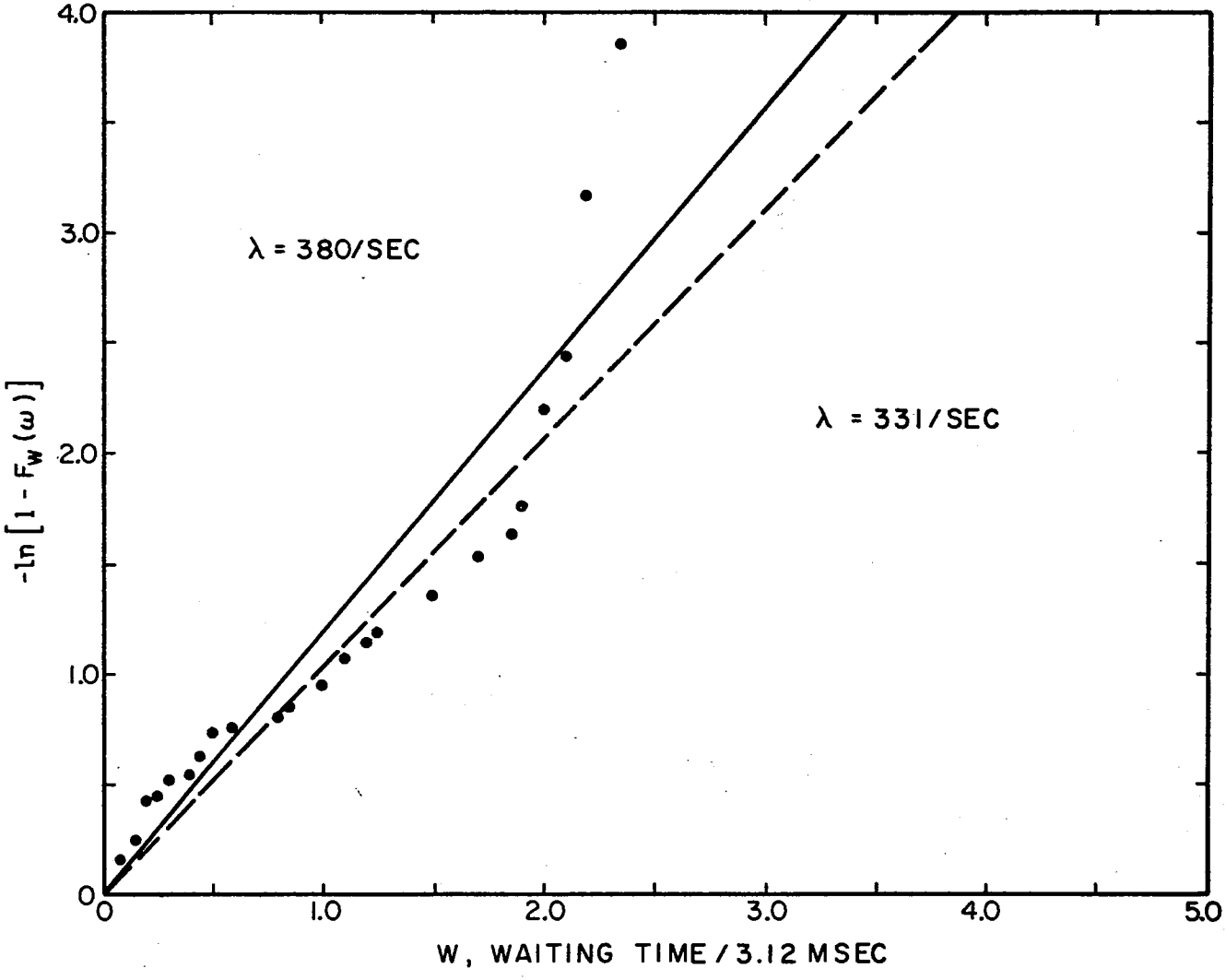


Figure 14. Distribution of Waiting Times, Segment 2, Record #1

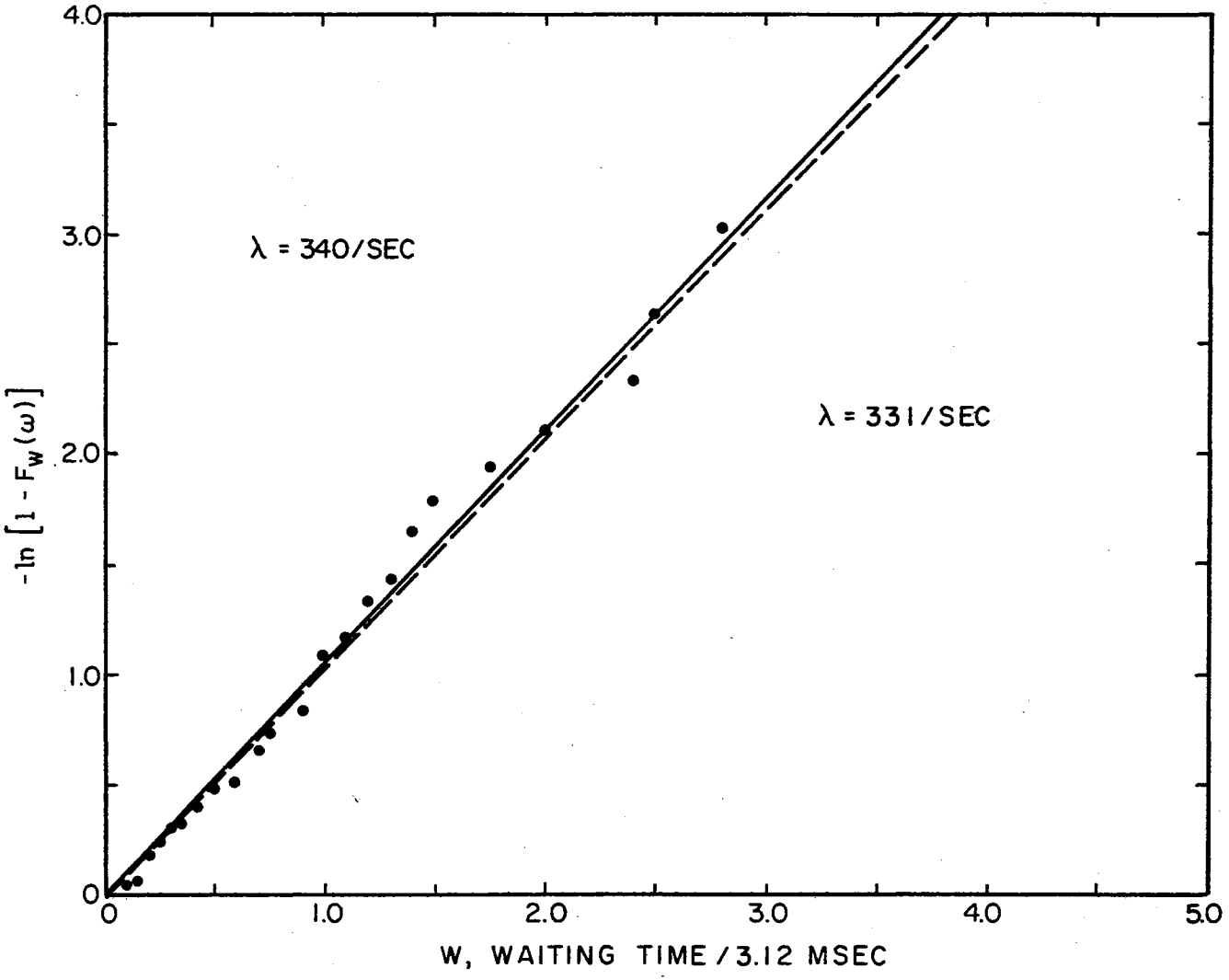


Figure 15. Distribution of Waiting Times, Segment 3, Record #1

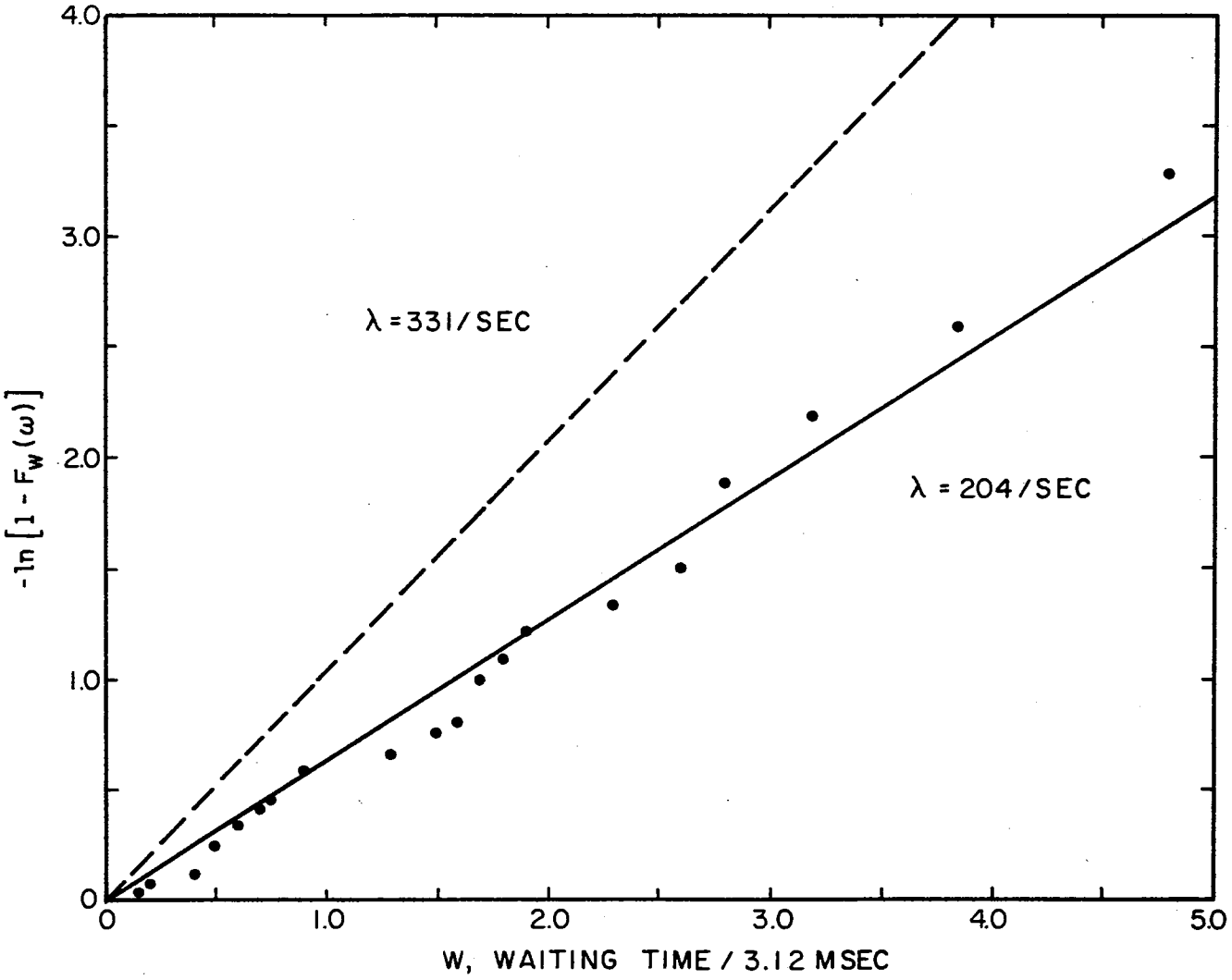


Figure 16. Distribution of Waiting Times, Segment 4, Record #1

is directed to Figures 13 through 16. These are plots of $-\ell n[1 - F_W(\omega)]$ as a function of waiting times for the four segments of Figure 12. A straight line fit was approximated for each set of data points, and an estimate of λ was obtained by computing the slope of the line. The straight line approximation is shown as a solid line in each graph, while a second line, shown dashed, was obtained by averaging the four λ's. As expected, the λ estimated by this approach is somewhat greater in value than that estimated by counting pulses. Viewing these four graphs in a composite sense, it is observed that the straight line approximations of Figures 13 and 14 are probably the least appropriate. The data appear linear for waiting times of 6.25 milliseconds or less, but the slope increases sharply above this point. A plausible explanation of this effect could lie in a change in λ within the data segments. Another possible explanation is that these sharp slopes occur at the tails of the distribution function, and may thus be statistical fluctuations not unexpected of a small sample size.

Consider the graphs of Figures 15 and 16. Although the straight line approximations seem to provide a reasonable fit for the data in these segments, there exists a fairly significant difference in the slope of the lines. This indicates a substantial change in the density of pulses. Thus, it is believed that a claim of stationarity based on these two segments of data, alone, is unreasonable.

Reviewing the data presented in light of the preceeding discussion, the following is concluded: (1) A Poisson model of the triggering mechanism seems reasonable, particularly if $\lambda = \lambda(t)$; (2) the case for constant λ, in a strict sense, lacks credibility; however, (3) it appears that the first three segments of data can be approximated by a

constant λ, and in spite of the sharp drop in pulse distribution, there is a large portion of the record over which a stationary Poisson approximation seems reasonable. This conclusion is reached by observing the apparent credibility of the straight line approximations to the waiting time data and using the assumption of mutual independence of the waiting times. Hence, it is believed that stationary methods of analysis will produce credible results for Record #1.

The corresponding measurements for Record #2 are presented graphically in Figures 17 through 22. Comparing the distribution of pulse widths shown in Figure 17 with the corresponding measurements of Record #1, it is noted that of the 177 pulses identified in Record #2, approximately 75 percent of the widths were 600 microseconds or less. If the 177 pulses were uniformly spread over the record, one would expect a Poisson distribution with λ = 236/sec. If t were taken to be 31.2 milliseconds, this would imply a peak in the Poisson density function at 7 pulses. Shown in Figure 18 is this plot in solid lines as compared with the dotted histograms obtained empirically. The histograms do not present the case for a Poisson distribution as convincingly as does the corresponding plot for Record #1. This is due primarily to the substantial variation between the data segments of the number of pulses and the pulse widths. It is believed that because of the one or two lengthy pulses found in each of the first three segments, a multiplicity of triggering impulses are hidden within the pulses. Thus, based on the graphs of Figure 18, it appears that the results are inconclusive. However, the quandry can be resolved by a review of the distributions of waiting times shown in Figures 20 through 22. Generally, the straight line approximations seem to provide a fairly credible fit to

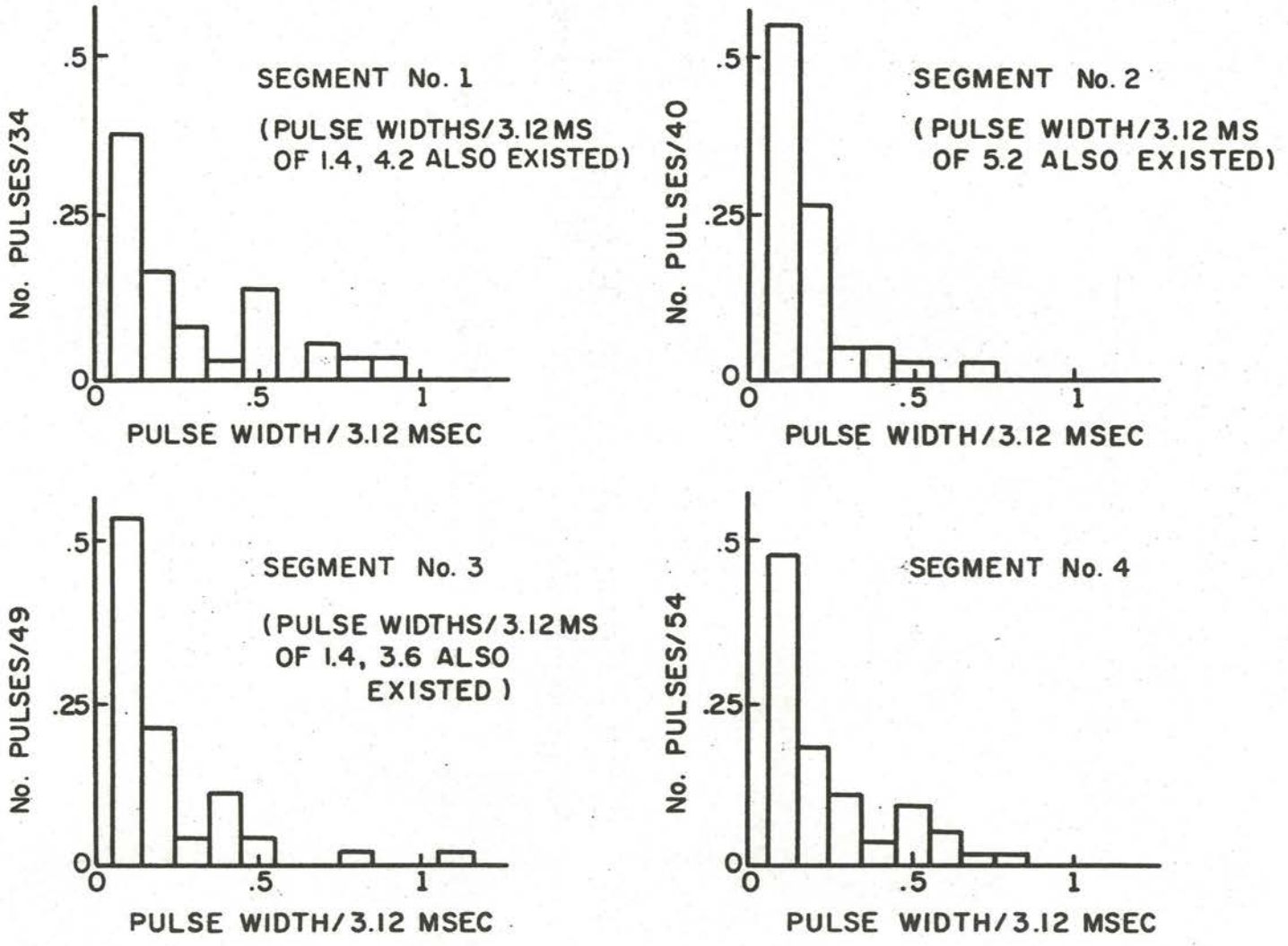


Figure 17. Distribution of Pulse Widths for Record #2

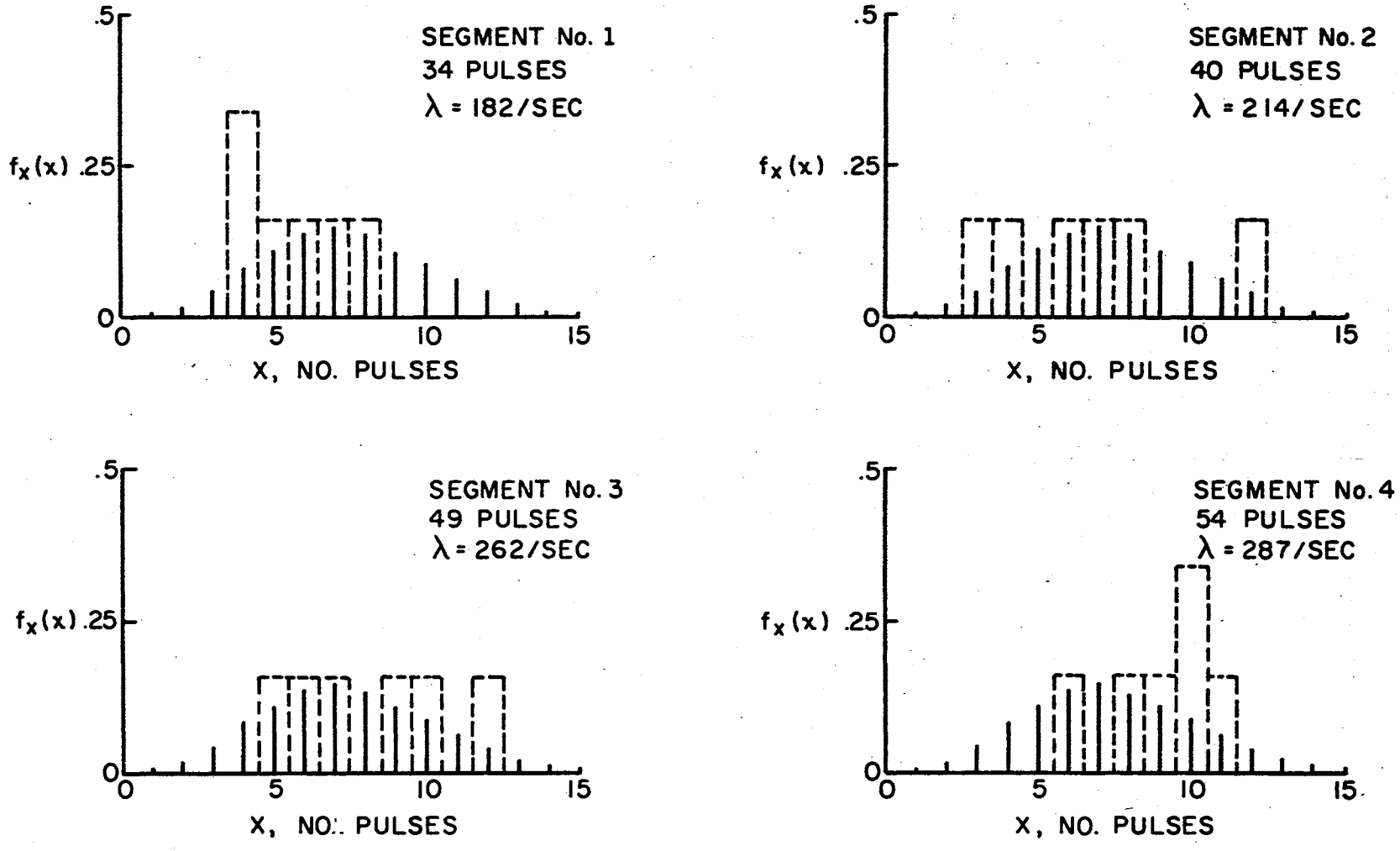


Figure 18. Distribution of Pulses for Record #2

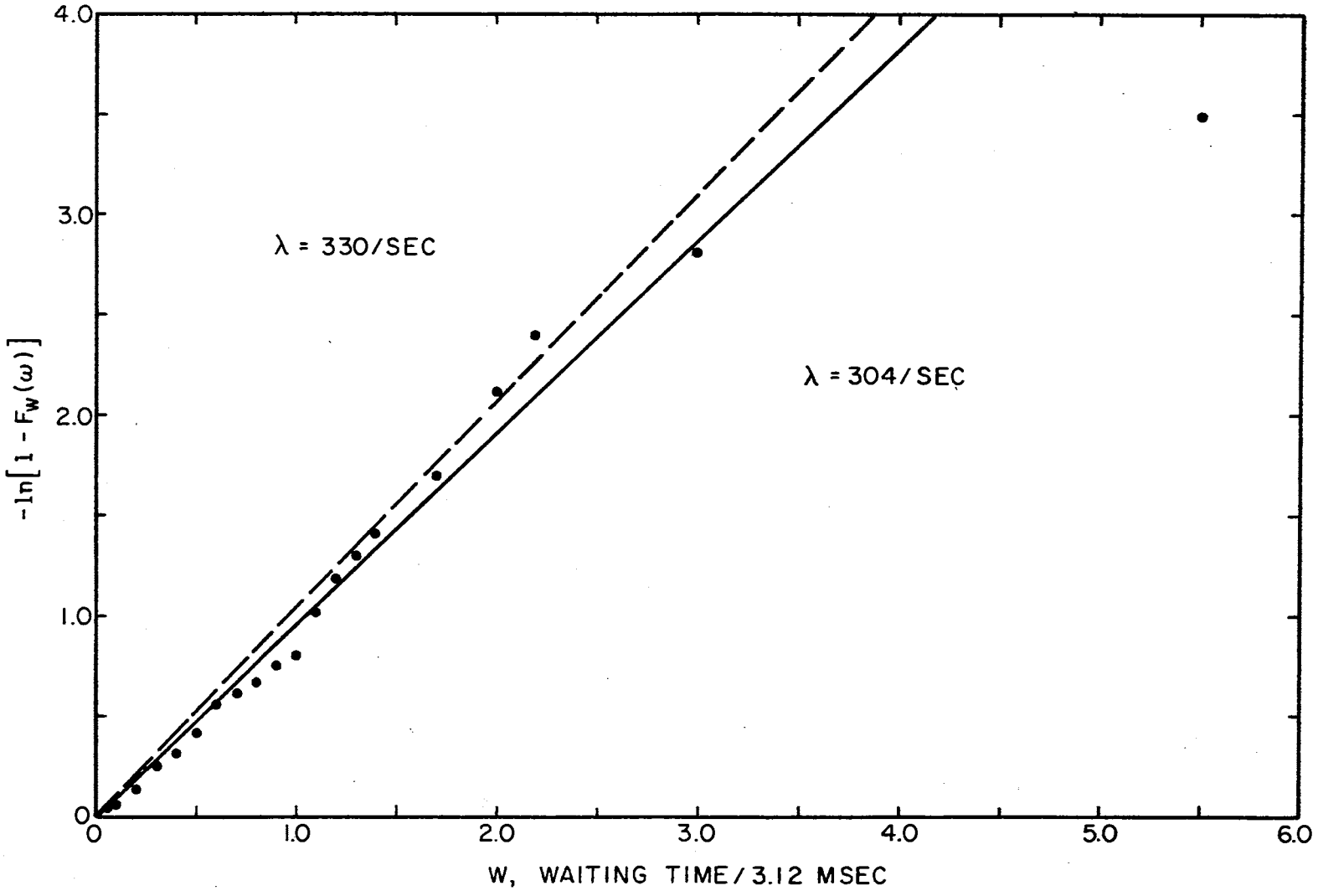


Figure 19. Distribution of Waiting Times, Segment 1, Record #2

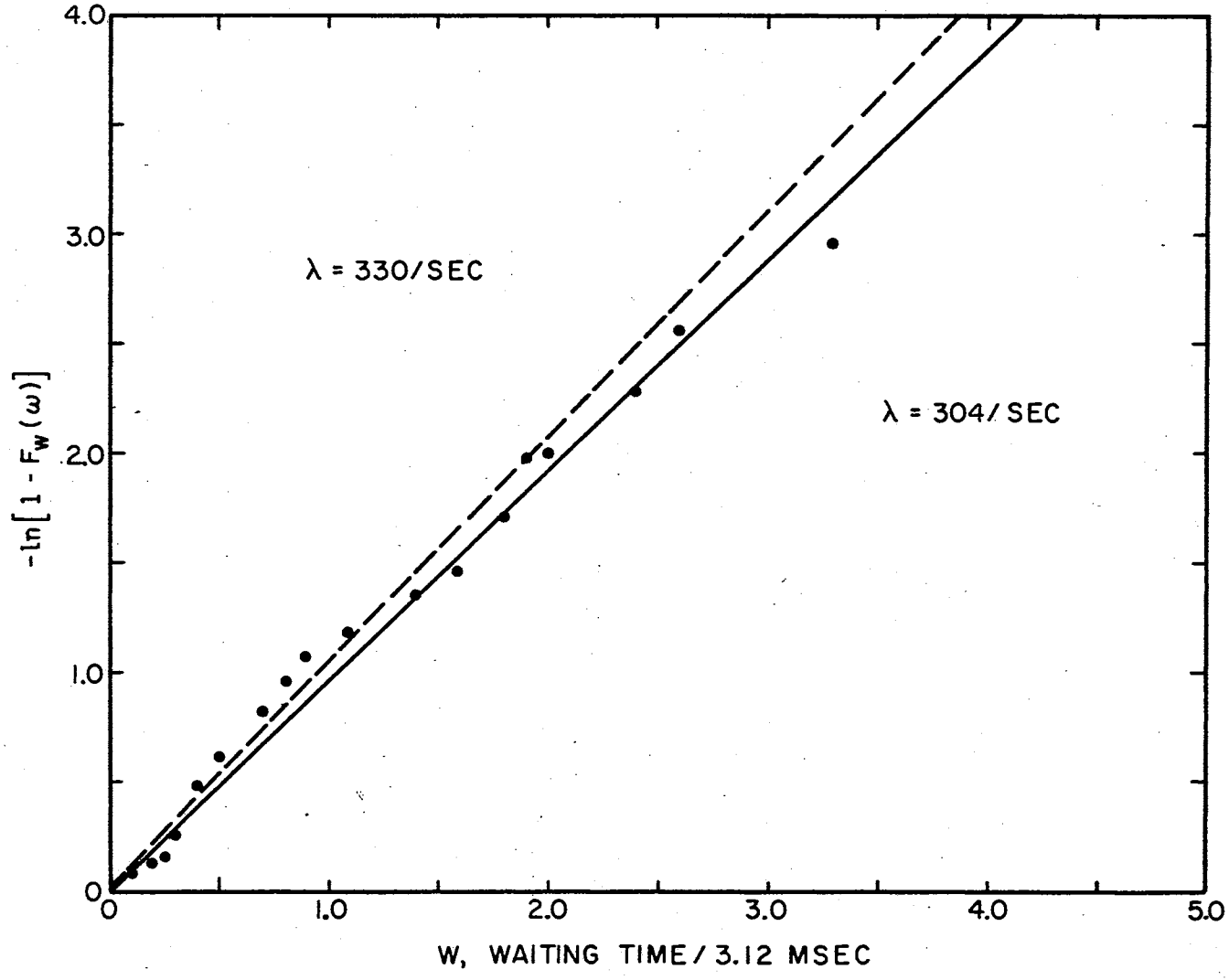


Figure 20. Distribution of Waiting Times, Segment 2, Record #2

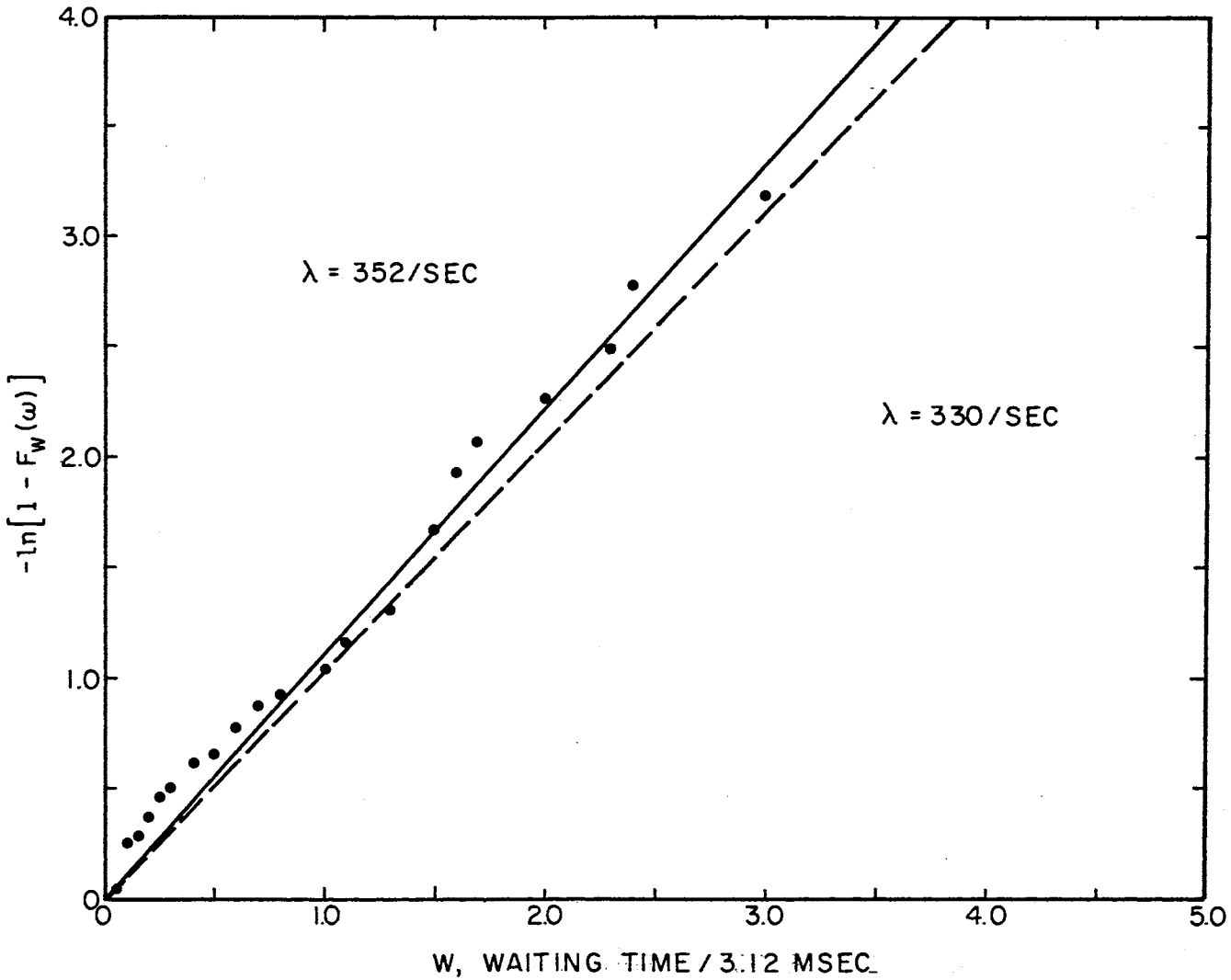


Figure 21. Distribution of Waiting Times, Segment 3, Record #2

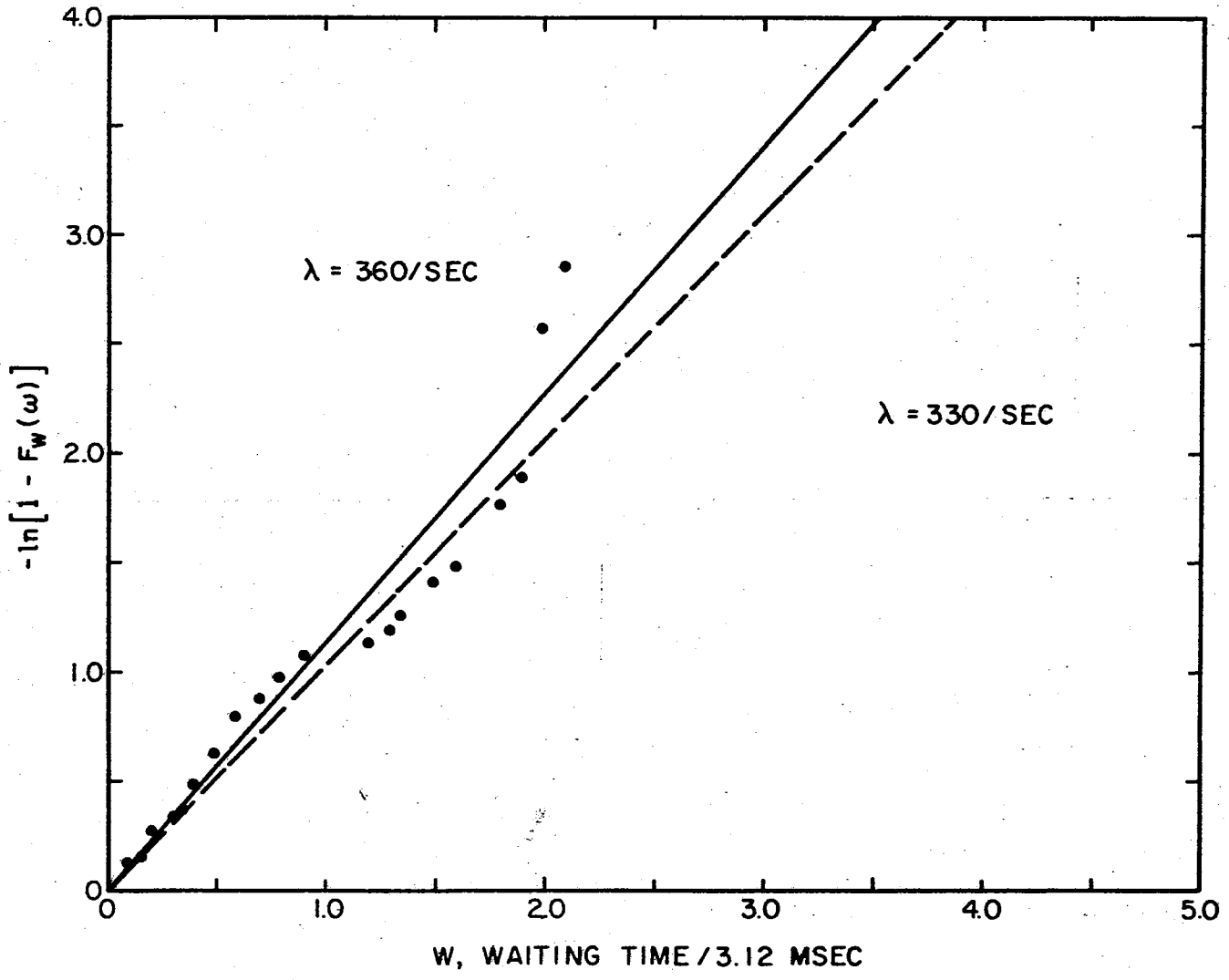


Figure 22. Distribution of Waiting Times, Segment 4, Record #2

the data, the worst case being the fourth segment. It appears that the final data segment is exhibiting a nonstationarity of the type found in Record #1; however, it seems that the straight line fit is nonetheless reasonable. Because of the relatively close agreement of the slope of the lines to a median slope, within say $\pm$ 10%, there appears to be a fairly credible case for a stationary model. Reviewing the data presented in a composite sense, the following is concluded: (1) again, there seems to be a reasonable likelihood that the trigger mechanism can be modeled by a Poisson distribution, and (2) there appears to be a reasonable case for an assumption of uniform density, or constant λ, throughout the record. Hence, it is concluded that Record #2 is also amenable to analysis in a stationary sense.

A final set of measurements for Record #3 are presented graphically in Figures 23 through 28. A comparison of the distribution of pulse widths shown in Figure 23, with the corresponding measurements of the two previous records, reveals that of the 197 pulses identified in Record #3, approximately 75 percent of the widths were 600 microseconds or less. If the 197 pulses were uniformly spread over the record, one would expect a Poisson distribution with λ = 263/sec. If t were taken to be 31.2 milliseconds, this would imply a peak in the Poisson density function at 8 pulses. This plot is shown in Figure 24 in solid lines as compared with the dotted histograms obtained empirically. Viewing the histograms, it appears that the Poisson assumption is not altogether unreasonable; however, a constant λ cannot be supported by this evidence alone. To further investigate the behavior of λ, attention is directed to the distribution of waiting times shown in Figures 25 through 28. Overall, it is believed that the straight line approximations in these

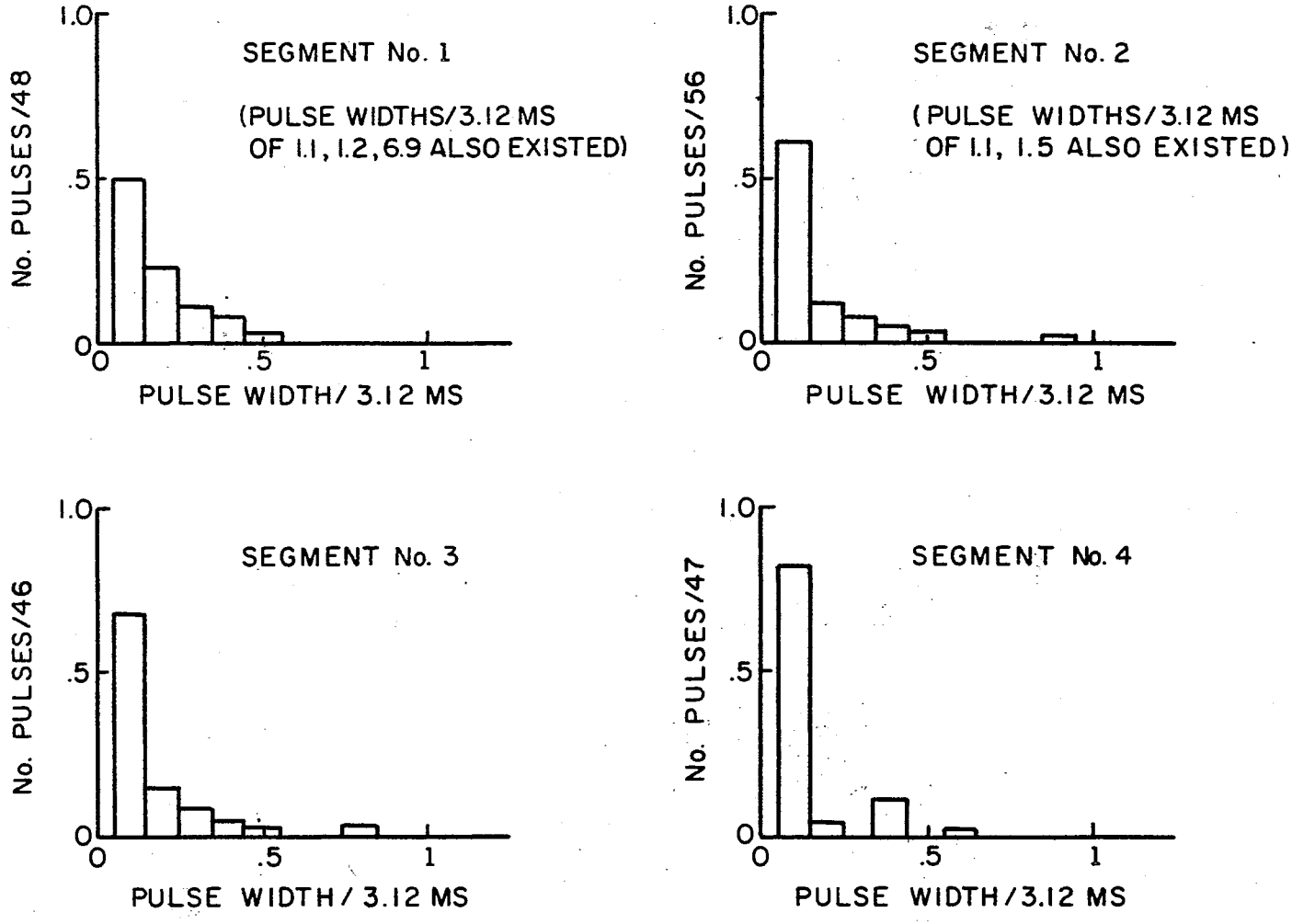
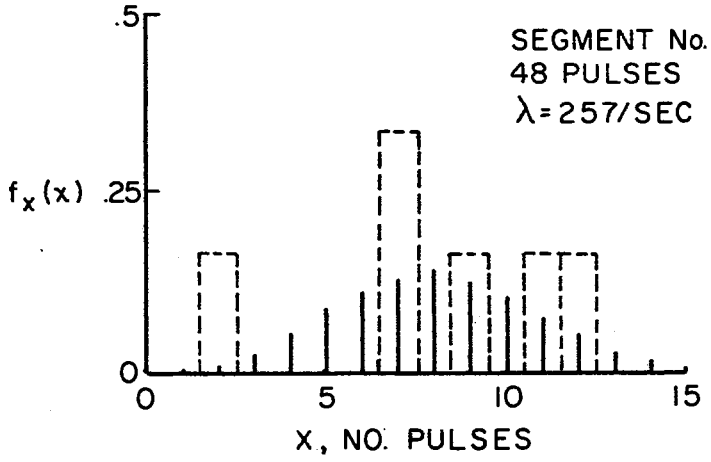
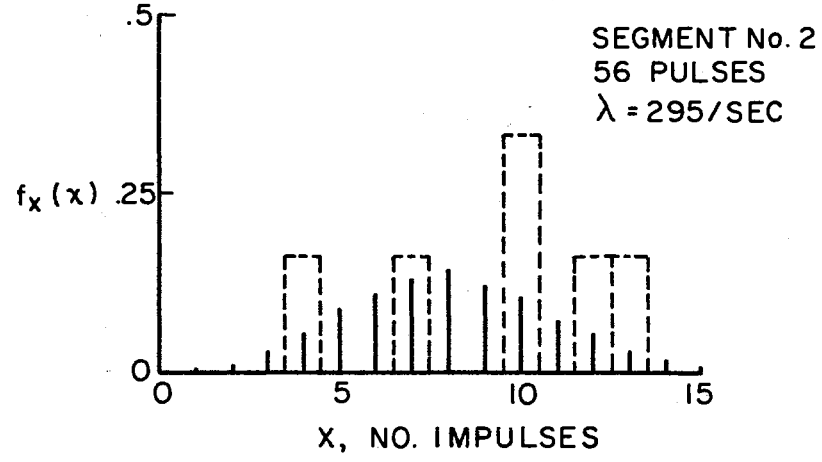
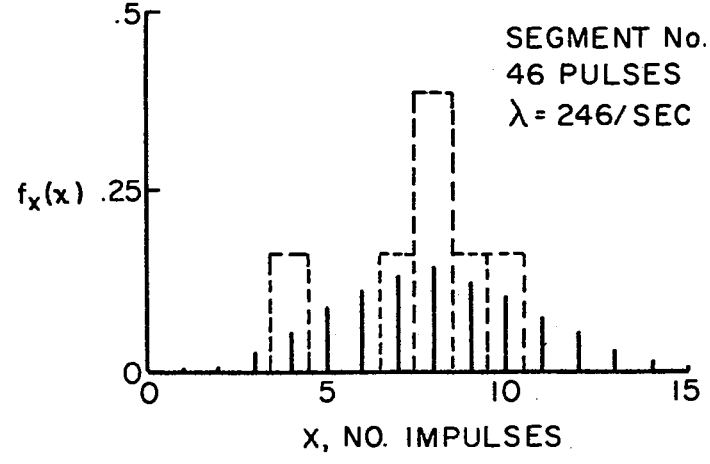


Figure 23. Distribution of Pulse Widths for Record #3







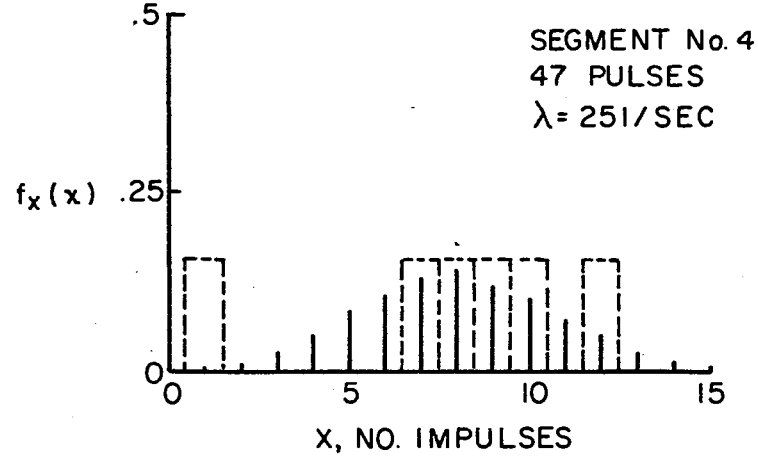


Figure 24. Distribution of Pulses for Record #3

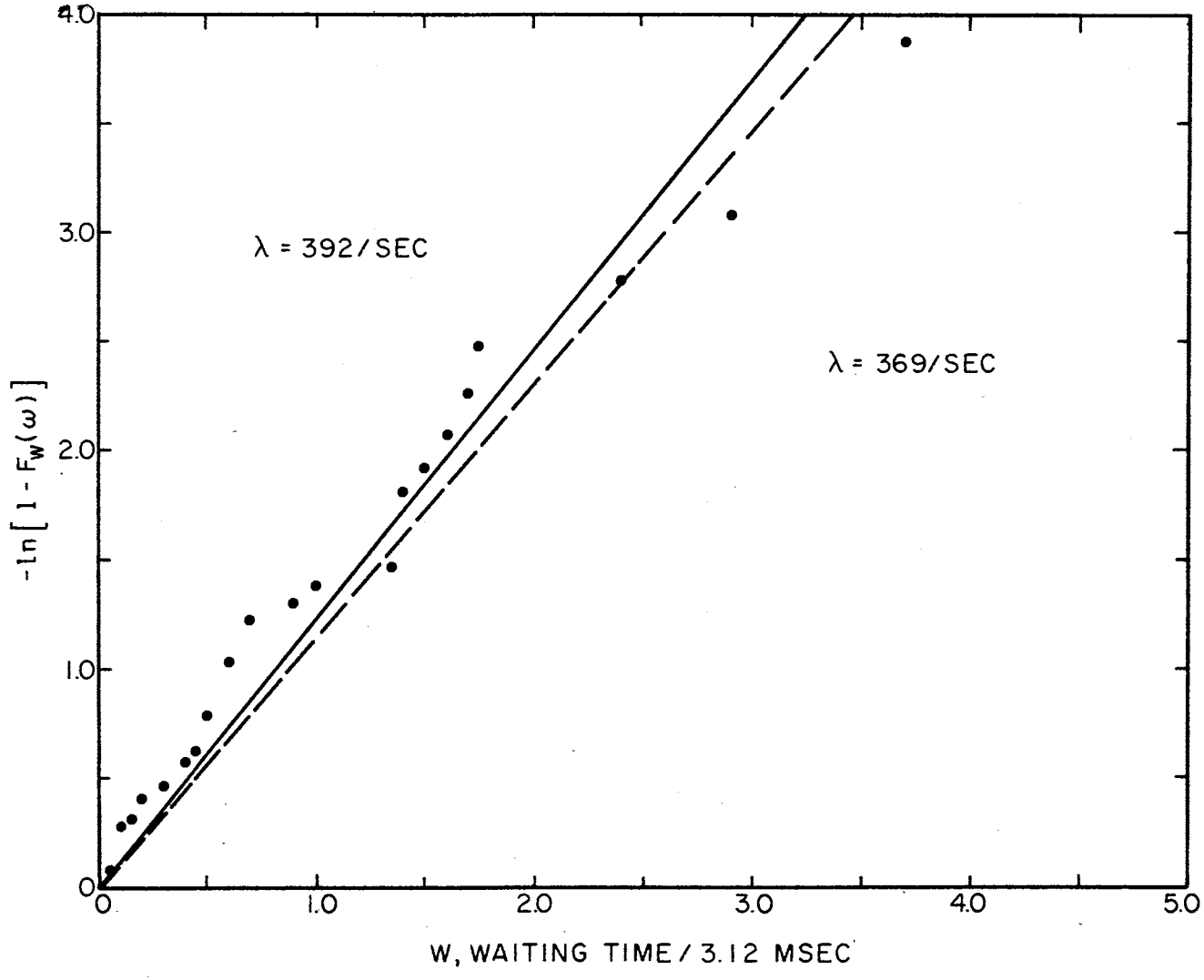


Figure 25. Distribution of Waiting Times, Segment 1, Record #3

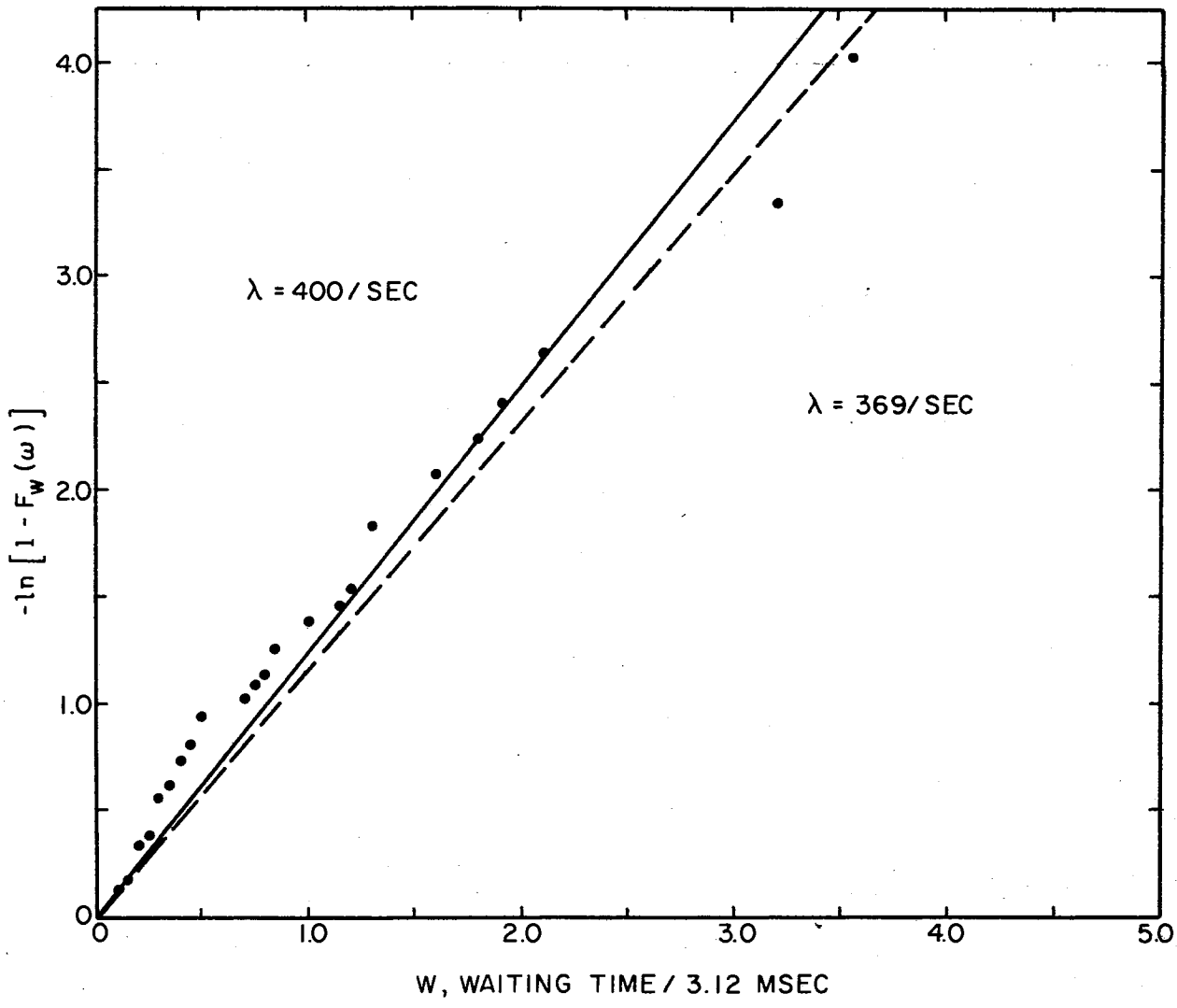


Figure 26. Distribution of Waiting Times, Segment 2, Record #3

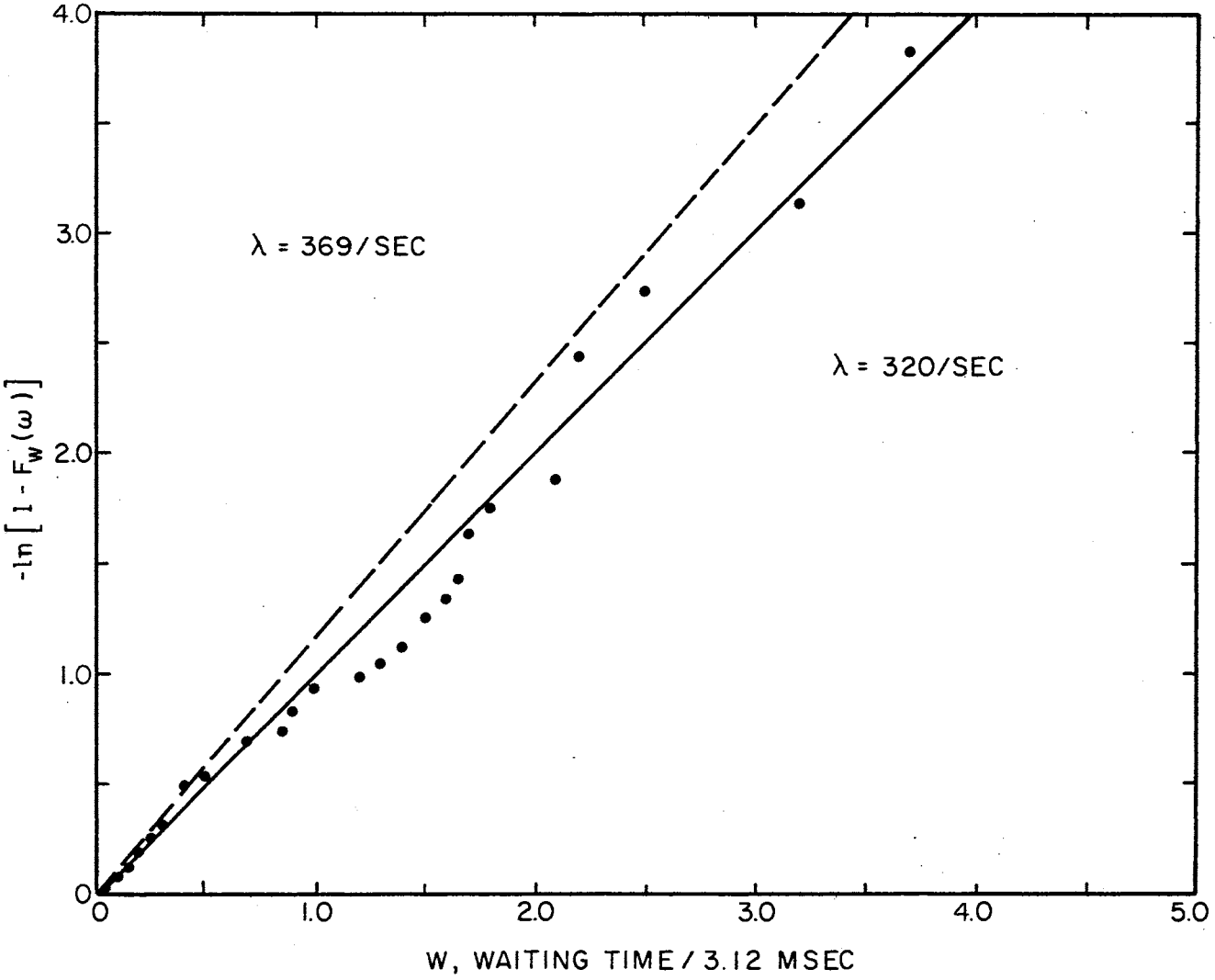


Figure 27. Distribution of Waiting Times, Segment 3, Record #3

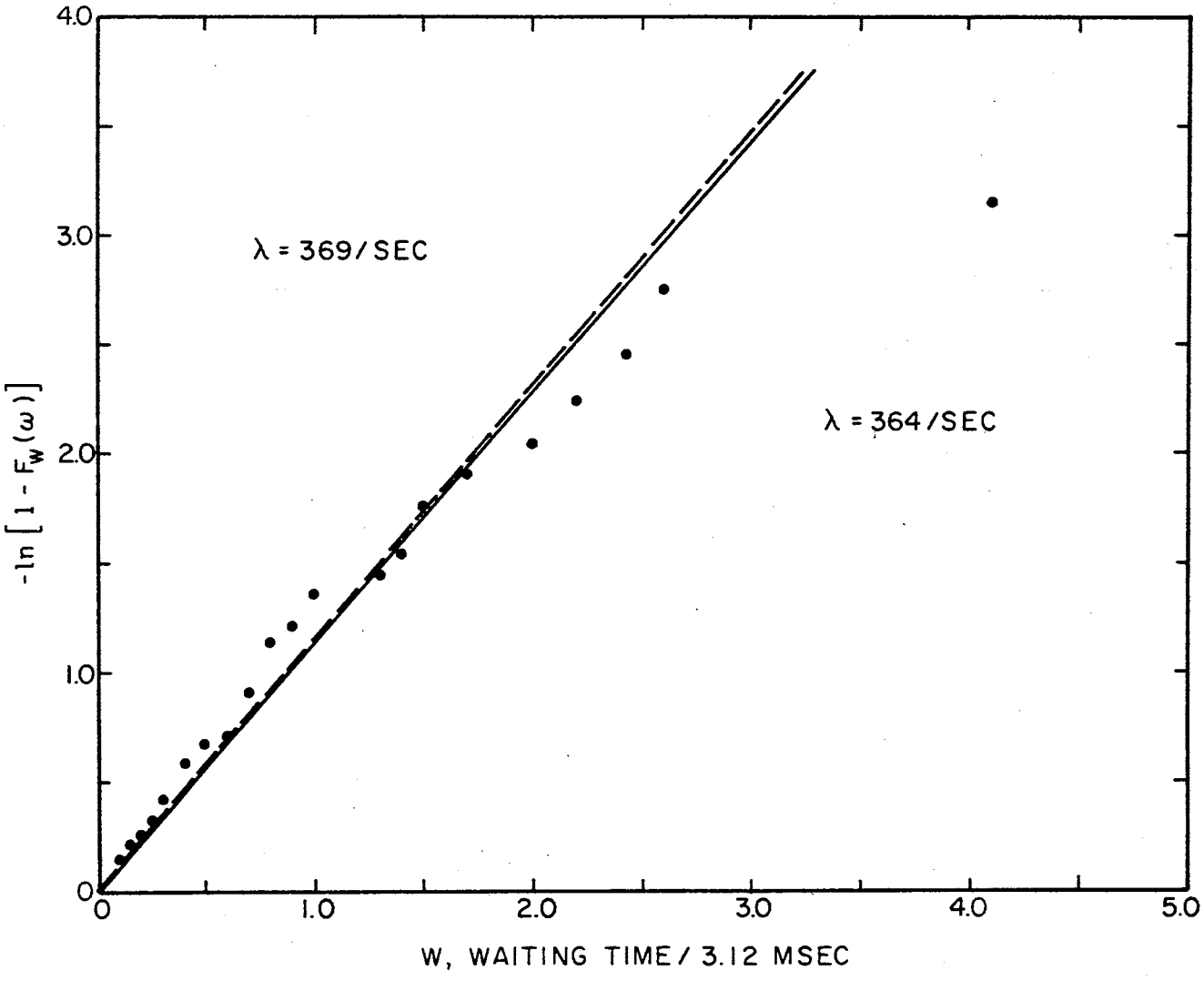


Figure 28. Distribution of Waiting Times, Segment 4, Record #3

graphs provide a credible fit to the data. It is observed that there exists a slight reduction in the slope of the data suggested above 6.25 milliseconds as compared with shorter waiting times. This effect is particularly predominate on the data shown in Figures 25 and 26. It is believed that this effect is, again, due either to a nonuniform density of pulses or a statistical fluctuation of a small sample size. Because of the relatively close agreement of the slope of the straight lines to a median slope, within say $\pm$ 10%, there again appears to be a fairly credible case for use of a stationary model. Reviewing the data presented in a composite sense once again, the following conclusions are drawn: (1) there appears to be adequate justification to model the triggering mechanism by a Poisson distribution, and (2) there seems to be a reasonable case for an approximation of uniform density, or constant λ, throughout the record. Hence, it is concluded that the record is amenable to analysis using stationary methods.

4.4 Summary and Conclusions of the Investigation of Stationarity. To the author's knowledge there have been no studies of this type made on thunderstorm data. There have been studies conducted with regard to the waiting time between lightning flahses, and there have been investigations of the distribution of waiting time between short term changes. However, the author knows of no reported attempt to determine whether or not the thunderstorm signature is stationary over short time intervals. One assumption which seems valid with regard to the short term changes and appears questionable with regard to the luminous flashes, is that of mutual independence among the waiting times.

It is concluded that this study provides reasonable cause to believe that the three segments are amenable to stationary analysis.

Hence, the following conclusions are drawn with regard to the study of stationarity: (1) a Poisson model for the sferic triggering mechanism seems reasonable, particularly if $\lambda = \lambda(t)$; (2) with the exception of segment four of Record #1, the data sets selected have constant λ's within $\pm$ 10%; and (3) the data can thus be analyzed with the assumption of wide sense stationarity without serious loss of credibility or accuracy.

CHAPTER V

RESULTS OF THE DATA ANALYSIS

5.1 Introduction. The data reduction theory developed in Appendix A was implemented in the form of the computer program of Appendix C in order to analyze segments of data gathered from two thunderstorms of fairly homogeneous meteorological content. Specifically, selected data segments were serially reduced to yield an estimate, in a empirical sense, of the time averaged second-order density function. From this density function and based on the plausible assumption of wide sense stationarity developed in Chapter IV, an estimate of the autocorrelation function was computed. Applying the Fourier transform, the power spectral density function was estimated. A preliminary data reduction study was undertaken in an effort to determine the length of data segments amenable to the method of analysis. As a result of the preliminary study, three data segments of 0.75 second duration were selected for detailed analysis. Two of the segments were taken from one thunderstorm and the third segment was taken from a storm occurring three days later.

5.2 Meteorological Description of the Thunderstorms Being Studied. Using the data gathering instrumentation referred to in Chapter II, the electromagnetic signature of several thunderstorms occurring in the Rapid City, South Dakota area were recorded during the Summer of 1969. A meteorological description of two representative thunderstorms, as reported by a meteorologist who was a member of the flight crew

gathering data, follows (20):

> A polar front passed through the area of the Black Hills, South Dakota on Sunday the 13th of July. That front then stalled and the mP air behind it became the air which spawned the hail and tornadoes during the subsequent week.
>
> On Monday, 14 July 1969, storm activity began shortly after noon local Daylight time. Hail occurred early in the life of the storm but sferics remained of low intensity during that time, increasing rapidly only as the cloud height built rapidly.
>
> The early stage of the storm cell was well defined and cloud growth was readily visible. However, after about an hour and a half, convection was taking place everywhere along a roughly north-south line and fresh cumulus buildups seemed to merge with the older storm cell. Sferic activity became quite widespread, although not of great intensity after two hours. By that time (about 2 p.m. MDT) the storm system had become quite widespread and was a general area thunderstorm rather than isolated cells. Cloud tops eventually reached 35,000 feet and higher in some cases (50,000 feet during hail at Ellsworth AFB).
>
> The storm of Thursday, July 17, occurred under conditions similar to those of the preceeding days. Cyclonic vorticity from the west came into the region at upper levels (10,000 to 20,000 feet) and triggered activity which used the surface heating and slightly negative stability to cause the storms. During this storm, however, a much more intense line of activity was evident west and north of the Black Hills region.
>
> Hail and funnels were reported during this storm. Cloud tops went above 35,000 feet and occasionally to 50,000 feet. To an observer, this storm looked very much like the storm of 14 July. However, the mature and dying phases of this storm were not as widespread as the earlier storm.

Various segments of storm data from these two thunderstorm records were subjected to computer data reduction, and the results of these preliminary trials are described in the next section.

5.3 Preliminary Data Reduction Study. The computer program described in Appendix C is actually a fourth generation program. It evolved as the investigation became more detailed. At the onset of the study three rather significant questions had to be answered concerning the handling of data. First of all, one had to determine how long a segment of data could be handled in a reasonable amount of computer

processor time. If this time were excessive, then an alternate investigative approach had to be sought. A second question, which would play an important part in resolving the question of computer processor time, was that of determining how wide the amplitude windows should be in computing the probability density function, where the amplitude windows are slicing widths of the data amplitudes. On the one hand, one would like windows as small as possible in order to get a "highly resolved" probability density function and, hence, reduce the bias in the estimate; however, each additional window imposes additional computer processor time and requires additional storage space for the probability density function. Thus, there was the problem of minimizing computational time, making most efficient use of computer storage, and making the amplitude windows as small as possible. A third question which imposed a significant impact on the two questions above was, how long a time lag must be considered in forming an estimate of the autocorrelation function? Clearly, for every increment in τ, there must exist a corresponding estimate of the second-order probability density function. Hence, the preliminary data reduction study was undertaken to optimize, in a cut-and-try investigative approach, the computer processor time, the probability density resolution and the resolution of the autocorrelation as a function of τ.

Several initial computational runs were made, using artificially created data, to certify the performance of the program. For reasons of economy of time and money, no large scale attempt was made at this point to resolve the problems mentioned earlier, in particular the processing time. After verifying the performance of the computer program, a digitized data tape was generated from the thunderstorm signature

recorded July 14 during a period of fairly intense sferic activity. The first 320,000 samples, 0.5 seconds of storm signature, were analyzed using the computer program. Arbitrarily, time lags of 12.5, 25, 50 and 100 milliseconds were chosen. Reflecting back on the choice of time lags, this was a fairly poor selection in view of the frequency response of the measurement system. Note that the lags correspond to frequencies of 80, 40, 20 and 10 Hz. Amplitude windows of 2.50 volts centered about 0, $\pm$ 2.5, $\pm$ 5.0, $\pm$ 7.5 and $\pm$ 10 volts were used. This voltage is measured at the A/D output after replay amplification. The minimum amplitude window width was dictated by the amplitude of the 18.75 KHz sinusoid added to the data. As shown in Appendix B, the effect of the 18.75 KHz can be minimized by making the amplitude window equal to the peak-to-peak amplitude of the sinusoid. For these 9 matrices characterizing the second-order probability density function. The most striking characteristic of these matrices was that 98% of the probability mass was concentrated at the origin or

$$P\{-1.25 \leq x(t) \leq 1.25, -1.25 \leq x(t + \tau) \leq 1.25\} = .98$$

for every τ. The second point of interest was that the program used 52 minutes of processor time. Hence, the following conclusions were drawn: (1) the poor spread in probability mass was apparently due to a relatively low intensity of sferic activity; and (2) if the processor time were directly proportional to the number of time lags, then to increase the sample size of τ, the number of amplitude windows may have to be reduced.

The second generation program was written at this point since several significant changes were suggested. In the first program, disc

storage was used to accommodate the anticipated lengthy time lags. Since the lags used appeared to be much too large, the program was modified to do all computation in the high speed memory, doing away with the access time on the disc and, hence, reducing the processor time involved. A new digitized data tape was generated from a period of extremely intense sferic activity. On the next series of runs the amplitude windows were left unchanged but the program was expanded to accomodate 8 samples of τ, these being 1.56, 3.12, 7.81, 15.62, 156.25, 781.25 μseconds and 1.56, 3.12 milliseconds. This run indicated a more desirable spread in the probability mass, typically 75 to 80% of the mass being concentrated at the origin. The data segment used for this trial was, again, 0.5 second in duration. The processor time used was 25 minutes. The following conclusions were drawn as a result of the second generation run: (1) the study is limited to data segments of 1 second or less due to processor time involved; (2) the data appear to become essentially uncorrelated somewhere between 156.25 and 781.25 μseconds; (3) data segments of extremely intense sferic activity should be used to obtain a reasonable spread in probability mass; and (4) it appeared that the amplitude data windows were reasonable and a further widening to reduce the number of windows did not seem necessary or desirable.

A third generation program was written at this point to conduct a comparative investigation of several data segments from both storms. Specifically, the first two trials involved only two segments of 0.5 seconds duration, and it was desired to investigate the behavior of the second-order statistics over representative samples from both storm records. Six data segments of 0.5 seconds to 1.5 seconds duration were

selected from each of the two storm records. These twelve data segments represented periods of extremely intense sferic activity. The program was modified to consider six data segments at a time, hence, two runs being necessary. Time lags of 0, 1.56, 3.12, 7.8, 15.6, 23.4, 31.25 and 39 μseconds were considered. Thus, eight 9 x 9 matrices had to be generated for each of the six data segments, and then the autocorrelation function computed. The processor time used in these runs were 1 hour and 23 minutes for the six segments of the first storm, and 1 hour and 47 minutes for those of the second storm. The results indicated that when the autocorrelation function was normalized to the average measured power, the functions were nearly identical in spite of considerable variance in the distribution of probability mass. The probability mass concentrated at the origin ranged from 50% to 78%. Because of the close agreement of the autocorrelation functions out to lags of 39 μseconds, it was hypothesized that this similarity existed for all time lags. To further investigate this, the autocorrelation function had to be computed for time lags out to the point that $x(t)$ and $x(t + \tau)$ became uncorrelated. It was concluded earlier that this occurred somewhere between 156.25 and 781.25 μseconds. One data segment of 0.75 seconds duration was arbitrarily selected and by cut-and-try application of the computer program, it was found that the process became essentially uncorrelated after 230 μseconds. Further, the autocorrelation could be characterized by 30 points corresponding to τ = 0, 7.8, ..., 227 μseconds in increments of 7.8 μseconds. Because of the close resemblance observed in the estimates among the twelve records studied, it appeared reasonable to refine the investigation to a representative sample of three. Hence, it was decided to investigate three segments

of 0.75 second duration, two segments from the first storm and the third from the second storm. At this point a fourth generation computer program resulted from the modification of the existing program. The new program considered data segments of 0.75 second duration and computed probability densities for 30 increments of τ. This resulted in thirty 9 x 9 matrices, 30 sample points of the autocorrelation function and a power spectral density out to 64 KHz in increments of 1 KHz. Processor time involved was 1 hour and 33 minutes for each data segment.

Summarizing the results of the preliminary data reduction study, four important problems were solved. The first three problems were those mentioned at the beginning of this section, namely, (1) how long a segment of data could be handled? It was encouraging to find this result, 0.75 second, consistent with the theory of Appendix A; (2) how wide should the amplitude windows be? The encouraging result here was that the compromise was no greater than it turned out to be. Specifically, it would have been desirable to estimate a highly resolved probability density function, but from the standpoint of processor time and the interfering sinusoid some resolution had to be sacrificed; and (3) how long a time lag must be considered? The relatively short time lag was a fortunate result because, now, all computation could be done in high speed memory without having to use the disc and incurring high access times. The fourth problem solved was that of focusing the scope of the investigation. Specifically, the second-order statistics of three data segments appeared to be representative of those studied earlier, and it was determined to conduct a detailed investigation of these three.

5.4 Analog Content of Three Data Segments Selected for Study.

The first two data segments, Record #1 and Record #2 were taken from approximately the mid-point of the magnetic tape recording of the storm on July 14. These segments were separated in time by approximately five minutes. The motivation for the selection of these two records was first, that they were both 0.75 seconds in duration, and secondly they would indicate variations in the statistics on one member function. It was conceivable that while each segment, separately, might appear to be wide sense stationary, they may vary considerably from segment to segment. The third segment, Record #3, was taken from the final third of the recording of the storm occurring July 17. The description of the analog content, which follows, is a qualitative discussion as opposed to the quantitative study made in Chapter IV.

Because of its representative nature, the entire signature of Record #2 is shown in Figures 29, 30, 31 and 32. The data are displayed with increasing time proceeding from left to right and continuously from the top line to the bottom line of each figure. The time scale is 5.20 milliseconds per inch. Assessing the sferic activity of this data segment, there appear to be two sustained strokes, i.e. greater than 30 milliseconds in duration, occurring on lines 1, 2 and 3 of Figure 29, one sustained stroke on lines 1 and 2 of Figure 30, and a final sustained stroke on lines 1 and 2 of Figure 31. Scattered throughout this record there appear numerous short bursts of activity, typically lasting 1 millisecond or less. It is observed that saturation effects in the equipment are exhibited in the data, particularly in line 3 of Figure 29. Because this effect occurs relatively infrequently, it would appear to have been inadvisable to reduce the sensitivity of the measurement

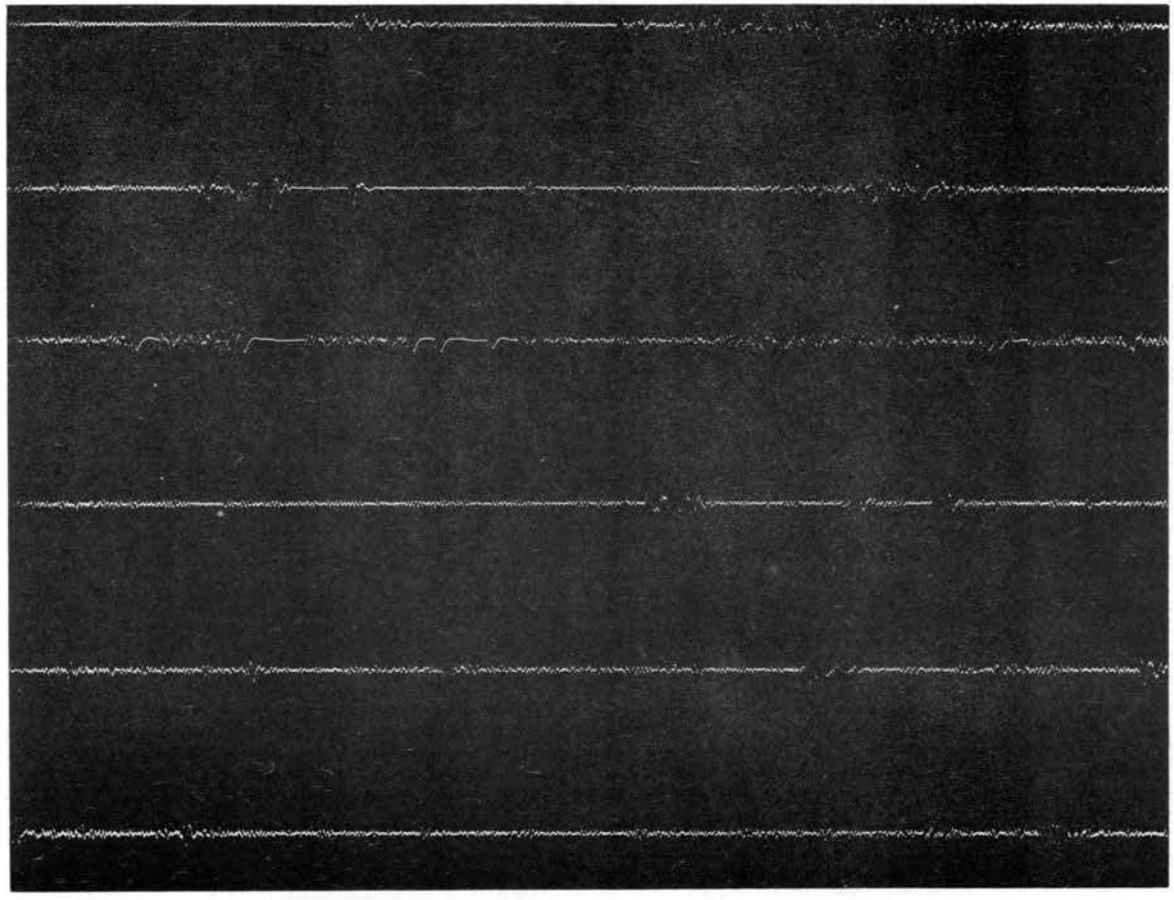


Figure 29. Analog Signature of Segment 1, Record #2

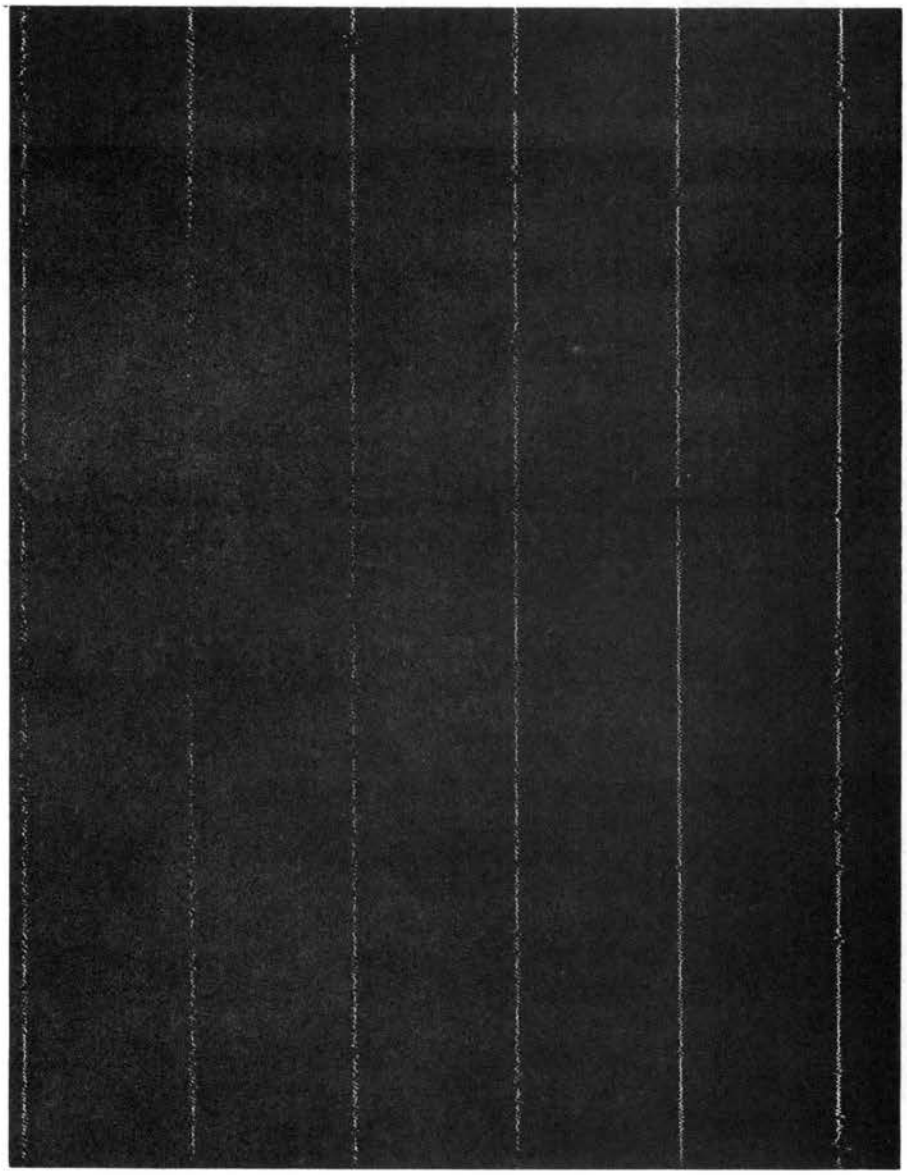


Figure 30. Analog Signature of Segment 2, Record #2

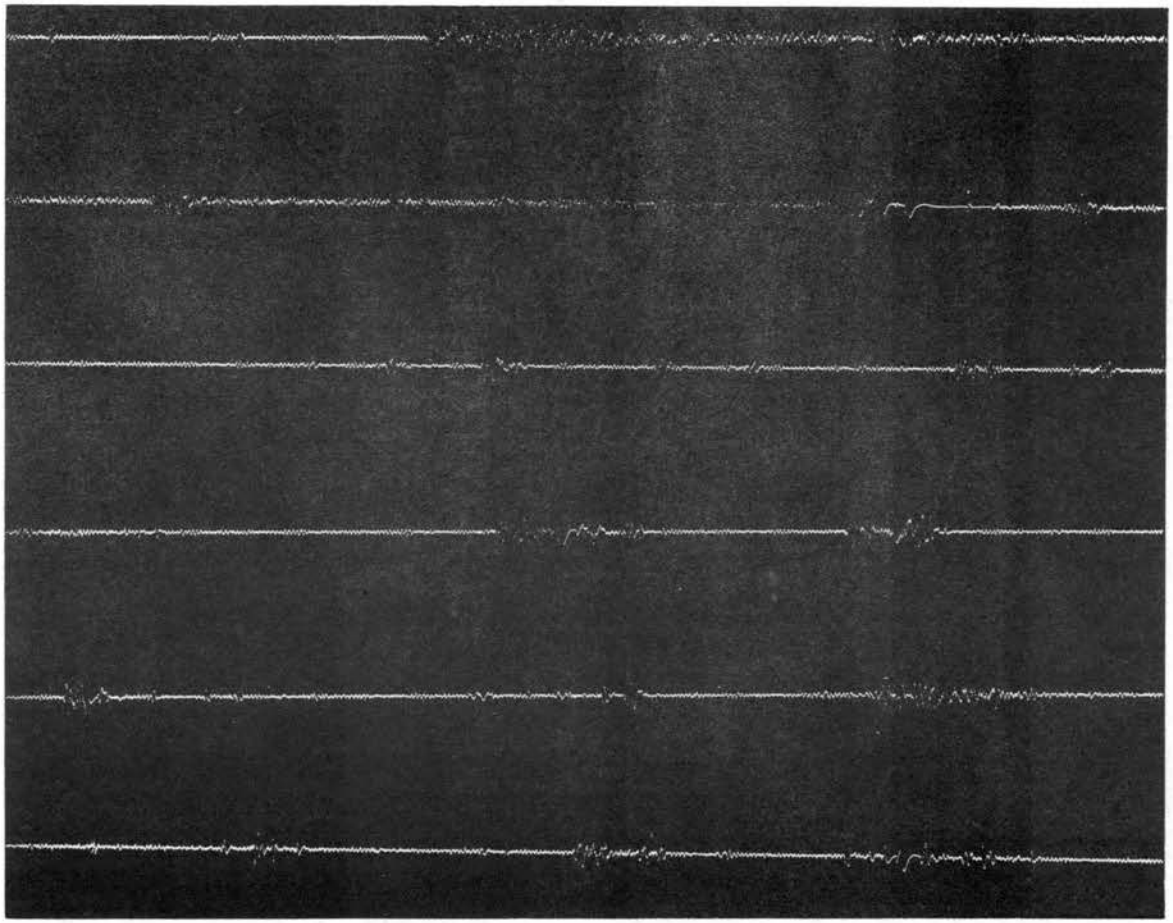


Figure 31. Analog Signature of Segment 3, Record #2

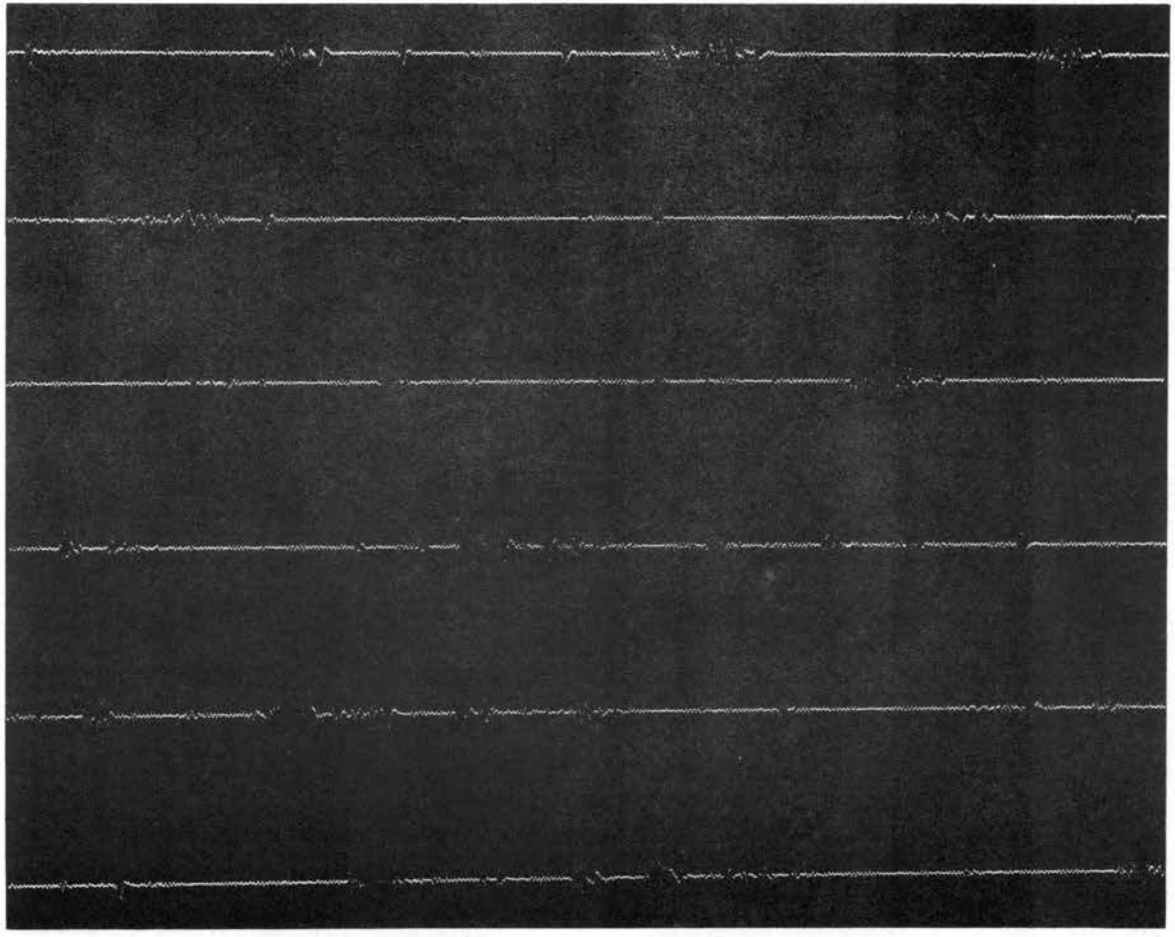


Figure 32. Analog Signature of Segment 4, Record #2

to reduce the incidence of saturation. It may seem an apparent paradox that lengthy strokes are being discussed from the same data that in Chapter IV evidenced no pulse widths greater than 16 milliseconds. This point is resolved by the fact that, in the quantitative study, the continuous strokes appeared to evidence surges which were separable into a succession of pulses. A case in point is line 3 of Figure 29. This line, although evidencing more or less continuous sferic activity, was nevertheless separated into 8 pulses.

The sferic activity of Record #1 consisted entirely of relatively short bursts and exhibited no sustained strokes. Figure 32 is representative of Record #1 in its entireity. The sferic activity of Record #3 exhibited two sustained strokes of approximately 45 milliseconds duration. The conspicuous characteristic of this record was the high incidence of extremely short bursts as compared with the two previous records. This is evidenced in the study of Chapter IV where a preponderance of burst durations were 300 μseconds or less.

A comparison of these analog data with the electromagnetic signature recording reported by Arnold and Pierce (2), Malan (16), Kitagawa and Kobayski (15) and Uman (24), indicates two immediately apparent differences. First, a characteristic peculiar to the data gathering equipment is the modulated sinusoid appearing during quiet periods. This signal originates in the aircraft electrical system and is coupled through the antenna into the measurement system. The carrier frequency of this signal has been found to be 18.75 KHz. A second difference is in the saturation effects of the measurement system. In the data gathered by Arnold and Pierce (2), there is a considerable positive excursion of the recorded field upon the occurrence of a return stroke,

as compared to the field changes associated with the leader and junction streamer processes. It appears likely that the return strokes were the cause of the occasional equipment saturation. Because of the relatively low incidence of return strokes and the relatively high incidence of lower amplitude rapid changes, it is believed that the elimination of near return strokes is not catastrophic.

5.5 Detailed Investigation of Three Segments of Storm Data. It seems appropriate at this point to discuss in general the development of this section and to relate some of the considerations given in this particular investigative approach. First, each record will be discussed separately in the context of its spread in "probability mass" and its autocorrelation and power spectral density functions. Then, the autocorrelation and power spectral density functions will be averaged to form an estimate of the time averaged autocorrelation and power spectral density functions of the process.

Obviously, there is no intention of offering the time averaged density function as a characterization of the storm process in an ensemble sense. Why, then, was this estimate computed in the first place? Essentially for two reasons, (1) at the onset of the study, the outcome was fairly uncertain, and because the second-order density function contains more statistical data than the autocorrelation function, it was computed to give the scope of the investigation greater breadth; and (2) although there exists the quantitative study of sferic activity in Chapter IV and the visual assessment in the preceeding section, the computation of the empirical second-order density function, in a serial sense, enables one to make a comparative study of the distribution of sferic amplitudes. As shown in Appendix A, it is possible that two

segments of data have identical autocorrelation functions but exhibit different sferic contents. Thus, the results of the time averaged density function estimate should be interpreted as a description of the particular data segment and not necessarily as a description of the process in an ensemble sense. A histogram of a typical marginal density estimate is shown in Figure 33. It is noted that even for these segments exhibiting extremely intense sferic activity, 70 to 80 percent of the probability mass is concentrated at the origin. Shown in Figure 34 are contours of constant probability. The autocorrelation function indicates the rate at which the ellipse approaches a circle, but it tells nothing at all about the comparative shapes. These representative plots are presented for reference in the later discussions of the individual records. The study of Chapter IV indicated that an assumption of wide sense stationarity was admissable on each of these three records separately, but it should be pointed out that this does not imply that the assumption holds in an ensemble sense. It is possible that each of these records could be used to define distinct processes which have entirely different autocorrelation functions. The results of the preliminary study indicate that this is not the case; however, this is the question to be answered in this portion of the study. Specifically, do the time averaged autocorrelation functions, estimated by time averaging over member functions, maintain close similarity over the range of τ under consideration?

Results of the data reduction of Record #1 indicated that 77.83% of the probability mass was concentrated at the origin for $\tau = 0$ and this value diminished to 63.05% for $\tau = 227$ μseconds. For comparison with the probability mass spread of later records, it was noted that

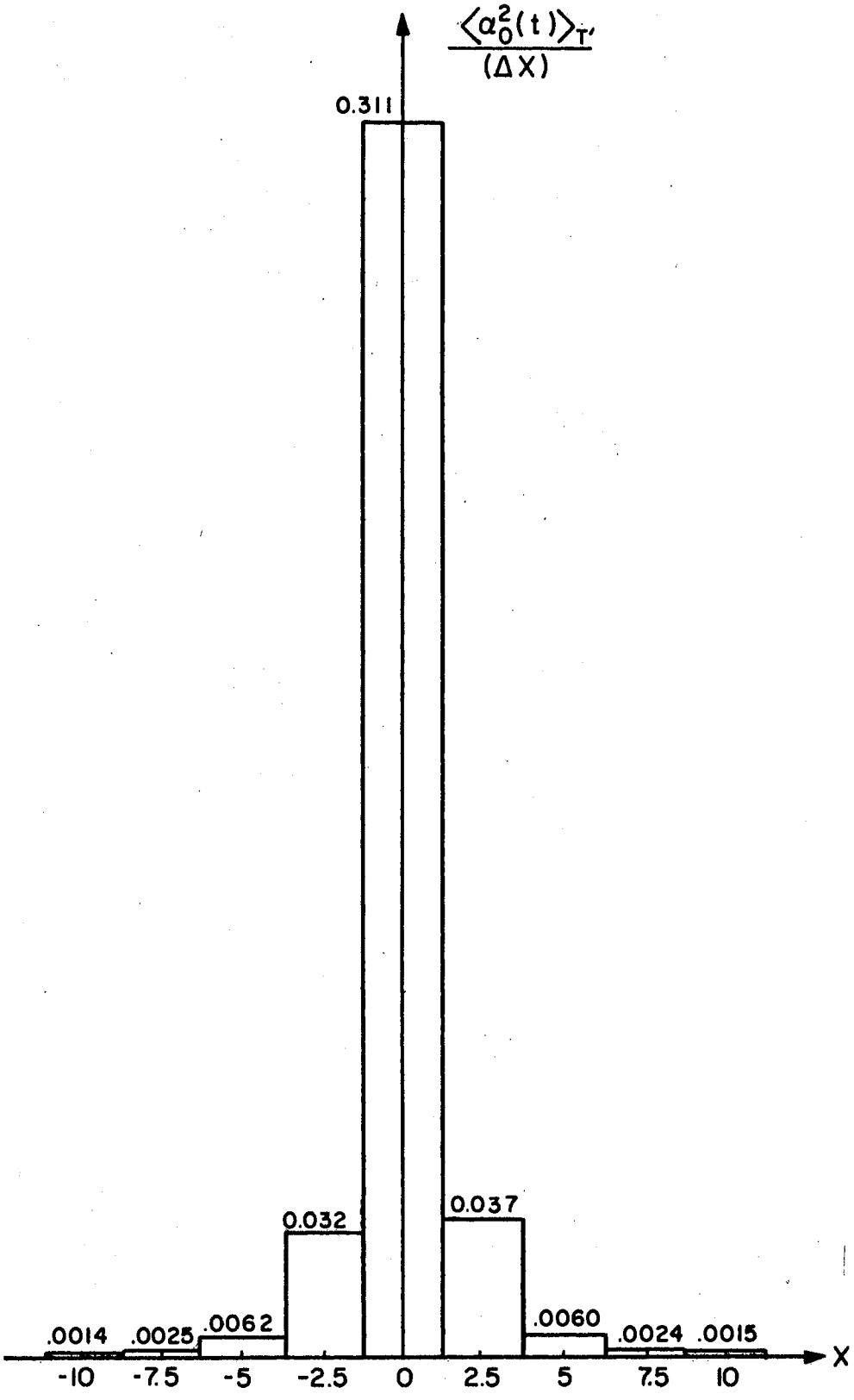


Figure 33. Histogram of Typical Marginal Density

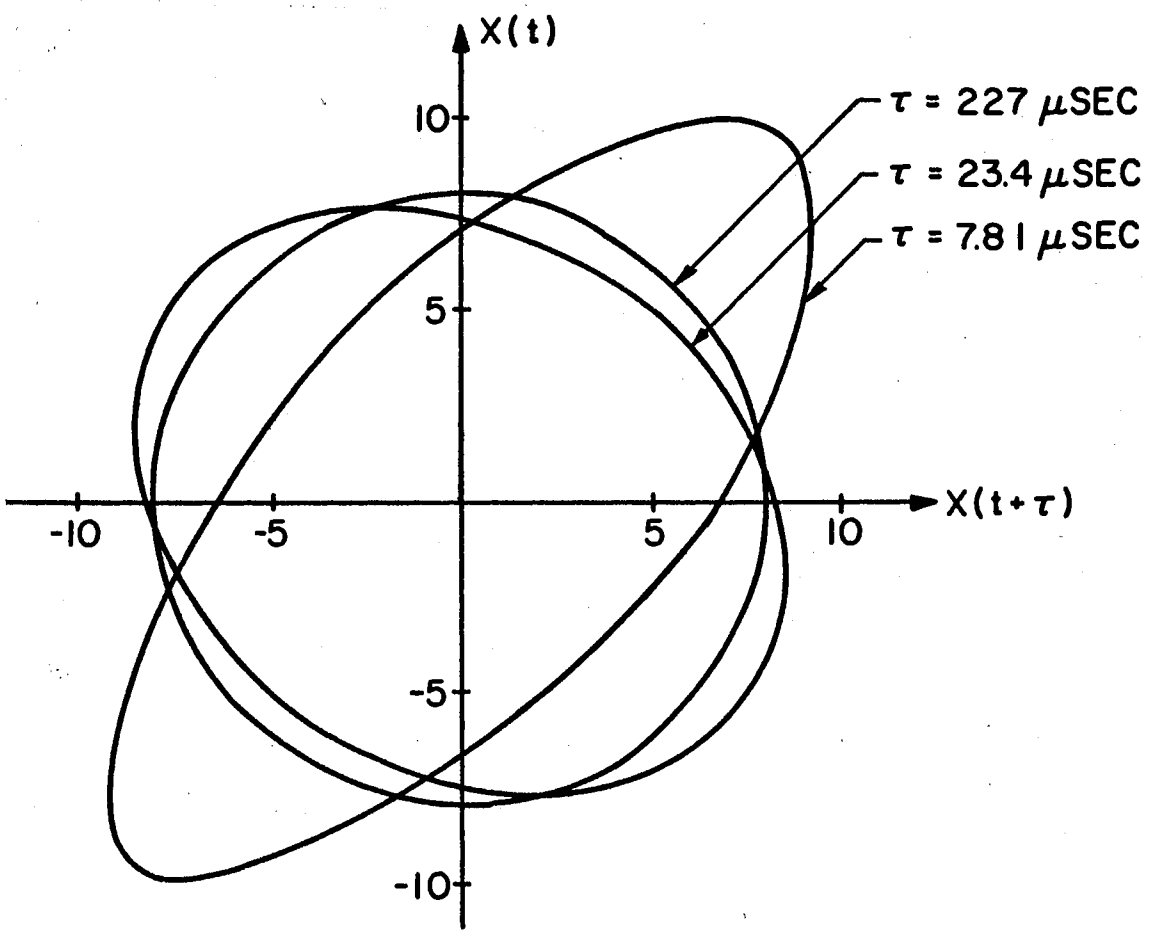


Figure 34. Contours of Constant Probability

$$< P\{x(t) = 0,\ x(t + 23.4\ \mu\text{seconds}) = 8.94\} > = .001$$

with 66.84% of the mass at the origin. Little more can be added at this point about the spread in probability mass, but in the discussion of the results of the remaining two records, some remarks can be made concerning the sferic activity of this record in comparison with the activity of those records. The autocorrelation function of Record #1 is shown in Figure 35 and the corresponding power spectral density is shown in Figure 36. The curves have been normalized to the average received power of the data segment in order to provide a basis for comparison among the three records. The autocorrelation function indicates that the data are essentially uncorrelated for time lags greater than approximately 230 μseconds. Effectively, this implies that the lower frequency limit of the measured data is 4.35 KHz, and in view of the calibration curve presented in Chapter II, it seems reasonably certain that this effect is attributable to the measurement system frequency response as opposed to the actual radiated spectrum of the thunderstorm. The power spectral density function indicates the existence of frequency components as low as zero. This is apparantly due to the approximation of the autocorrelation curve by its sample values. Specifically, the power spectrum shows a zero component whose amplitude is $.3108 \times 10^{-5}$, and the area under the autocorrelation curve is $.2980 \times 10^{-5}$. Hence, this computation evidences the existence of a zero frequency density in the power spectrum and also serves as a cross check on the result obtained by Fourier transforming the sampled autocorrelation function. A second computation which seems worthwhile is that of obtaining the area under the power spectral density function. This was done, and the result indicated that this curve, out to 64 KHz accounts for 100% of the

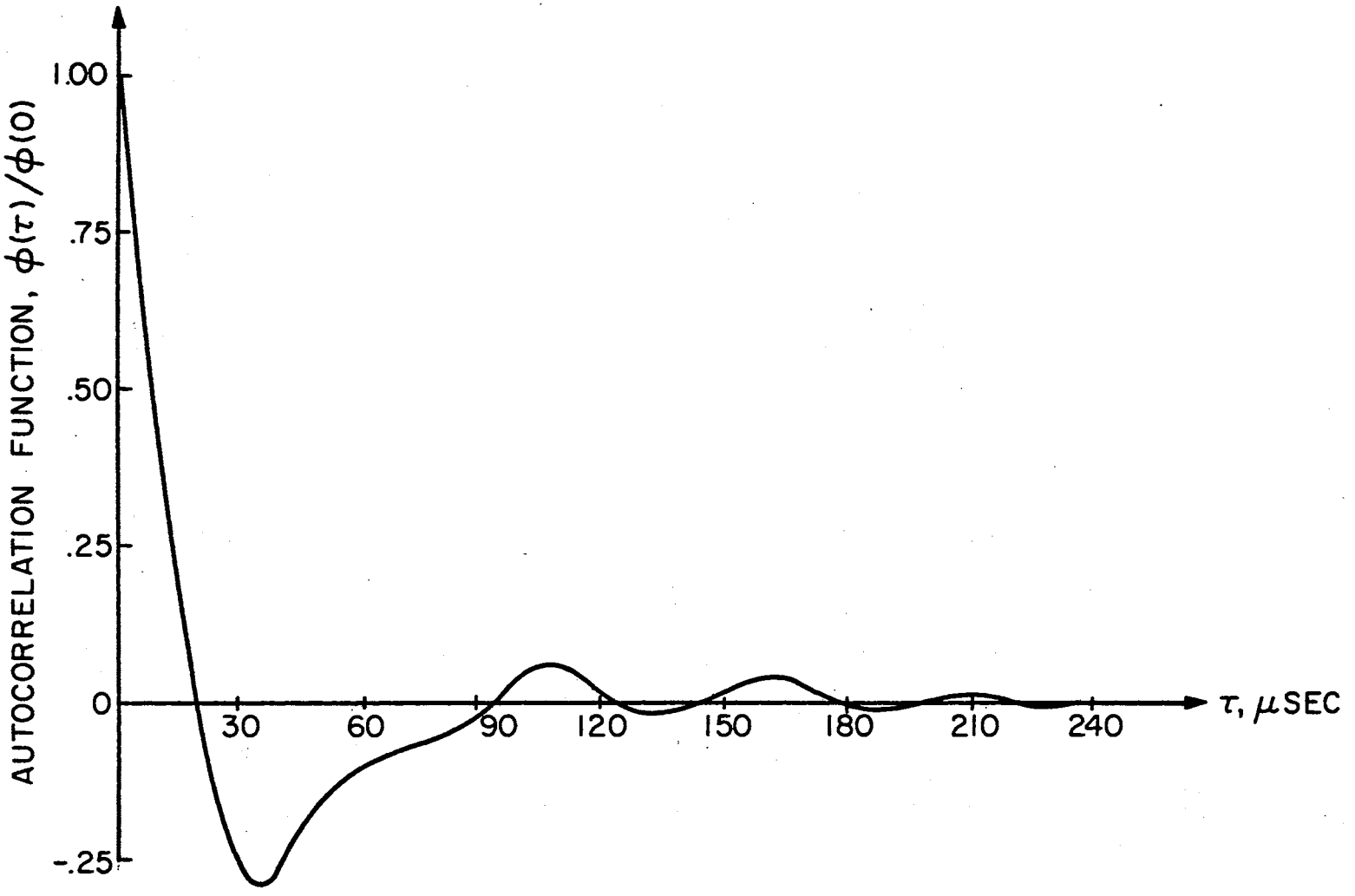


Figure 35. Time Averaged Autocorrelation Estimate From Record #1

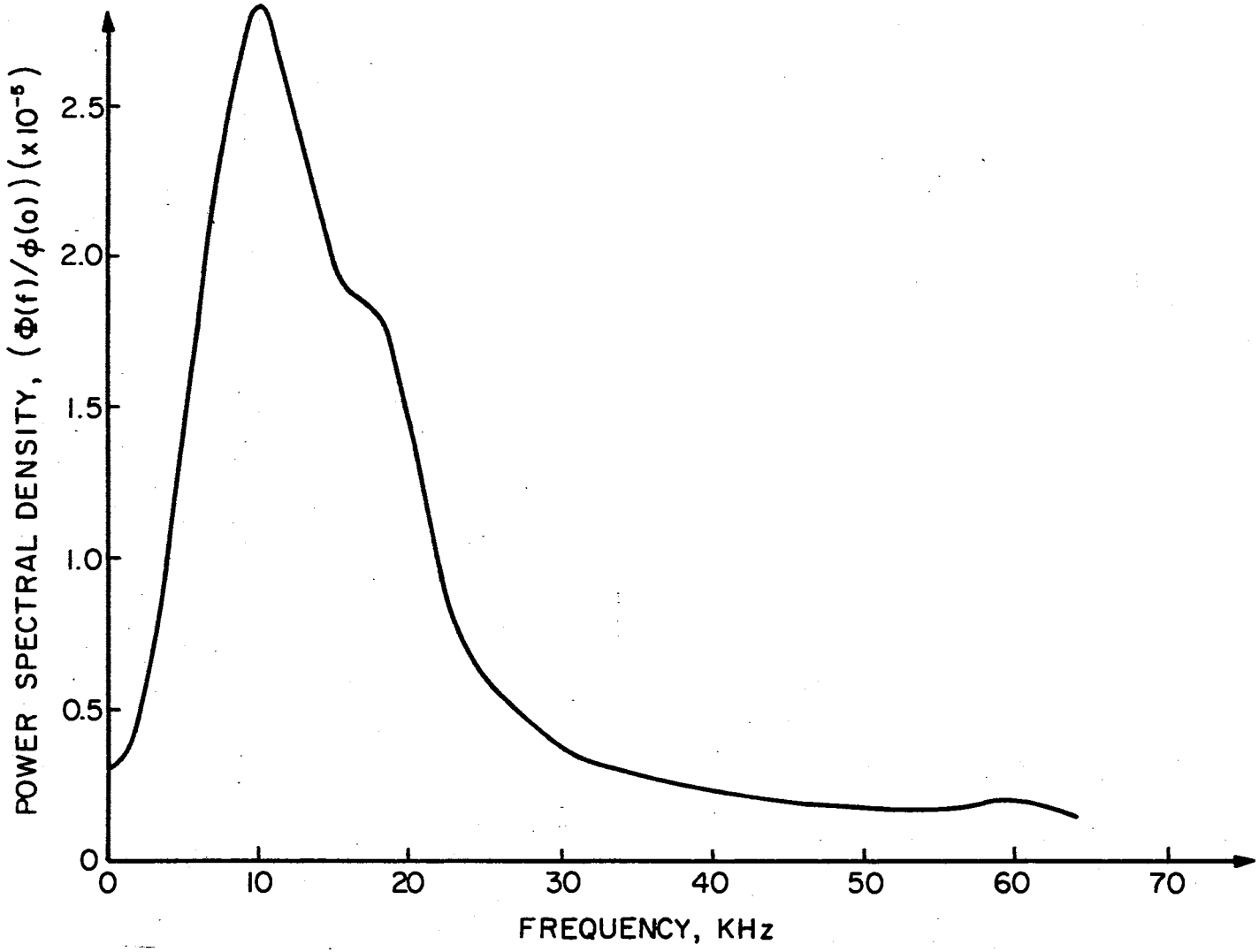


Figure 36. Time Averaged Power Spectral Density Estimate From Record #1

average power of the process, as predicted by the estimated autocorrelation function. This number at first appears incredible, but by recalling that the frequencies above 64 KHz are folded back into the spectrum (aliasing), this answer is a reassuring cross check of the computed spectral amplitudes. Because of the rapid fall off of the power spectral density, it is believed that the aliasing effect is reasonably bounded with acceptable error. To reduce the aliasing error, one must sample $\phi(\tau)$ at a higher rate which would involve correspondingly more lengthy processor times. Owing to the deteriorated high frequency response of the measurement system, it was concluded that the added cost would not be a worthwhile expenditure. Viewing the curve of the power spectral density, it is observed that a peak occurs at 10 KHz and the interfering 18.75 KHz appears as a bulge in the curve. There also occurs a slight rise in the curve at 60 KHz. The peak at 10 KHz is consistent with established results reported in the literature. The rise in the curve at 60 KHz is predicted by Taylor (23), however, he gives no account for it in a phenomenological sense.

The data reduction of Record #2 indicated that 74.82% of the probability mass was concentrated at the origin for $\tau = 0$ and this value diminished to 61.16% as τ increased. Corresponding to Record #1, it was observed that

$$< P\{x(t) = 0,\ x(t + 23.4\ \mu\text{seconds}) = 9.24\} > = .001$$

with 62.18% of the probability mass at the origin. Interpreting this in the context of the intensity of sferic activity, it appears that because of the decreased mass concentration at the origin with little apparent shift in the probability sample point, there were slightly

more sferics with moderate amplitudes in Record #2 than Record #1. This is in close agreement with the study of Chapter IV, since Poisson parameters of 331/second and 330/second were estimated for Records #1 and #2, respectively. The autocorrelation function of Record #2 is shown in Figure 37 and the corresponding power spectral density is shown in Figure 38. Again, the curves have been normalized to the average received power. The autocorrelation function indicates that the data are essentially uncorrelated for time lags greater than approximately 225 μseconds. It appears that the comments regarding the low frequency response of the measurement system are applicable here also. A visual comparison of the autocorrelation of Figures 35 and 37 reveal striking similarity. The power spectral density function indicates a zero frequency component of $.3775 \times 10^{-5}$ while the area under the autocorrelation function is $.4053 \times 10^{-5}$. Again, there exists reasonably good agreement between the two computational methods. The area under the power spectral density curve was computed and found to be unity, again verifying the transformed values. The amplitude of the power spectral density exhibits a peak at approximately 9 KHz, a bulge around 18 KHz and a hump at 34.5 KHz. The peak value at 9 KHz and the bulge at 18 KHz was expected, but the existence of the hump at 34.5 KHz was a surprising result. Since it did not appear in the power spectral density function of Record #1, it would seem that its presence could be attributed to the occurrence of the sustained strokes found in Record #2 and which were not found in Record #1.

A review of the data reduction of Record #3 reveals that 83.62% of the probability mass was concentrated at the origin for $\tau = 0$ and this value decreased to 71.79% for $\tau = 227$ seconds. Corresponding to the

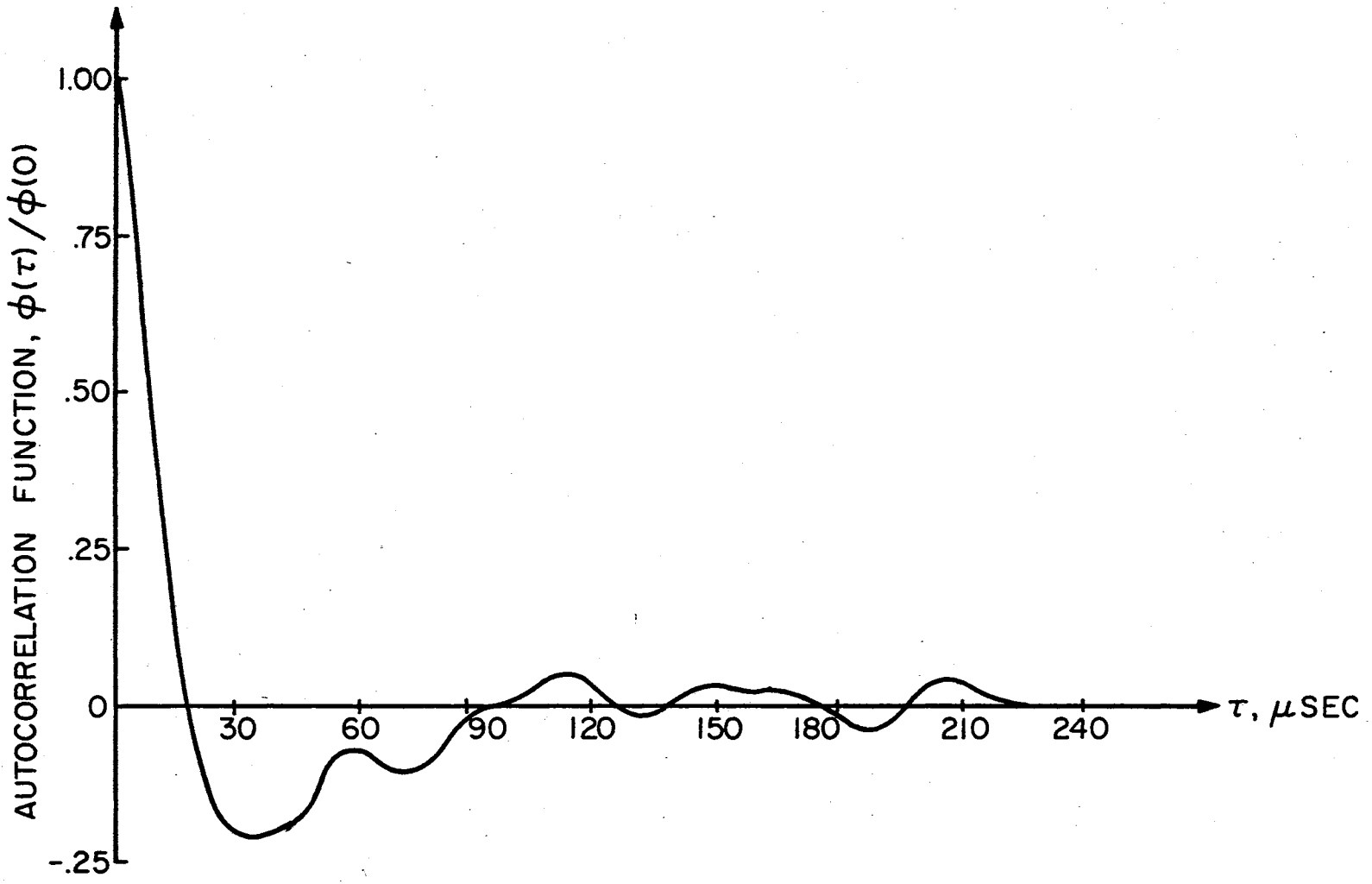


Figure 37. Time Averaged Autocorrelation Estimate From Record #2

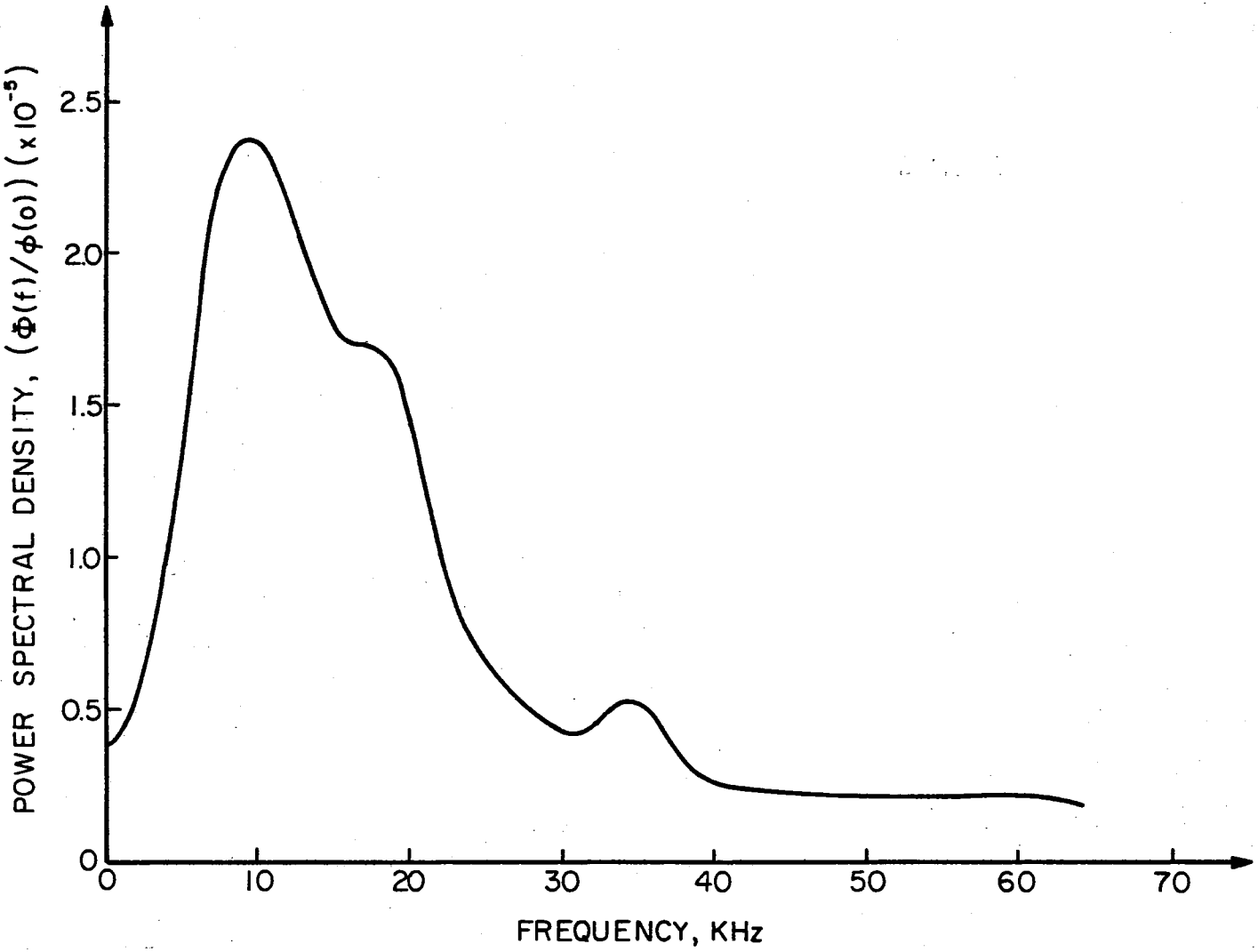


Figure 38. Time Averaged Power Spectral Density Estimate From Record #2

previous records, it was observed that

$$< P\{x(t) = 0,\ x(t + 23.4\ \mu\text{seconds}) = 6.99\} > = .001$$

with 73.69% of the probability mass at the origin. In order to interpret this spread in mass properly, it must be added that the output gain of the tape replay unit was lower for this record than for the two preceeding records. This was necessary becasue of the increased amplitude of several sfereics in the data. Because of the relatively poor spread in probability mass, it appears that this event did not occur with a relative frequency to warrant the reduction in gain. Comparins the sferic activity of this record with that of the first two records, it is recalled from Chapter IV that a Poisson parameter of 369/second was predicted. Since, this is in fairly close agreement with the parameters predicted for Records #1 and #2 and since the autocorrelation functions, estimated for τ = 23.4 μseconds for Records #1, #2 and #3 are -.45216, -.49229 and -.16669, respectively, one would expect that if the Poisson model holds in an ensemble sense, the sferic intensity has been reduced by approximately the square root of three (reference pages 113 and 114, Appendix A). It should be pointed out that this reasoning is valid even though the replay gains for the records were different. If the contour of constant probability is changed from .001 to .0003, then the probability sample point should be in fairly close agreement with the first two records. Computation reveals that

$$< P\{x(t) = 0,\ x(t + 23.4\ \mu\text{seconds}) = 9.03\} > = .0003 \quad .$$

This result then supports the belief that the Poisson model holds in an ensemble sense. Hence, combining these facts, then, leads to the

conclusion that the sferic activity was similar to Records #1 and #2, but that the replay gain was set too low. The autocorrelation function of Record #3 is shown in Figure 39 and the corresponding power spectral density is shown in Figure 40. These curves have also been normalized to the average received power. The autocorrelation function indicates that the data are approximately uncorrelated for time lags greater than 230 μseconds. A visual comparison of the autocorrelation function of Figures 35, 37 and 39, again, reveal striking similarity. The similarity of Figures 35 and 37, since they are both from the same member function of the ensemble, suggest that a sample average might be a credible description for one storm. However, the similarity of the third autocorrelation function, from a different member of the ensemble, suggests the possibility that a sample average of the time averaged autocorrelation and power spectral density functions could be computed from this sample of size three with some degree of credibility in an ensemble sense. Turning attention to the power spectral density function, it is observed that a zero frequency component of $.2941 \times 10^{-5}$ exists as compared to the area of $.3298 \times 10^{-5}$ under the autocorrelation function. Again, the area under the power spectral density curve was computed and found to be unity. Inspection of the power spectral density curve reveals a peak at 8 KHz, a pronounced 18.75 KHz component and a hump at 34.5 KHz. This record evidenced the sustained stroke activity found in Record #2 and also exhibits the 34.5 KHz component which would tend to reenforce the earlier discussion accounting for its presence.

Averaging the ordinates of Figures 35, 37 and 39, there results the estimate $\overline{\phi(\tau)/\phi(0)}$ of $\langle R(\tau)/R(0)\rangle_T$ shown in Figure 41. The sample

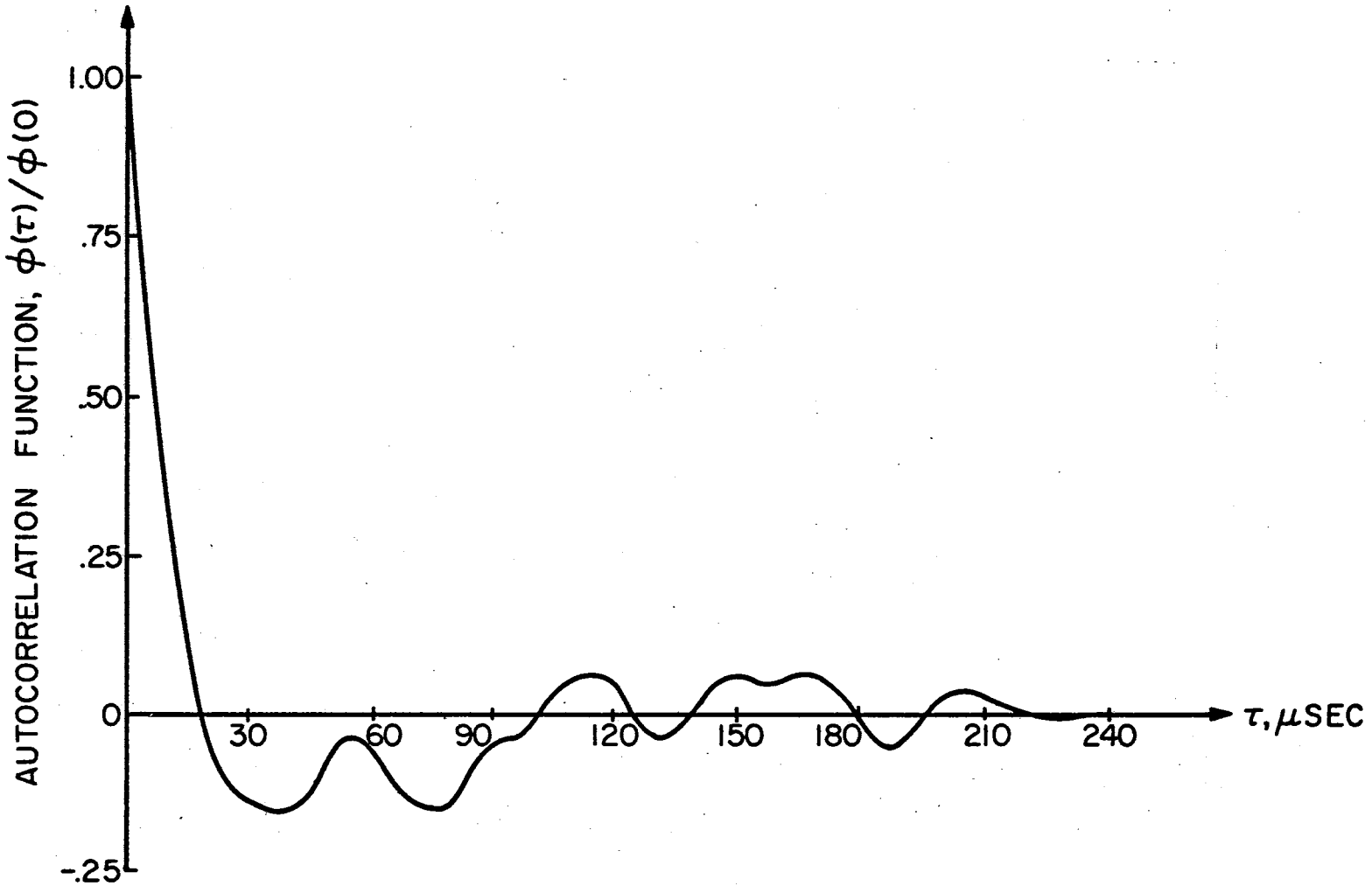


Figure 39. Time Averaged Autocorrelation Estimate From Record #3

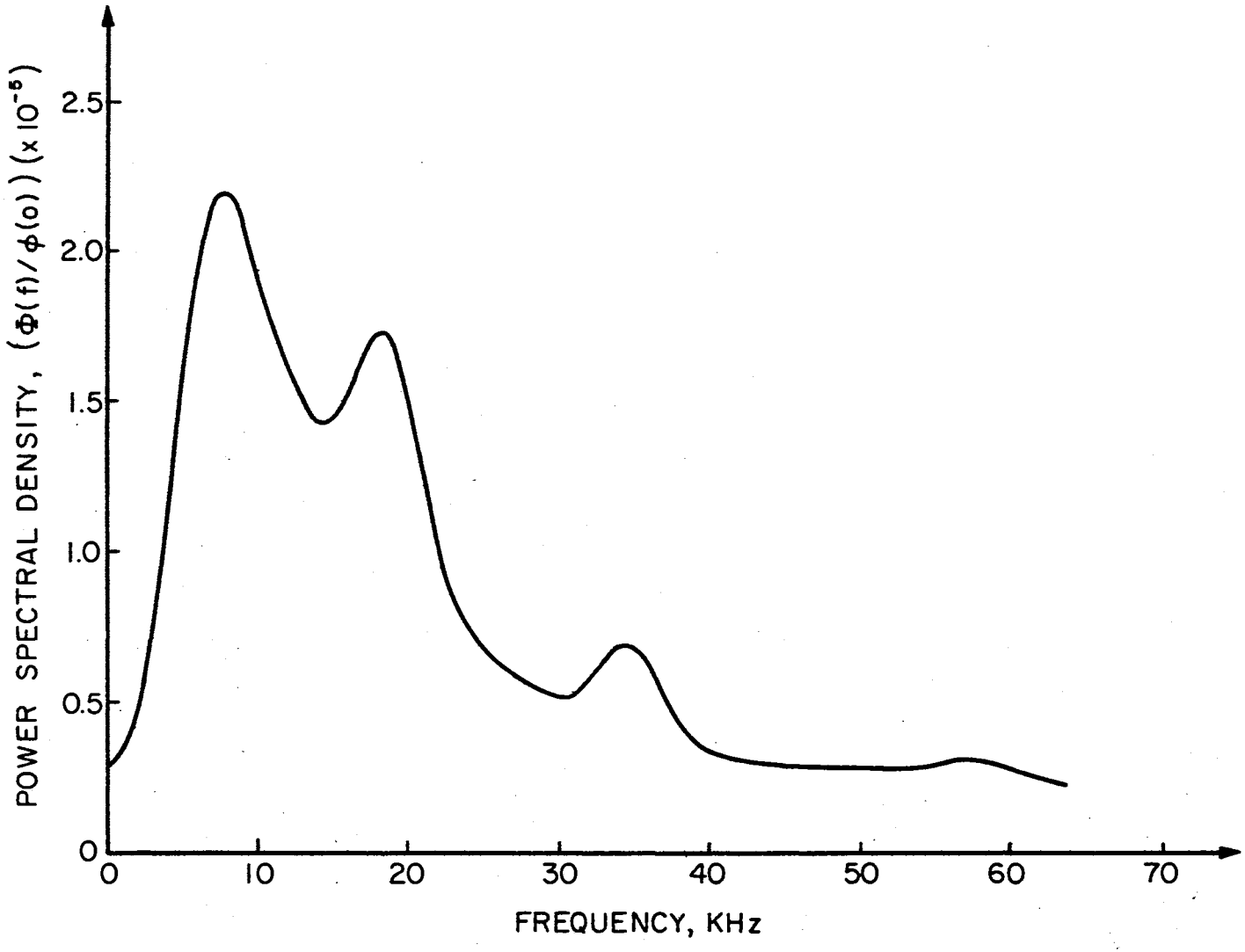


Figure 40. Time Averaged Power Spectral Density Estimate From Record #3

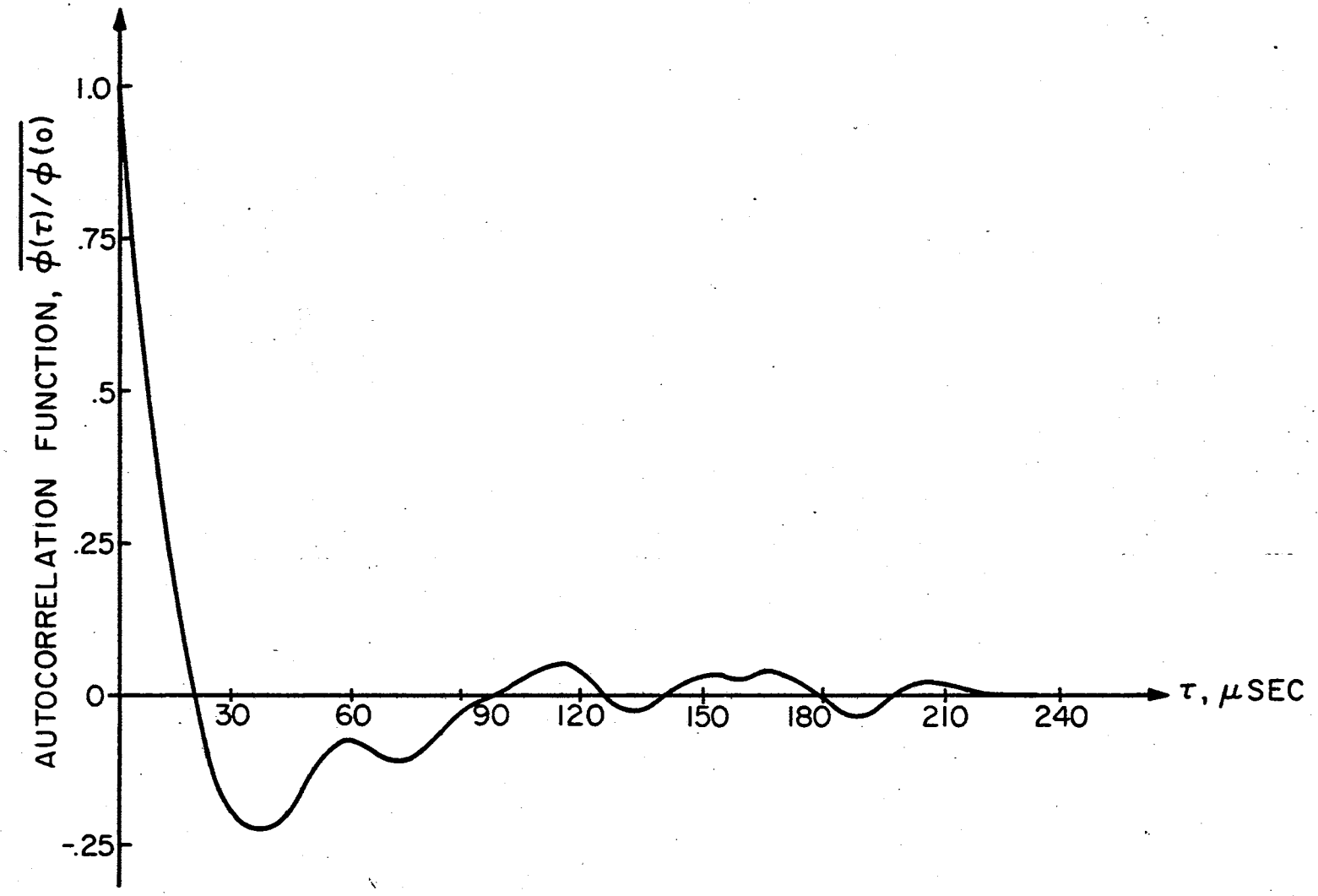


Figure 41. Averaged Autocorrelation Estimate

variance, normalized to the sample mean, is plotted in Figure 42 as a function of τ. Admittedly, a sample size of three may not present an extremely convincing basis for an argument of convergence of the estimate, but on the other hand, it was surprising to discover the striking similarity in the autocorrelation functions. The point at τ = 227 μseconds might exhibit greater percent variance than one would like (although its actual magnitude is small), but the remaining points tend to support a conclusion that this averaged estimate is a credible description of the time averaged autocorrelation of the three records in an ensemble sense. The average of the power spectral density functions is shown in Figure 43.

5.6 Summary and Conclusions of the Detailed Investigation. Before summarizing the results obtained, it seems appropriate at this point to illuminate the limitations imposed on the study at the onset. One limitation is an effective measurement bandwidth of 6 KHz to 30 KHz, taking into account the 32:1 time base expansion referred to in Chapter II. A second limitation results from the interference of the 18.75 KHz signal coupled into the measurement system from the aircraft electrical system. Thus, storm noise amplitudes which are approximately 18 db below saturation level are masked by the sinusoidal signal. A third limitation is the saturation of the measurement system, presumably caused by a near return stroke. It is thought that none of these three limitations are catastrophic. The frequency bandwidth restricts the study to the VLF spectrum. The minimum resolution restriction requires that the investigation concentrate on data segments exhibiting intense sferic activity, and the saturation of the measurement system is believed to occur relatively infrequently so that the study, realistically, concentrates on

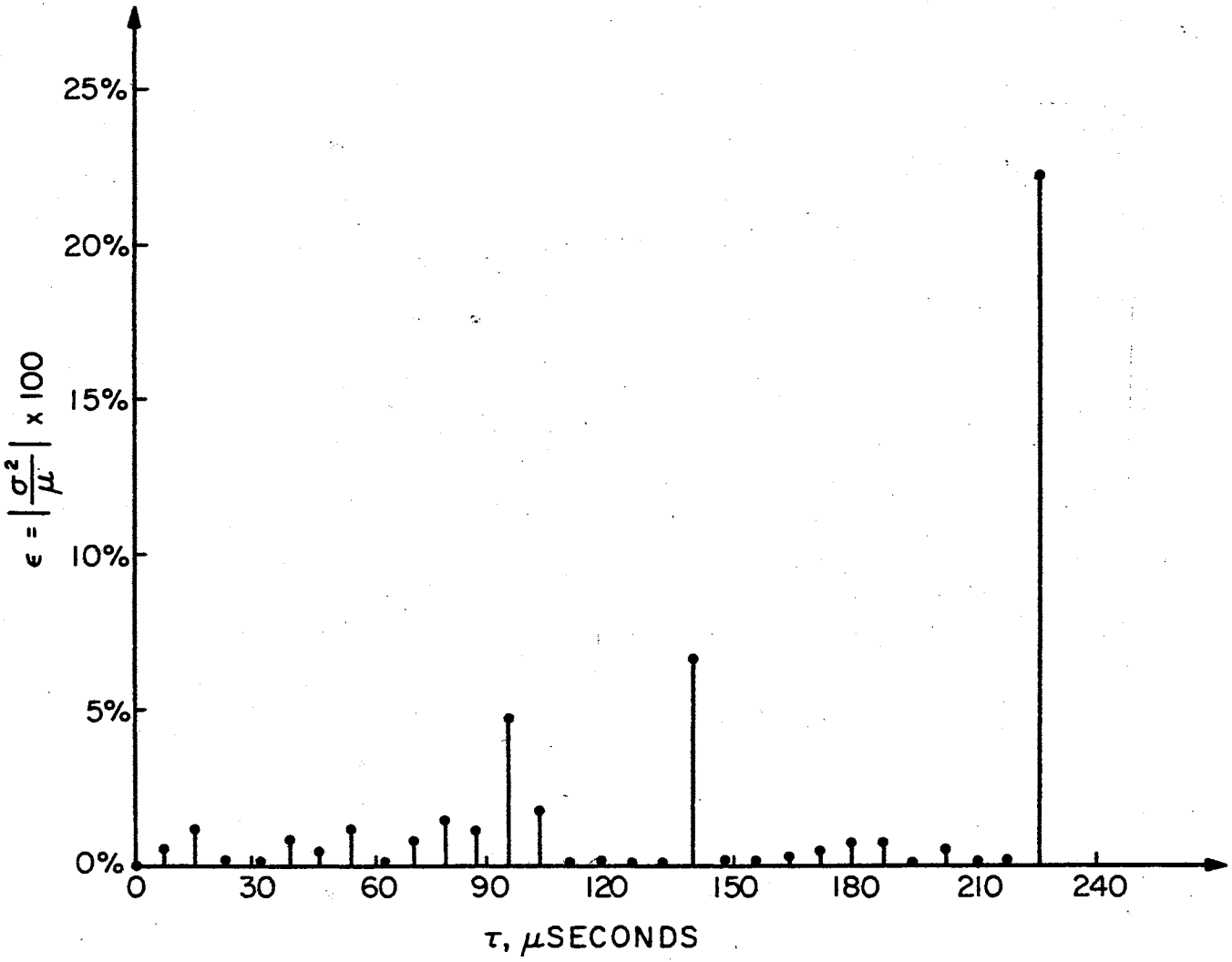


Figure 42. Sample Variance of the Averaged Autocorrelation Estimate

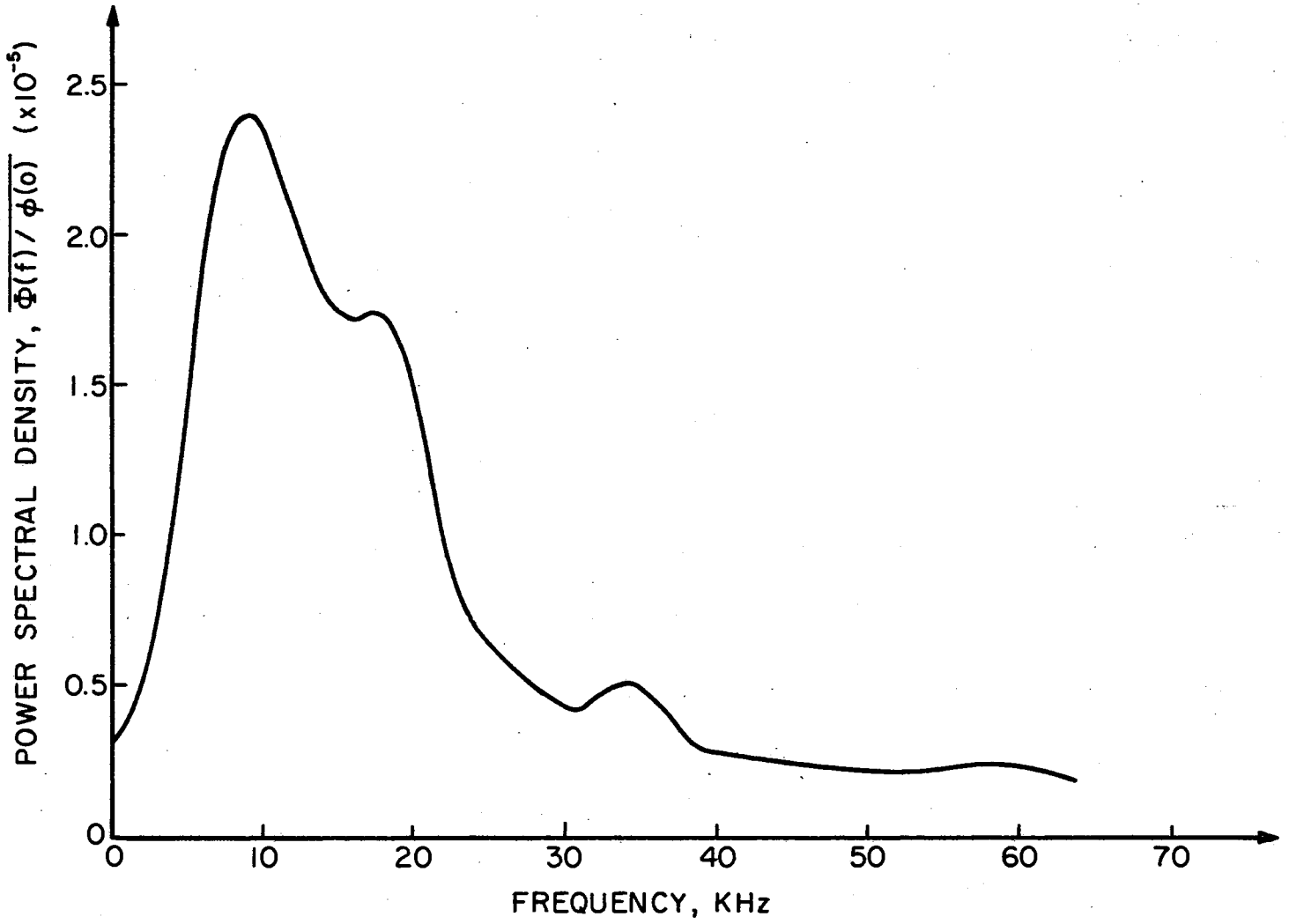


Figure 43. Averaged Power Spectral Density Estimate

the more frequently occurring leader and junction streamer field changes.

Summarizing the study of the time averaged second-order density functions, it appears that while they do reveal valuable information about the spread in the amplitude of the sferics, there is no supporting evidence apparent that indicates a time averaged second-order density function is a credible description of the process in an ensemble sense. The empirical density function reveals knowledge first of all, concerning the setting of the output gain of the magnetic tape replay unit. For example, if the probability that the product $x(t)x(t + \tau)$ assumes extremum values is relatively high, say .05, this would imply that the gain setting is too high and many of the sferics are exceeding the peak allowable input of the analog to digital converter. While this effect is observable through the use of an oscilloscope, it is impossible to visually estimate the rate of occurrence. On the other hand, if the probability that the product $x(t)x(t + \tau)$ is heavily concentrated within the lower seven amplitude windows and is zero for the $\pm$ 10 volt windows, then the replay gain is too low. Hence, in this respect the empirical second-order density function is essential. Secondly, having once set the gain for a particular record, the second-order density function reveals knowledge concerning the relative sferic intensity between two records on one data tape. This relieves some of the necessity of having to transfer the data from magnetic tape to an analog plot and then visually assessing the relative sferic intensity. In both of these examples of the use of the second-order empirical density function, it should be observed that the function actually describes one member function of the process in a serial sense contrary to the usual

connotation in an ensemble sense. In order to justify the time averaged second-order density function as a description of the process in an ensemble sense, one would have to present a credible argument for strict sense stationarity of at least degree two and further establish ergodicity of the second-order density function. This would imply that

$$f(x_1,x_2;t + \varepsilon,\ t + \tau + \varepsilon) = f(x_1,x_2;t,\ t + \tau)$$

for every ε. Then one would have to compute the mean squared error of an averaged estimate as a function of x_1, x_2 and τ, assuming that the estimate was formed in accordance with the theory of Appendix A. Based on the evidence of some nonstationarity in all three records studied, the author can see little hope of presenting a credible argument for a strict sense stationary approximation much less ergodicity. Hence, it is concluded that an attempt to model the process in terms of a time averaged second-order density function is not a worthwhile endeavor.

Summarizing the results of the study of the autocorrelation and power spectral density functions, it appears that there is cause to believe a credible approximation of the time averaged autocorrelation and power spectral density functions of the process has been estimated, at least for those intervals in which intense sferic activity is occurring. It should also be pointed out that these estimates have been normalized to the average received power in order to provide a basis for comparison between the data segments. Also, since the data were gathered at approximately 30 kilometers distance from the storm cell, the curves would have to be modified in accordance with known atmospheric propagation effects in order to describe the storm process at points more distant from, or nearer, the storm cell.

Because the autocorrelation and power spectral density functions are a Fourier transform pair and one uniquely specifies the other, a discussion of either function implicitly deals with the other function also. Since there have been several reported studies made of the spectra of lightning discharges, it seems reasonable then to consider the power spectral density in the remaining portion of this summary.

Comparing the time averaged power spectral density estimate with the spectral content computed by Taylor (23), it is observed that the low frequency content offered here falls off much more rapidly than he predicts, and the high frequency content reported here does not fall off as rapidly as he predicts. Specifically, Taylor predicts a low frequency rise going as $\sqrt{\text{frequency}}$ from 1 KHz to 5 KHz and then a fall off as 1/frequency from 5 KHz to 50 KHz with a slight hump at 60 KHz and then falling again between 60 KHz and 64 KHz. Extrapolating this result to the power spectral density curve, one would expect a rise proportional to frequency and a fall going as $1/(\text{frequency})^2$. Actually, the low frequency response falls off approximately five times faster than Taylor predicts. It is believed that this is primarily attributable to the low frequency response of the measurement system. Arnold and Pierce (2) report peak amplitudes of the frequency spectra for stepped-leader radiation, K-change radiation, and return-stroke radiation at 20, 8 and 5 KHz, respectively. Because of the saturation of the measurement system occurring during return strokes and the corresponding loss of the 5 KHz component, it seems reasonable that, here, the spectral density will peak around 8 KHz. This compares with a 9 KHz peak estimated by the time averaged power spectral density. It is believed, too, that the stepped-leader and K-change radiation occurring

in the measured data account for the less rapid fall off with high frequency as compared with Taylor's study. Thus, the power spectral density function seems reasonably consistent with the work of other investigators with the exception of the hump at 34.5 KHz. There exists no reference to this component in the literature surveyed. At first it was thought that this component was due to a reflected wave from the earth or from the ionosphere, but its absence in Record #1 would lead one to the belief that its source is in the discharge mechanism of the storm. A review of the three records indicate that Records #2 and #3 exhibited two or more sustained discharges while Record #1 was composed exclusively of short bursts. Hence, it is speculated that the sustained type discharges promote the radiation of this component although the exact cause for this phenomenon is unknown.

As a result of the data analysis, the following conclusions are drawn: (1) because of the striking similarity in the three estimates of the time averaged autocorrelation and power spectral density functions, it is concluded that the sample average of these functions is a reasonable characterization of the three records in an ensemble sense; (2) based on the preliminary study it appears that these three records are a representative sampling of those data segments evidencing intense sferic activity of the type occurring immediately before a discharge or while a discharge is in progress; and (3) based on the two conclusions above, it appears reasonable to assume that the sample average of the time averaged autocorrelation and power spectral density function estimates is a representative description of the process for segments evidencing extremely intense sferic activity.

CHAPTER VI

SUMMARY AND CONCLUSIONS

6.1 Summary. A review of this study reveals that the investigation is separable into two major areas. First, there is an investigation of the degree of stationarity and secondly, there is an investigation of the time averaged second-order statistics of the process. In order to attach any significance to the study of time averaged second-order statistics, in particular the autocorrelation and power spectral density functions, the process had to be shown to approximate the wide sense stationary condition. It was clear from the onset that the sferics are not time independent in a global sense; however, it appeared possible that the process might be stationary during the time periods evidencing intense sferic activity, i.e., in a local sense. Hence, the validity of this hypothesis was investigated on three data segments of 0.75 second duration. Essentially, the premise was made that the sferic triggering mechanism could be modeled by a Poisson impulse process, and that by identifying sferic bursts in the data record with triggering impulses, one could estimate the Poisson parameter. If the parameter could be shown to be independent of time, then there would exist reason to believe that the assumption of wide sense stationarity was credible. Conducting the test, it was concluded that because of the relatively high incidence of short sferic bursts compared with the length of separable segments of each data record, the study was composed of a

reasonably large sample size, and hence, yielded credible results. The most vulnerable point in the investigation of stationarity is that of identifying individual bursts within each data segment. Admittedly, this involved highly subjective judgement, but in as far as humanly possible, the judgment was uniformly applied.

After the investigation indicated wide sense stationarity was a credible assumption, it appeared that the study of the time averaged autocorrelation and power spectral density functions was a reasonable investigative approach in the search for a description of thunderstorm interference. In order to estimate the time averaged autocorrelation and power spectral density functions, an estimate of the time averaged second-order density function was made by time averaging in a serial sense over each record. Hence, for each record segment a probability density function was obtained in an empirical sense. Inferences could be made from this density function regarding the data segments in a serial sense, but no inferences beyond the first and second moments could be credibly made about the process in an ensemble sense. From the effort to determine stationarity, it appeared that the second-order density function did provide a mechanism for the comparative assessment of sferic intensity among the data records analyzed. Using each estimate of the density function, the time averaged autocorrelation function was estimated. The Fourier transform of the autocorrelation function was taken to produce an estimate of the time averaged power spectral density function.

As a preliminary study, the time averaged autocorrelation function was estimated for time lags out to 39 μseconds from twelve data segments. The result of the analysis revealed that the estimates bore

close resemblance. Because of the close resemblance among the several estimates, it was decided to refine the study to three representative data segments and consider time lags out to 227 μseconds, or the point at which the process essentially became uncorrelated. Estimates of the time averaged autocorrelation and power spectral density functions from two data segments of 0.75 second duration each, and separated in time by five minutes, revealed striking similarities after being normalized to the average power of the corresponding segments. Analysis of a third segment of 0.75 second duration from a thunderstorm occurring three days later, resulted in time average autocorrelation and power spectral density function estimates which exhibited the same general characteristics of the estimates made from the first two data records.

6.2 Conclusions. As a result of the study of stationarity it has been concluded that (1) a Poisson model for the sferic triggering mechanism seems reasonable, particularly if λ is taken as a slowly varying function of time; (2) with the exception of the final quarter of the first data segment, the three data sets selected have constant Poisson parameters within $\pm$ 10%; and (3) the data can thus be analyzed with the assumption of wide sense stationarity without serious loss of credibility.

Proceeding from the assumption of wide sense stationarity, the study of the time averaged autocorrelation and power spectral density function estimates, reveals the following conclusions: (1) because of the striking similarity in the three estimates of the time averaged autocorrelation and power spectral density functions, it appears that the sample average of these estimates is a reasonable characterization of the three records in an ensemble sense; (2) based on the results of

the preliminary data analysis, it appears that these three records are a representative sampling of those data segments evidencing intense sferic activity of the type occurring immediately before a discharge or while a discharge is in progress; and (3) thus, it appears reasonable to assume that the mean of the time averaged autocorrelation and power spectral density estimates is a representative description of the process in an ensemble sense, for segments evidencing extremely intense sferic activity.

In a composite sense, it is believed that the study has culminated in an approximate description of the autocorrelation and power spectral density functions, which may be useful in the design of filters for the purpose of minimizing thunderstorm noise interference in a mean square sense. The applicability of the characterization offered in this study, in a practical sense, suffers three constraints. First, it is restricted to the VLF spectrum owing to the frequency response of the measurement system. Secondly, the description has been normalized to the average received power to provide a basis for comparison. The normalization was necessary because of the variation in intensity of sferics, although distance from the storm remained fixed. In order to utilize the curves in an absolute sense, one would have to make some estimate of the average received power. Third, the data were gathered at a distance of thirty kilometers from the storms. If the curves were to be used for other distances, one would have to modify the curves in accordance with known propagation characteristics.

6.3 Recommendations for Further Study. The frustrations and successes experienced in this study suggest three areas of proposed investigation. It is believed that a similar study over a broader portion

of the frequency spectrum would be extremely worthwhile because it appears that the primary source of noise interference in the LF spectrum is thunderstorm noise. This study describes the process over only 20% of this portion of the spectrum. A second area of suggested investigation arises from the need for an estimate of the average radiated power during periods of intense sferic activity. It is believed that a statistical description as a function of distance from the storm would be useful in a practical sense. Perhaps this statistic could also be related to the meteorological content of the storm. A third area is suggested from the study of the time averaged density function estimates. It is believed that if the 18.75 KHz interference were eliminated from the recorded data, one could conduct a study of the fine structure of the sferic signal and perhaps gain some knowledge of the process in a phenomenological sense.

BIBLIOGRAPHY

1. Apostol, T. M. Mathematical Analysis. Reading Mass.: Addison-Wesley Publishing Company, Inc., 1964.

2. Arnold, H. R. and E. T. Pierce. "Leader and Junction Processes in the Lightning Discharge as a Source of VLF Atmospherics." Radio Science Journal of Research NBS/USNC-URSI. Vol. 68D, No. 7, (July, 1964).

3. Bendat, J. S. and A. G. Piersol. Measurement and Analysis of Random Data. New York: John Wiley & Sons, Inc., 1966.

4. Berglund, G. D. "A Guided Tour of the Fast Fourier Transform." IEEE Spectrum. Vol. 6, No. 7, (July, 1969).

5. Bingham, C., M. D. Godfrey, and J. W. Tukey. "Modern Techniques of Power Spectrum Estimation." IEEE Transactions on Audio and Electroacoustics. Vol. AU-15, No. 2, (June, 1967).

6. Blackman, R. B. and J. W. Tukey. The Measurement of Power Spectra From the Point of View of Communication Engineering. New York: Dover Publications Incorporated, 1958.

7. Chatterjee, B. Propagation of Radio Waves. New York: Asia Publishing House, 1963.

8. Davenport, W. B., Jr. and W. L. Root. An Introduction to the Theory of Random Signals and Noise. New York: McGraw-Hill Book Co., 1958.

9. Hildebrand, F. B. Introduction to Numerical Analysis. New York: McGraw-Hill Book Company, Inc., 1956.

10. Hogg, R. V. and A. F. Craig. Introduction to Mathematical Statistics. New York: The Macmillan Company, 1965.

11. Horner, F. "A Review of Information on Atmospherics From Near Lightning Discharges." In Radio Noise of Terrestrial Origin. Ed. F. Horner. New York: Elsevier Publishing Co., 1962.

12. Horner, F. "Radio Noise From Thunderstorms." In Advances in Radio Research. Ed. J. A. Saxton. New York: Academic Press, 1964.

13. Jordan, E. C. Electromagnetic Waves and Radiating Systems. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1960.

14. Kitagawa, N. and M. Brook. "A Comparison of Intracloud and Cloud-to-Ground Lightning Discharges." Journal of Geophysical Research. Vol. 65, No. 4, (April, 1960).

15. Kitagawa, N. and M. Kobayashi. "Field-Changes and Variations of Luminosity due to Lightning Flashes." In Recent Advances in Atmospheric Electricity. Ed. L. G. Smith. New York: Pergamon Press, 1958.

16. Malan, D. J. Physics of Lightning. London: The English Universities Press Ltd., 1963.

17. Papoulis, A. Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill Book Co., 1965.

18. Papoulis, A. The Fourier Integral and its Applications. New York: McGraw-Hill Book Co., 1962.

19. Prabhu, N. U. Stochastic Processes. New York: The Macmillan Company, 1966.

20. Pybus, E. J. Untitled Memo. March 16, 1970.

21. Richards, P. I. "Computing Reliable Power Spectra." IEEE Spectrum. Vol. 4, No. 1, (January, 1967).

22. Rao, C. R. Linear Statistical Inference and its Applications. New York: John Wiley & Sons, Inc., 1965.

23. Taylor, W. L. "Waveforms." In Problems of Atmospheric and Space Electricity. Ed. S. C. Coroniti. New York: Elsevier Publishing Co., 1965.

24. Uman, M. A. Lightning. New York: McGraw-Hill Book Co., 1969.

APPENDIX A

MATHEMATICAL THEORY OF THE DATA ANALYSIS

A.1 Introduction. This appendix deals with the problem of estimating the second-order statistics of a random process by a time average taken over an arbitrary member function. The principal advantage in considering an estimate based on time averaged data is that data records can be treated in a serial format as opposed to the generation of samples in an ensemble sense. Computationally speaking, this feature is of extreme importance. The development of this appendix begins with three general theorems concerning the estimation of the time averaged second-order probability distribution function, autocorrelation function and power spectral density. This theoretical discussion is then specialized to the particular problem of the application to thunderstorm data analysis and subsequent implementation into a computer program.

A.2 Theory of Time Averaged Estimates. The second-order statistics which are of interest in this study are defined as follows: the probability distribution function,

$$F(x_1,x_2;t,t+\tau) = P\{X(t) \leq x_1,\ X(t+\tau) \leq x_2\} \quad ,$$

the probability density function

$$f(x_1,x_2;t,t+\tau) = \frac{\partial^2 F(x_1,x_2;t,t+\tau)}{\partial x_1 \partial x_2} \quad ,$$

the autocorrelation function,

$$R_{xx}(t,t+\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1,x_2;t,t+\tau)dx_1 dx_2 = E\{X(t)X(t+\tau)\} \quad ,$$

and the power spectral density function,

$$S(t,\omega) = \int_{-\infty}^{\infty} R_{xx}(t,t+\tau)e^{-j\omega\tau}d\tau \quad .$$

Using the definitions given above, three theorems are presented concerning the estimation of these second-order statistics based on a time average taken over an arbitrary member function of the ensemble. Consideration is first directed toward estimating the time averaged second-order distribution function.

Theorem A.2.1. Let the process $Y_i(x_1;t)$ and $Z_i(x_2;t+\tau)$ be defined on a member function $x_i(t)$ of the random process $X(t)$ as follows,

$$Y_i(x_1;t) = \begin{cases} 1 & \text{if } x_i(t) \leq x_1 \\ 0 & \text{if } x_i(t) > x_1 \end{cases}$$

$$Z_i(x_2;t+\tau) = \begin{cases} 1 & \text{if } x_i(t+\tau) \leq x_2 \\ 0 & \text{if } x_i(t+\tau) > x_2 \end{cases} \quad .$$

If the average product of these two derived processes exists for arbitrary x_1 and x_2,

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} Y_i(x_1;t)\, Z_i(x_2;t+\tau)dt$$

then the averaged product is an unbiased estimate of the time averaged second-order distribution function such that,

$$E_i\{\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} Y_i(x_1;t)\, Z_i(x_2;t+\tau)dt\} = \lim_{T\to\infty} < F(x_1,x_2;t,t+\tau) >_T$$

$$= < F(x_1,x_2;t,t+\tau) >$$

where the symbol <> denotes an average over all t.

Proof. Consider the time average

$$G_i(x_1,x_2;\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} Y_i(x_1;t)\, Z_i(x_2;t+\tau)dt \quad .$$

For fixed x_1, x_2 and τ, $G_i(x_1,x_2;\tau)$ is a random variable which maps the ith event from the sample space onto the real line between [0,1]. Further, its expectation over the ensemble is

$$E_i\{G_i(x_1,x_2;\tau)\} = E_i\{\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} Y_i(x_i;t)\, Z_i(x_2;t+\tau)dt\}$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E_i\{Y_i(x_1;t)\, Z_i(x_2;t+\tau)\}dt$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} P\{X(t) \leq x_1,\ X(t+\tau) \leq x_2\}dt$$

$$= < F(x_1,x_2;t,t+\tau) > \quad .$$

Thus the estimate $G_i(x_1,x_2;\tau)$ is unbiased.

This theorem, then, provides one technique for forming an estimate of the time averaged second-order probability distribution function. Interpreting this theorem, it says the following: If x_1, x_2 and τ are fixed but otherwise arbitrary, then the value of $G_i(x_1,x_2;\tau)$ depends solely on which member function of the process it was computed from. It changes with parameters x_1, x_2 and τ. Thus, a collection of sample values of $G_i(x_1,x_2;\tau)$ are generated by performing the time average operation over member functions of the ensemble of X(t). It should be clear that these samples are distributed about the mean given in Theorem A.2.1. Although any one $G_i(x_1,x_2;\tau)$ is an unbiased estimate of $< F(x_1,x_2;t,t+\tau) >$, the estimate will be improved by forming an average in the sense of a sample mean. For example, if the samples are uncorrelated, then

$$\mathrm{VAR}\{\overline{G_i(x_1,x_2;\tau)}\} = \frac{1}{n}\,\mathrm{VAR}\{G_i(x_1,x_2;\tau)\}$$

for a sample of size n. This, then, is how the estimate would be implemented: (1) generate a number of individual estimates, $G_i(x_1,x_2;\tau)$; and (2) compute an estimate of $< F(x_1,x_2;t,t+\tau) >$ using the sample mean of the $G_i(x_1,x_2;\tau)$'s.

A relevant question at this point is, under what conditions is $< F(x_1,x_2;t,t+\tau) >$ a useful representation of the process in an ensemble sense? Clearly, $< F(x_1,x_2;t,t+\tau) >$ will be a complete description in the event $F(x_1,x_2;t,t+\tau)$ is independent of t. Stating this more precisely, a necessary and sufficient condition for the process to be described by its time averaged second-order probability distribution function is that the process be strict sense stationary of order two and the sample size contributing to the estimate be unlimited. This

condition is said to exist if

$$F(x_1,x_2;t,t+\tau) = F(x_1,x_2;t+\varepsilon,t+\tau+\varepsilon)$$

for every ε and n approaching infinity.

Next, the mixed partial derivative of the YZ product can be used to obtain the time averaged autocorrelation function.

Theorem A.2.2. Let $Y_i(x_1;t)$ and $Z_i(x_2;t+\tau)$ be defined on an arbitrary member function of the process X(t) as in Theorem A.2.1. If for each i, the mixed partial derivative

$$\frac{\partial^2 < Y_i(x_1;t)\ Z_i(x_2;t+\tau) >}{\partial x_1 \partial x_2}$$

exists for every x_1, x_2 and τ and the process has finite nonzero average power

$$0 < \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X^2(t)dt < \infty$$

then the function

$$\phi_{xx}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 \frac{\partial^2 < Y_i(x_1;t)\ Z_i(x_2;t+\tau) >}{\partial x_1 \partial x_2} dx_1dx_2$$

is a nontrivial, unbiased estimator of the time averaged autocorrelation function such that,

$$E_i\{\phi_{xx}(\tau)\} = < R_{xx}(t,t+\tau) > \quad .$$

Proof. It will first be shown that $\phi_{xx}(\tau)$ is the time average of the product $X(t)X(t+\tau)$:

$$\phi_{xx}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 \frac{\partial^2}{\partial x_1 \partial x_2}\{\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} Y_i(x_1;t)Z_i(x_2;t+\tau)dt\}dx_1dx_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\{\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \frac{\partial^2[Y_i(x_1;t)Z_i(x_2;t+\tau)]}{\partial x_1 \partial x_2} dt\}dx_1dx_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\{\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \frac{\partial Y_i(x_1;t)}{\partial x_1} \cdot \frac{\partial Z_i(x_2;t+\tau)}{\partial x_2} dt\}dx_1dx_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\{\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \delta[x_1-X(t)]\delta[x_2-X(t+\tau)]dt\}dx_1dx_2$$

$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\delta[x_1-X(t)]\delta[x_2-X(t+\tau)]dx_1dx_2\}dt$$

$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} X(t)X(t+\tau)dt \quad .$$

Let $\tau = 0$. The condition of finite nonzero average power, then, implies

$$0 < \phi_{xx}(0) < \infty \quad .$$

Since $\phi_{xx}(0) \geq \phi_{xx}(\tau)$ for every τ, the estimate is nontrivial and, furthermore, finite. Taking the ensemble expectation over the sample space results in,

$$E\{\phi_{xx}(\tau)\} = E\{\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} X(t)X(t+\tau)dt\}$$

$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} E\{X(t)X(t+\tau)\}dt$$

$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_{xx}(t,t+\tau)dt$$

$$= < R_{xx}(t,t+\tau) > \quad .$$

The development on the right hand side of the equality simply states that the expected value of the time average is the time average of the expected value. This statistic, namely $< R_{xx}(t,t + \tau) >$, is the expected value of $\phi_{xx}(\tau)$, the latter being the expectation of the product in terms of an estimate of the time averaged second-order probability density function.

This theorem demonstrates the equivalence of the derived process method with the familiar technique of lagged products. An interpretation of this theorem is analogous to the discussion of Theorem A.2.1. Using the mixed partial derivative of the results of Theorem A.2.1, or equivalently the average lagged product, a collection of autocorrelation functions can be generated and then averaged for an improved estimate in the ensemble sense.

Again, the question arises, under what conditions is $< R(t,t + \tau) >$ a useful representation of the process in an ensemble sense? The time averaged autocorrelation function will be a complete description if $R(t,t + \tau)$ is independent of t. Hence, for the sample mean of the time averaged autocorrelation function to be a valid description of the process in an ensemble sense, the process must be wide sense stationary and the sample size leading to the mean must be unlimited.

Using the estimate of the time averaged autocorrelation function, its Fourier transform is an estimate of the time averaged power spectral density. Furthermore, this estimate is related to the deterministic Fourier transform of X(t).

Theorem A.2.3. Let $\phi_{xx}(\tau)$ be an estimate of the time averaged autocorrelation function. If the function $\Phi_{xx}(\omega)$ exists on a member function as follows,

$$\Phi_{xx}(\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\tau) e^{-j\omega\tau} d\tau$$

then $\Phi_{xx}(\omega)$ is an unbiased estimate of the time averaged power spectral density function such that,

$$E\{\Phi_{xx}(\omega)\} = < S(t,\omega) >$$

and it is related to the deterministic Fourier transform of $X(t)$, say $F(\omega)$, by

$$\Phi_{xx}(\omega) = \lim_{T\to\infty} \frac{1}{2T} \left|F_T(\omega)\right|^2 \quad .$$

Proof. The first part of the theorem follows by taking the expectation of $\Phi_{xx}(\omega)$,

$$\begin{aligned}
E\{\Phi_{xx}(\omega)\} &= E\{\int_{-\infty}^{\infty} \phi_{xx}(\tau) e^{-j\omega\tau} d\tau\} \\
&= \int_{-\infty}^{\infty} E\{\phi_{xx}(\tau)\} e^{-j\omega\tau} d\tau \\
&= \int_{-\infty}^{\infty} < R_{xx}(t,t+\tau) > e^{-j\omega\tau} d\tau \\
&= \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R_{xx}(t,t+\tau) dt \; e^{-j\omega\tau} d\tau \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \int_{-\infty}^{\infty} R_{xx}(t,t+\tau) e^{-j\omega\tau} d\tau dt \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} S(t,\omega) dt \\
&= < S(t,\omega) > \quad .
\end{aligned}$$

Hence, the estimate is unbiased, and moreover, $< R_{xx}(t,t+\tau) >$ and

$< S(t,\omega) >$ are a Fourier transform pair. Proof of the second part of the theorem is given by Papoulis (18) and an outline of that work is offered for reference. Let $F_T(\omega)$ denote the Fourier transform of a member function $x_i(t)$ over the finite interval $(-T,T)$,

$$F_T(\omega) = \int_{-T}^{T} x_i(t) e^{-j\omega t} dt \quad .$$

The average power over the interval $(-T,T)$ for an arbitrary frequency ω is

$$\Phi_T(\omega) = \frac{1}{2T} |F_T(\omega)|^2 \quad .$$

The Helly-Bray theorem asserts

$$\Phi_{xx}(\omega) = \lim_{T\to\infty} \Phi_T(\omega) = \lim_{T\to\infty} \frac{1}{2T} |F_T(\omega)|^2 \quad .$$

The interpretation of this theorem is analogous to that of Theorem A.2.2. The second part of this theorem will be used to relate the power spectral density estimate to the investigations made by other research workers in the area of thunderstorm spectra.

From the preceeding discussion of the time averaged autocorrelation and owing to the uniqueness of the Fourier transform, $< S(t,\omega) >$ will be a valid representation of the process in an ensemble sense if the process is wide sense stationary. Elaborating on this point, consider the autocorrelation function $R(t,t+\tau)$ and define the Fourier transform with respect to t and τ by

$$\Delta(\omega_1,\omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(t,t+\tau) e^{-j(\omega_1\tau - \omega_2 t)} d\tau dt$$

and its inverse transform by,

$$R(t,t+\tau) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta(\omega_1,\omega_2) e^{j(\omega_1\tau-\omega_2 t)} d\omega_1 d\omega_2 \quad .$$

Consider the time average,

$$\frac{1}{2T} \int_{-T}^{T} R(t,t+\tau)dt = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta(\omega_1,\omega_2) e^{j\omega_1\tau} \frac{1}{2T} \int_{-T}^{T} e^{-j\omega_2 t} dt d\omega_1 d\omega_2$$

$$< R(t,t+\tau) > = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta(\omega_1,\omega_2) e^{j\omega_1\tau} \lim_{T\to\infty} \frac{\sin \omega_2 T}{\omega_2 T} d\omega_1 d\omega_2$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} [\frac{1}{2\pi} \int_{-\infty}^{\infty} \Delta(\omega_1,\omega_2) \lim_{T\to\infty} \frac{\sin \omega_2 T}{\omega_2 T} d\omega_2] e^{j\omega_1\tau} d\omega_1 .$$

From the proof of Theorem A.2.3, it is observed that the portion of the integrand enclosed in brackets must be identically equal to $< S(t,\omega) >$ due to the uniqueness of the inversion integral. If $\Delta(\omega_1,\omega_2)$ is interpreted as mass density in the $\omega_1\omega_2$ plane and since

$$\lim_{T\to\infty} \frac{\sin \omega_2 T}{\omega_2 T} = \begin{cases} 1 & \omega_2 = 0 \\ 0 & \omega_2 \neq 0 \end{cases}$$

then $< S(t,\omega) >$ is the line mass residing on the ω_1 axis. This implies that $< R(t,t+\tau) >$ is uniquely described by the line mass on the ω_1 axis, and that all masses off this axis play no part in this description. Clearly, if the line mass on the ω_1 axis is zero, then $< R(t,t+\tau) > = 0$. The condition of finite nonzero average power, then, ensures that $< S(t,\omega) >$ is nonzero. Let the process be wide sense stationary with autocorrelation and power spectral density functions given by $R(\tau)$ and $S(\omega)$, respectively. Computing the double Fourier transform $\Delta(\omega_1,\omega_2)$ results in

$$\Delta(\omega_1,\omega_2) = 2\pi S(\omega_1)\delta(\omega_2) \quad .$$

Thus, for the stationary case, $\Delta(\omega_1,\omega_2)$ is localized on the ω_1 axis. If $R(t,t+\tau)$ changed slowly as a function of t, then this would imply that $\Delta(\omega_1,\omega_2)$ would be concentrated about the ω_1 axis. For this case, then, $< S(t,\omega) >$, exhibiting all the properties of $S(\omega)$, would be a quasi-stationary approximation of $\Delta(\omega_1,\omega_2)$.

The three theorems offered in this section provide a mathematical basis for the technique of estimating a time averaged second-order distribution function, a time averaged autocorrelation function and a time averaged power spectral density function. A prime consideration in the problem of estimation is the interaction of the interval length (-T,T) and the number of member functions of the ensemble which are considered. As a practical matter, both of these parameters must be finite, and clearly they will depend upon the nature of the process itself. The author knows of no analytical way to optimize the selection of these parameters without a priori knowledge of the statistics which are being estimated.

A.3 Practical Considerations in Applying the Estimation Theory.

In Section A.2, it was shown that if a process were wide sense stationary, then the time average of the autocorrelation and power spectral density functions are valid representations of the process. In Section A.2, there are essentially three approaches suggested for use in the estimation of the time averaged statistics. One method is suggested in the proof of Theorem A.2.2, namely that of an averaged lagged product, i.e.

$$\phi_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)X(t+\tau)dt \quad .$$

Using this result, then, the method implied in Theorem A.2.3 would produce an unbiased estimate of the time averaged power spectral density. This procedure is analogous to the "indirect method" outlined by Blackman and Tukey (6). A second technique is implied in Theorem A.2.3, namely that of computing the power spectral density function from X(t), i.e.

$$\Phi_{xx}(\omega) = \lim_{T\to\infty} \frac{1}{2T} |F_T(\omega)|^2 \quad .$$

Using this result, then, the inverse Fourier transform of $\Phi_{xx}(\omega)$ could be taken to produce an estimate of $\phi_{xx}(\tau)$. This procedure is analogous to the "direct method" outlined by Blackman and Tukey. The third approach suggested begins with the estimation of the probability density function, and then successively computing estimates of the time averaged autocorrelation and power spectral density functions.

The third method was selected as the investigative approach to be used in this study. It was concluded that by including an investigation of the time averaged probability density function, a further quantitative measure of the sferic activity could be assessed. In illustration, it is possible that two segments of data have identical autocorrelation functions and yet evidence entirely different sferic activity. Consider the autocorrelation of zero mean shot noise,

$$R_{nn}(\tau) = \lambda \int_{-\infty}^{\infty} h(\tau+\beta)h(\beta)d\beta$$

where λ is the density of triggering impulses. Then, let $\lambda_1 = \lambda_2/a^2$

where a is an arbitrary constant and $h_1(t) = ah_2(t)$, for every t. Thus,

$$R_{11}(\tau) = R_{22}(\tau) \quad .$$

However, the first case differs from the second in that the impulses are more numerous but less intense.

The quantitative measure of the sferic activity will be assessed by the spread in probability mass. Note that although the time averaged distribution function is adequate for computing the time averaged autocorrelation and power spectral density functions, it is not necessarily a valid approximation to the distribution function of a wide sense stationary process. Wide sense stationarity is a necessary but not sufficient condition to ensure validity. Second-order stationarity in the strict sense is both necessary and sufficient.

The second-order probability density function is given by

$$f(x_1,x_2;t,t+\tau) = \lim_{\substack{\Delta x_1 \to 0 \\ \Delta x_2 \to 0}} \frac{P\{x_1 - \frac{\Delta x_1}{2} \leq X(t) \leq x_1 + \frac{\Delta x_1}{2},\ x_2 - \frac{\Delta x_2}{2} \leq X(t+\tau) \leq x_2 + \frac{\Delta x_2}{2}\}}{\Delta x_1 \Delta x_2} \quad .$$

Hence, to estimate $f(x_1,x_2;t,t+\tau)$, an estimate of

$$P\{x_1 - \frac{\Delta x_1}{2} \leq X(t) \leq x_1 + \frac{\Delta x_1}{2}\ ,\ x_2 - \frac{\Delta x_2}{2} \leq X(t+\tau) \leq x_2 + \frac{\Delta x_2}{2}\}$$

is formed and the result is divided by $\Delta x_1 \Delta x_2$. The estimate is formed using the method outlined in Theorem A.2.1 by defining the random process $\alpha_k(t + k\tau)$ as follows,

$$\alpha_k(t + k\tau) = \begin{cases} 1 & \text{if } x_k - \frac{\Delta x_k}{2} \leq X(t + k\tau) \leq x_k + \frac{\Delta x_k}{2} \\ 0 & \text{otherwise} \end{cases} .$$

Consider the product $\alpha_0(t)\alpha_1(t + \tau)$ and compute the expectation,

$$E\{\alpha_0(t)\alpha_1(t + \tau)\} =$$

$$P\{x_0 - \frac{\Delta x_0}{2} \leq X(t) \leq x_0 + \frac{\Delta x_0}{2} , x_1 - \frac{\Delta x_1}{2} \leq X(t + \tau) \leq x_1 + \frac{\Delta x_1}{2}\} .$$

Thus,

$$f(x_0,x_1;t,t + \tau) = \lim_{\substack{\Delta x_0 \to 0 \\ \Delta x_1 \to 0}} \frac{E\{\alpha_0(t)\alpha_1(t + \tau)\}}{\Delta x_0 \Delta x_1} .$$

Now, consider the finite time-average $< \alpha_0(t)\alpha_1(t + \tau) >_T$. Suppose $X(t)$ is known only for $|t| < T$. If $\tau > 0$, then $t + \tau > T$ for $t > T - \tau$, and $\alpha_0(t)\alpha_1(t + \tau)$ can be integrated only on the interval $(-T,T - \tau)$. Hence, assume $X(t)$ is specified over $(-T,T + \tau)$ and there results

$$< \alpha_0(t)\alpha_1(t + \tau) >_T = \frac{1}{2T} \int_{-T}^{T} \alpha_0(t)\alpha_1(t + \tau)dt .$$

This form must be modified, however, because $X(t)$ is sampled at regular intervals in the computational process. Assume that $X(t)$ differs from zero only for the time interval $(-T,T + \tau)$ seconds, and that its Fourier transform $\Phi_{xx}(\omega)$ exists only in a frequency interval from -B to B Hz, then the sampling theorem implies that $2B(2T + \tau)$ distinct samples are required to describe $X(t)$ completely. (The dual assumption of a time limited and band limited signal is not theoretically possible; however, for large time-bandwidth products, this condition is closely

approximated.) Therefore, let X(t) be periodically sampled at a rate of 2B samples per second. This implies a discrete form of the time average as follows,

$$< \alpha_0(t)\alpha_1(t + \tau) >_T = \frac{1}{n} \sum_{j=1}^{n} \alpha_0(t_j)\alpha_1(t_j + \tau)$$

where n is the greatest integer equal to or less than 4BT. This expression will be used to form the estimate, $< \alpha_0(t)\alpha_1(t + \tau) >_T$.

Having formed an estimate, it now remains to investigate its properties. First the biasedness is studied. Consider,

$$E\{< \alpha_0(t)\alpha_1(t + \tau) >_T\} =$$

$$\frac{1}{n} \sum_{j=1}^{n} P\{x_0 - \frac{\Delta x_0}{2} \leq X(t_j) \leq x_0 + \frac{\Delta x_0}{2},\ x_1 - \frac{\Delta x_1}{2} \leq X(t_j + \tau) \leq x_1 + \frac{\Delta x_1}{2}\}$$

$$= <P\{x_0 - \frac{\Delta x_0}{2} \leq X(t) \leq x_0 + \frac{\Delta x_0}{2},\ x_1 - \frac{\Delta x_1}{2} \leq X(t + \tau) \leq x_1 + \frac{\Delta x_1}{2}\}>_T$$

$$= < \int_{x_1 - \frac{\Delta x_1}{2}}^{x_1 + \frac{\Delta x_1}{2}} \int_{x_0 - \frac{\Delta x_0}{2}}^{x_0 + \frac{\Delta x_0}{2}} f(\beta_1,\beta_2;t,t + \tau)d\beta_1 d\beta_2 >_T .$$

The Taylor series expansion of $f(\beta_1,\beta_2;t,t + \tau)$ about (x_0,x_1) can be written,

$$f(\beta_1,\beta_2) = f(x_0,x_1) + (\beta_1 - x_0)D_1 f(x_0,x_1) + (\beta_2 - x_1)D_2 f(x_0,x_1)$$

$$+ \frac{(\beta_1 - x_0)^2}{2} D_1^2 f(x_0,x_1) + (\beta_1 - x_0)(\beta_2 - x_1)D_{12} f(x_0,x_1) + \frac{(\beta_2 - x_1)^2}{2}$$

$$\cdot D_2^2 f(x_0,x_1) + \ldots .$$

Using the relations,

$$\int_{x - \frac{\Delta x}{2}}^{x + \frac{\Delta x}{2}} (\beta - x)\, d\beta = 0$$

$$\int_{x - \frac{\Delta x}{2}}^{x + \frac{\Delta x}{2}} \frac{(\beta - x)^2}{2}\, d\beta = \frac{(\Delta x)^3}{24}$$

the following is obtained for $\Delta x_0 = \Delta x_1 = \Delta x$,

$$\frac{E\{< \alpha_0(t)\alpha_1(t + \tau) >_T\}}{(\Delta x)^2} = < f(x_0,x_1;t,t + \tau) >_T$$

$$+ < \frac{\Delta x}{24} \{D_1^2 f(x_0,x_1;t,t + \tau) + D_2^2 f(x_0,x_1;t,t + \tau)\} >_T + \ldots$$

Hence, the estimate of the time averaged second-order density function is a biased estimate; however, assuming $f(x_0,x_1;t,t + \tau)$ is a fairly smooth function so that the second and higher order partials are relatively insignificant, the bias becomes negligible.

To investigate the mean squared error, the variance of the estimate is first computed. Consider

$$E\{< \alpha_0(t)\alpha_1(t + \tau) >_T^2\} = \frac{1}{n^2} E\{ \sum_{r=1}^{n} \alpha_0(t_r)\alpha_1(t_r + \tau) \sum_{s=1}^{n} \alpha_0(t_s)\alpha_1(t_s + \tau)\}.$$

Let $A(t_r)$ denote the event,

$$\{x_0 - \frac{\Delta x_0}{2} \leq X(t_r) \leq x_0 + \frac{\Delta x_0}{2}\ ,\ x_1 - \frac{\Delta x_1}{2} \leq X(t_r + \tau) \leq x_1 + \frac{\Delta x_1}{2}\}$$

and let $B(t_s)$ denote the event,

$$\{x_0 - \frac{\Delta x_0}{2} \leq X(t_s) \leq x_0 + \frac{\Delta x_0}{2}\ ,\ x_1 - \frac{\Delta x_1}{2} \leq X(t_s + \tau) \leq x_1 + \frac{\Delta x_1}{2}\} \quad .$$

Hence,

$$E\{< \alpha_0(t)\alpha_1(t + \tau) >_T^2\} = \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} P\{A(t_r), B(t_s)\}$$

$$= \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} P\{B(t_s) \,|\, A(t_r)\} P\{A(t_r)\}$$

$$= \frac{1}{n} \sum_{r=1}^{n} < P\{B(t) \,|\, A(t_r)\} >_T P\{A(t_r)\}$$

$$= << P\{B(t) \,|\, A(t)\} >_T P\{A(t)\} >_T \quad .$$

But $P\{B(t) \,|\, A(t)\} = 1$ for every t; therefore,

$$E\{< \alpha_0(t)\alpha_1(t + \tau) >_T^2\} =$$

$$< P\{x_0 - \frac{\Delta x_0}{2} \leq X(t) \leq x_0 + \frac{\Delta x_0}{2},\ x_1 - \frac{\Delta x_1}{2} \leq X(t + \tau) \leq x_1 + \frac{\Delta x_1}{2}\} >_T$$

and,

$$VAR\{< \alpha_0(t)\alpha_1(t + \tau) >_T\} =$$

$$< P\{x_0 - \frac{\Delta x_0}{2} \leq X(t) \leq x_0 + \frac{\Delta x_0}{2},\ x_1 - \frac{\Delta x_1}{2} \leq X(t + \tau) \leq x_1 + \frac{\Delta x_1}{2}\} >_T$$

$$- < P\{x_0 - \frac{\Delta x_0}{2} \leq X(t) \leq x_0 + \frac{\Delta x_0}{2},\ x_1 - \frac{\Delta x_1}{2} \leq X(t + \tau) \leq x_1 + \frac{\Delta x_1}{2}\} >_T^2 \quad .$$

The mean squared error can now be written as

$$E\{[< \alpha_0(t)\alpha_1(t+\tau) >_T - < f(x_0,x_1;t,t+\tau) >_T]^2\} =$$

$$\text{VAR}\{< \alpha_0(t)\alpha_1(t+\tau) >_T\}$$

$$+ E\{[E\{< \alpha_0(t)\alpha_1(t+\tau) >_T\} - < f(x_0,x_1;t,t+\tau) >_T]^2\} \quad .$$

It is noted that the second term on the right is the average squared bias in the estimate. If it is accepted that the bias term is negligible, then the normalized mean squared error is defined to be

$$\varepsilon^2 = \frac{\text{VAR}\{< \alpha_0(t)\alpha_1(t+\tau) >_T\}}{< \alpha_0(t)\alpha_1(t+\tau) >_T^2} \quad .$$

An estimate of the minimum sample record length in seconds required to achieve a desired error ε^2 when B is the bandwidth in Hz occupied by X(t), assuming a uniform power spectrum within B, offered by Bendat and Piersol (3) is

$$T_r = \frac{\frac{1}{2}}{\varepsilon^2 B < \alpha_0(t)\alpha_1(t+\tau) >_T} \quad .$$

Thus, for a normalized mean squared error, ε^2, of .01, a probability resolution of .001 and a frequency bandwidth of 300 KHz, computation reveals that a data segment of 0.167 seconds is required.

Following several trial computational runs, Δx was selected to produce nine increments in the amplitude range of $\alpha(t)$. The amplitude window widths were 2.5 volts centered about 0, $\pm$ 2.5, $\pm$ 5.0, $\pm$ 7.5, and $\pm$ 10.0 volts.

Once $< \alpha_0(t)\alpha_1(t+\tau) >_T$ has been determined, the autocorrelation function is computed using

$$\phi_{xx}(\tau) = \sum_{x_1} \sum_{x_2} x_1 x_2 < \alpha_0(t)\alpha_1(t + \tau) >_T \quad .$$

The collection of $\phi_{xx}(\tau)$'s are used to estimate the time averaged autocorrelation function as follows,

$$< R_{xx}(t,t + \tau) > = \frac{1}{n} \sum_{i=1}^{n} \phi_i(\tau) \quad .$$

Using each of the $\phi_{xx}(\tau)$'s, the Fourier transform is taken to obtain an estimate of the power spectral density function, $\Phi_{xx}(\omega)$.

A.4 Additional Errors Incurred in the Power Spectral Density Estimate. In addition to the errors incurred in the estimation of the time averaged autocorrelation function, there are several remaining sources of error associated with the power spectral density estimate. There are essentially two sources of error to be considered, the first of which is due to aliasing. It has been mentioned previously that because the frequency bandwidth is B Hz, X(t) is sampled at a rate of 2BHz. To illustrate why this is necessary, let f_s be the sampling frequency. This implies that the spectrum of X(t) is folded about $f_s/2$, and if there are frequency components higher than $f_s/2$, they appear to be low frequency components. This effect is shown graphically in Figure 44. The dots are sample points of a signal, shown by the solid line, whose frequency is greater than $f_s/2$. This set of data points could also represent a signal, shown by the dashed line, whose frequency is below $f_s/2$. By sampling at a rate of 2BHz, the effects of aliasing are negligible. For a detailed treatment of this error, the reader is referred to Blackman and Tukey (6), pp. 31-33. A second source of error is due to leakage effects shown graphically in Figure 45. This error is

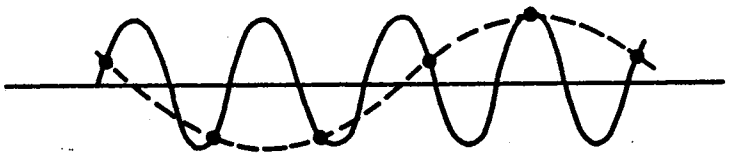


Figure 44. An Example of the Aliasing Effect

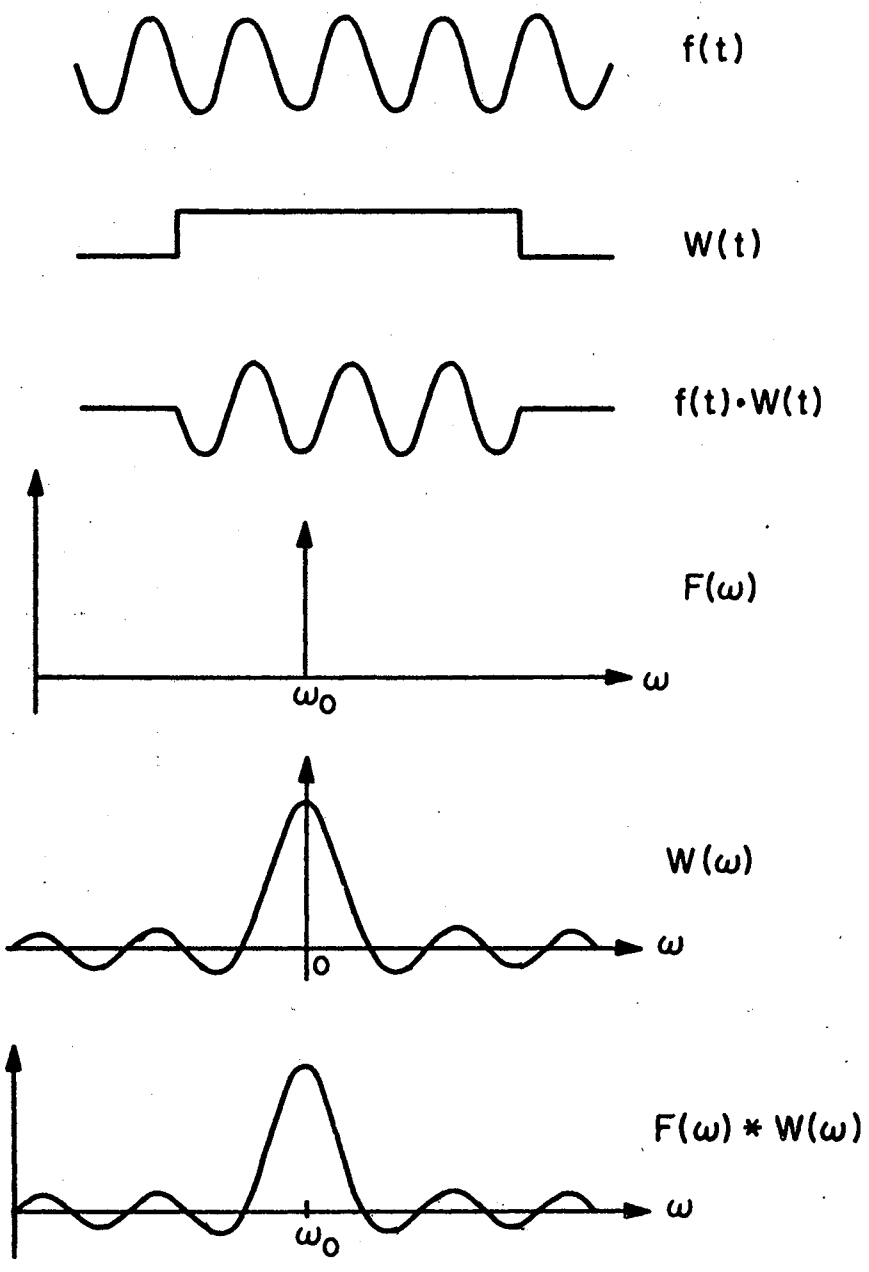


Figure 45. An Example of the Leakage Effect

due to the fact that the data record must be limited in time. If the one-sided Fourier transform were taken of the cosine wave, f(t), there would result a single impulse at ω_0; however, because the record has been truncated, its transform is not isolated but generates a series of spurious peaks called sidelobes. The objective is to localize the contribution of a given frequency by reducing the amount of "leakage" through the sidelobes. The usual approach consists of applying a lag window to the autocorrelation function, which has lower sidelobes in the frequency domain than the rectangular lag window. Specifically, one would like to concentrate the main lobe of $W(\omega)$ near $\omega = 0$, keeping the sidelobes as low as possible. In order to concentrate the main lobe, W(t) must be as wide as possible. In order to reduce the side lobes, W(t) must be smooth and slowly changing. Since W(t) must vanish for $|t| > T_M$, a compromise must be reached. A simple and convenient compromise recommended by Blackman and Tukey (6) is represented by the Hanning lag window

$$D_w(\tau) = \begin{cases} \frac{1}{2}\left(1 + \cos\frac{\pi\tau}{T_M}\right) & |\tau| < T_M \\ 0 & |\tau| > T_M \end{cases} .$$

This lag window function, then, will be used in the computation of $\Phi_{xx}(\omega)$.

Once the $\Phi_{xx}(\omega)$'s are estimated using the Fourier transform of the modified $\phi_{xx}(\tau)$'s, then the estimate of $< S(t,\omega) >_T$ is formed using

$$< \hat{S}(t,\omega) >_T = \frac{1}{n} \sum_{i=1}^{n} \Phi_i(\omega) .$$

APPENDIX B

EFFECT OF SINUSOIDAL COMPONENT UPON THE ESTIMATE OF HISTOGRAMS AND VARIANCE

B.1 Introduction. Reference has been made in the text of this study to the presence of an 18.75 KHz component in the recorded analog data. The purpose of this appendix is to investigate what effect the sinusoid has upon the empirical probability density function and the estimate of variance. In the following development, it is assumed that the analog data consists of the algebraic sum of a random process and a sinusoidal component whose phase is a random variable. Further, it is assumed that these two components are uncorrelated. This model is then used to investigate the influence of the sinusoid upon the estimate of variance for different amplitude window widths in the empirical probability density function.

B.2 Composite Density Function of the Sinusoid and the Random Component. Let the analog data be denoted by W(t) such that

$$W(t) = X(t) + Y(t)$$

where X(t) is a random process, and

$$Y(t) = A \sin(\omega_0 t + \phi) \quad .$$

Further, let

$$f_{\phi}(\phi) = \begin{cases} \frac{1}{2\pi} & |\phi| < \pi \\ 0 & \text{otherwise} \end{cases} .$$

Consider the autocorrelation function of W(t),

$$R_{ww}(t,t+\tau) = E\{[X(t) + Y(t)][X(t+\tau) + Y(t+\tau)]\}$$

$$= E\{X(t)X(t+\tau)\} + E\{Y(t)Y(t+\tau)\} .$$

The equality on the right side results from the assumption that X(t) and Y(t) are uncorrelated. Using a result derived by Papoulis (17), one can obtain,

$$R_{ww}(t,\ t+\tau) = R_{xx}(t,t+\tau) + \frac{A^2}{2}\cos\omega_0\tau .$$

The corresponding power spectral density is

$$S_w(\omega) = S_x(\omega) + \frac{A^2\pi}{2}[\delta(\omega+\omega_0) + \delta(\omega-\omega_0)] .$$

Hence, the presence of the sinusoid introduces a periodic component in the autocorrelation function and adds spectral components at $\pm\omega_0$. In order to reduce the effect of the sinusoid on the estimate of the time averaged autocorrelation and power spectral density functions, it was observed that the analog data were fairly inactive over 60% to 80% of the recording. Thus, it appeared reasonable to insert a threshold about the base line of the data, and thereby obtain a corresponding reduction in the effect of the sinusoid.

The upper plot of Figure 46 shows an approximate marginal probability density function of X(t) as deduced by a smoothing of the histograms reported in Chapter V. In order to quantify this density function

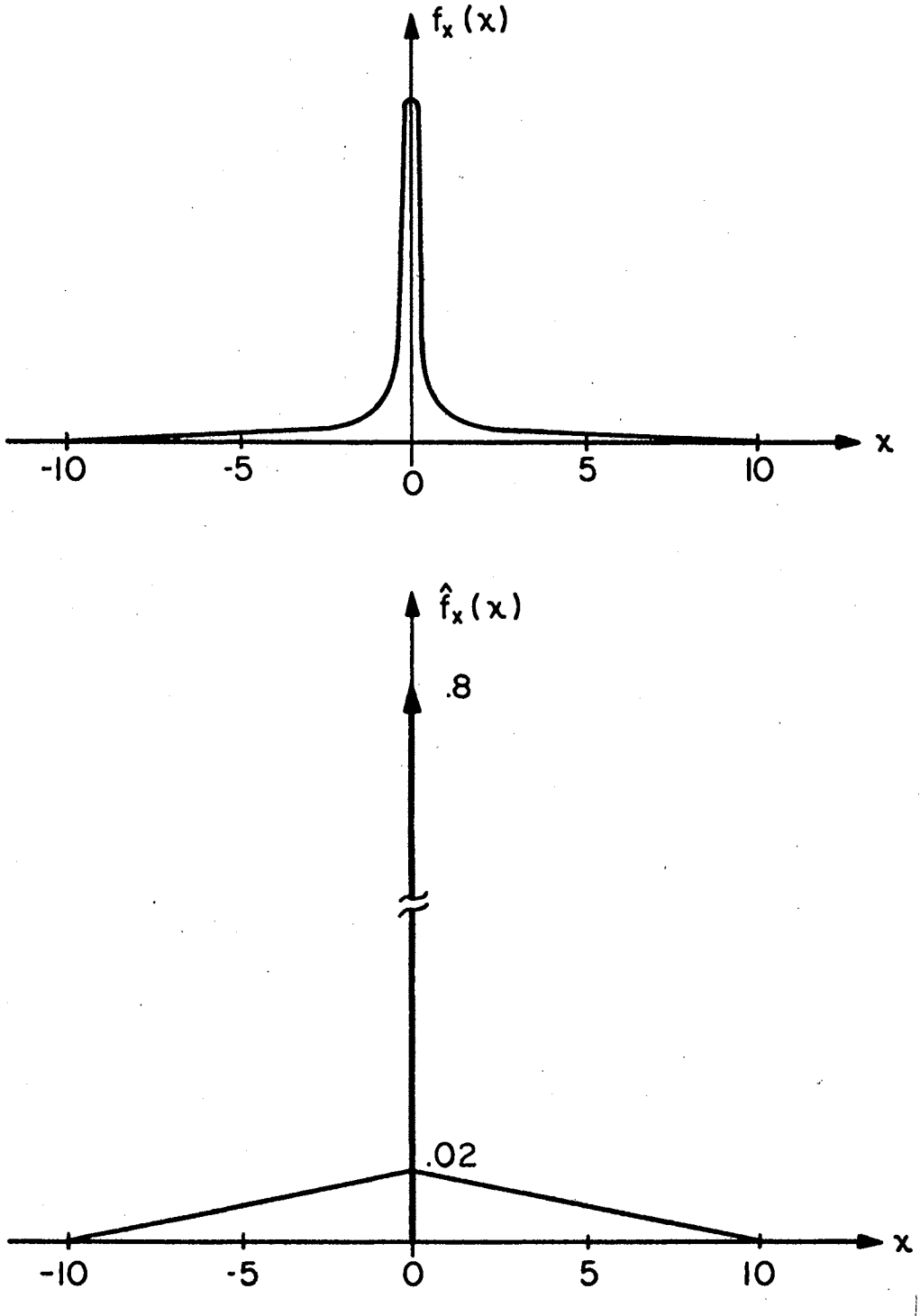


Figure 46. Plot of Marginal Probability Density of X(t) Compared With Approximate Density

and, also, to maintain computational simplicity, it is assumed that the density can be characterized by the lower plot of Figure 46. Here, the mass concentrated about the origin is described by an impulse with a weighting coefficient of .8. The remainder of the mass density is assumed to fall off linearly away from the origin out to peak amplitudes of $\pm$ 10 volts.

Papoulis (17) shows that $f_Y(y)$ is given by

$$f_Y(y) = \begin{cases} \dfrac{1}{\pi\sqrt{A^2-y^2(t)}} & |y(t)| < A \\ 0 & \text{otherwise} \end{cases} .$$

This distribution is shown in Figure 47, for A = 1 volt. Further, if it is assumed that X(t) and Y(t) are independent, the probability density function of W(t) can be determined by convolving $f_X(x)$ and $f_Y(y)$. Carrying out this computation results in,

$$f(\omega) = \begin{cases} \dfrac{1}{\pi 500}\left[(\omega+10)\sin^{-1}(\omega+10) + \sqrt{1-(\omega+10)^2} + \dfrac{\pi}{2}(\omega+10)\right] & -11 < \omega < -9 \\ \dfrac{1}{500}(\omega+10) & -9 < \omega < -1 \\ -\dfrac{2}{\pi 500}\left[\omega \sin^{-1}\omega + \sqrt{1-\omega^2} - \dfrac{\pi}{2}\cdot 10\right] + \dfrac{.8}{\pi}\dfrac{1}{\sqrt{1-\omega^2}} & -1 < \omega < 1 \\ \dfrac{1}{500}(10+\omega) & 1 < \omega < 9 \\ \dfrac{1}{\pi 500}\left[(\omega-10)\sin^{-1}(\omega-10) + \sqrt{1-(\omega-10)^2} - \dfrac{\pi}{2}(\omega-10)\right] & 9 < \omega < 11 \\ 0 & \text{elsewhere.} \end{cases}$$

A plot of this function is shown in Figure 48. It is interesting to

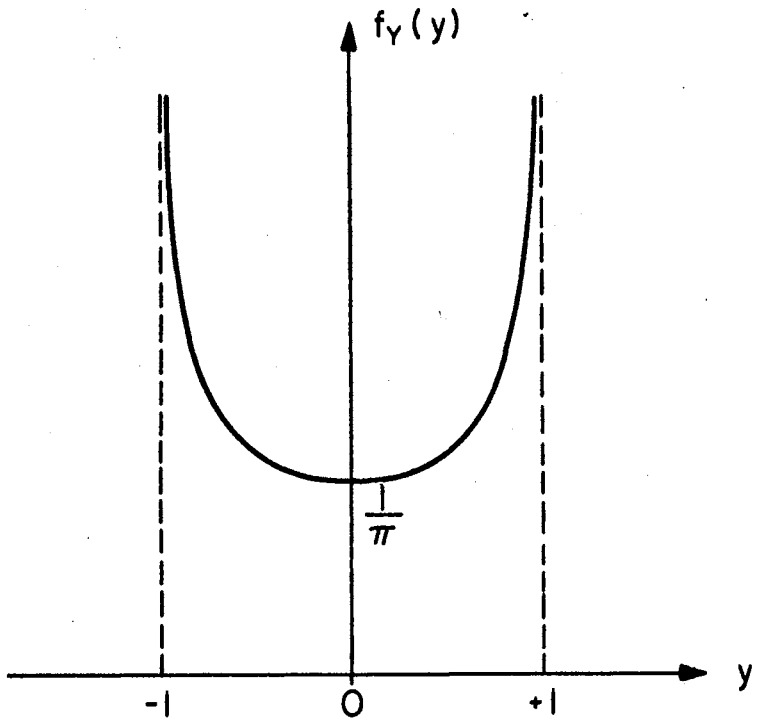


Figure 47. Density Function of $Y(t) = \cos(\omega_0 t + \phi)$

observe that, for this particular distribution of X(t), Y(t) has no influence on the distribution of W(t) in the intervals (-9,-1) and (1,9); however, in the intervals (-11,-9) and (9,11), Y(t) flattens and spreads the mass of X(t). The greatest effect of Y(t) is exerted about the origin. Here, Y(t) flattens the impulse and spreads the mass toward the amplitude extremum of the sinusoid. Hence, it appears that the major effect of the sinusoid is to subvert small signal resolution. For example, if one were to attempt to quantize the probability mass of X(t) in the interval (-.5,15), using the density of W(t), the result would be considerably lower than the true value. Correspondingly, the predicted density in the intervals (-1.5, -.5) and (.5,1.5) would be

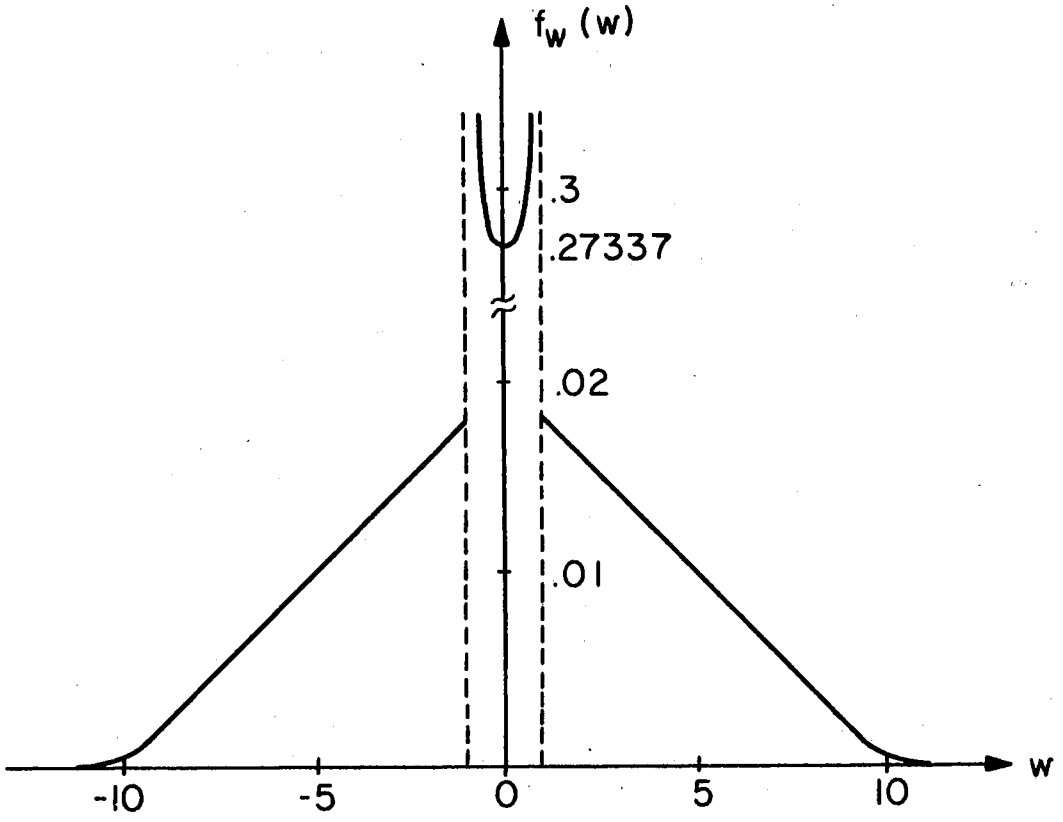


Figure 48. Density Function of W(t) = X(t) + Y(t)

considerably higher than the true value.

Since the study of the text is primarily concerned with the estimation of second-order statistics through the use of histograms, a comparative analysis will be made of the variance estimates for different bar widths of the histogram. From this investigation, it is hoped that some conclusions can be drawn concerning the selection of optimal bar widths. To conduct the comparative analysis of the variance estimates, the strategy will be to, first, compute the probability distribution function of W(t). Secondly, the domain of the distribution function will be subdivided in accordance with assumed bar widths. Third, the

values of the distribution at the division points will be subtracted to determine the probability that W(t) falls in any one particular bar. Finally, the variance will be computed using

$$\sigma^2 = \sum_i \omega_i^2 \, P\{W(t) = \omega_i\}$$

where ω_i is the midpoint of the bar.

Computation of the probability distribution function of W(t) results in the following expression,

$$f(\omega) = \begin{cases} 0 & \omega < -11 \\ \frac{1}{\pi 500}\{\frac{1}{2}[(\omega+10)^2 + \frac{1}{2}]\sin^{-1}(\omega+10) + \frac{3}{4}(\omega+10)\sqrt{1-(\omega+10)^2} \\ \quad + \frac{\pi}{4}[\frac{1}{2} + (\omega+10)^2]\} & -11 \leq \omega < -9 \\ \frac{1}{500}[\frac{\omega^2}{2} + 10\omega + 49.5] + \frac{3}{2000} & -9 \leq \omega < -1 \\ -\frac{2}{\pi 500}\{\frac{1}{2}\omega^2\sin^{-1}\omega - \frac{799}{4}\sin^{-1}\omega + \frac{3}{4}\omega\sqrt{1-\omega^2} - 5\pi\omega \\ \quad - \frac{837\pi}{8}\} + \frac{163}{2000} & -1 \leq \omega < 1 \\ \frac{1}{500}[10\omega - \frac{\omega^2}{2} - 9.5] + \frac{1837}{2000} & 1 \leq \omega < 9 \\ \frac{1}{\pi 500}\{\frac{1}{2}[(\omega-10)^2 + \frac{1}{2}]\sin^{-1}(\omega-10) \\ \quad + \frac{3}{4}(\omega-10)\sqrt{1-(\omega-10)^2} - \frac{\pi}{4}[(\omega+10)^2 - \frac{5}{2}]\} + \frac{1997}{2000} & 9 \leq \omega < 11 \\ 1 & 11 \leq \omega. \end{cases}$$

Amplitude window widths of .88, 2.0 and 4.4 were used to compute an estimate of the variance of X(t). The results of these computations

are shown in Table I. Reviewing the data contained in Table I, for twenty-five windows of width .88, the estimate of variance is in error by 16.3%. Correspondingly, when the window width is equal to the peak amplitude excursion of the sinusoid, the error in the estimate is 5%; however, if the window width is made larger, say 4.4, the error is 13.3%.

TABLE I

COMPARISON OF VARIANCE ESTIMATES AS A FUNCTION OF HISTOGRAM BAR WIDTH

Parameter	X(t)	.88	2.0	4.4
Variance	3.333	3.8862	3.5000	3.7760
Error	-----	16.3%	5.0%	13.3%

The comparison of the estimates of variance, then, lead to the conclusion that the optimal window width, to be used in minimizing the error of the variance estimate, is equal to the peak amplitude excursion of the sinusoid.

B.3 Considerations in Setting the Minimum Amplitude Window Width. As stated earlier, the objective in setting the minimum amplitude window width is to minimize the effect of the sinusoid on the autocorrelation and power spectral density estimates. It seems reasonable that the level should be set sufficiently high to mask the sinusoid during

periods of inactivity but not so high as to mask a significant portion of the data during active periods. This would imply that the minimum width should be set at the peak value of the sinusoid. The study just conducted leads one to the same conclusion.

A final consideration involved here is one developed in Appendix A; namely, the greater the amplitude window width, the greater will be the bias in the estimate of the density function. Thus, one would want to choose as small an amplitude window as possible.

This discussion, then, culminates in the following conclusion: For the problem of estimating a probability density function from the gathered thunderstorm data, one would want to set the amplitude window width as near the peak amplitude of the sinusoid as possible in order to minimize the error in the estimate of the autocorrelation function.

APPENDIX C

SAMPLE COMPUTER PROGRAM FOR THE DATA ANALYSIS USING THE IBM 360/50

C.1 Introduction. This appendix documents the computer program used to implement the data reduction technique developed in Appendix A on the IBM 360/50. A discussion of how the program carries out the computation is included along with a listing of the program itself.

C.2 Computer Implementation of the Data Reduction Technique. A flow chart of the sample computer program is shown in Figure 49. The function of the main program is to receive sampled thunderstorm data, order the data in a matrix, call subroutines and output the resulting estimates of the time averaged second-order statistics. The main program begins by quantizing the first twenty rows of the data input block. An unlisted READ statement advances the input tape reel to the data segment to be analyzed. Then, the data are read into the computer and given the variable name DATBLK. This variable name denotes a 120 x 100 matrix. The data on the input tape reel are in blocks of 2000 samples, and the program reads the first block into the bottom twenty rows of DATBLK. The main program then calls up subroutine DATCON. This subroutine classifies each data point in the bottom 100 rows of DATBLK according to the predetermined amplitude windows. Each amplitude window has a width of 2.50 with center points at -10, -7.5, -5, -2.5, 0, 5, 7.5 and 10. If the sample point value is in the range -8.75 to -11.25,

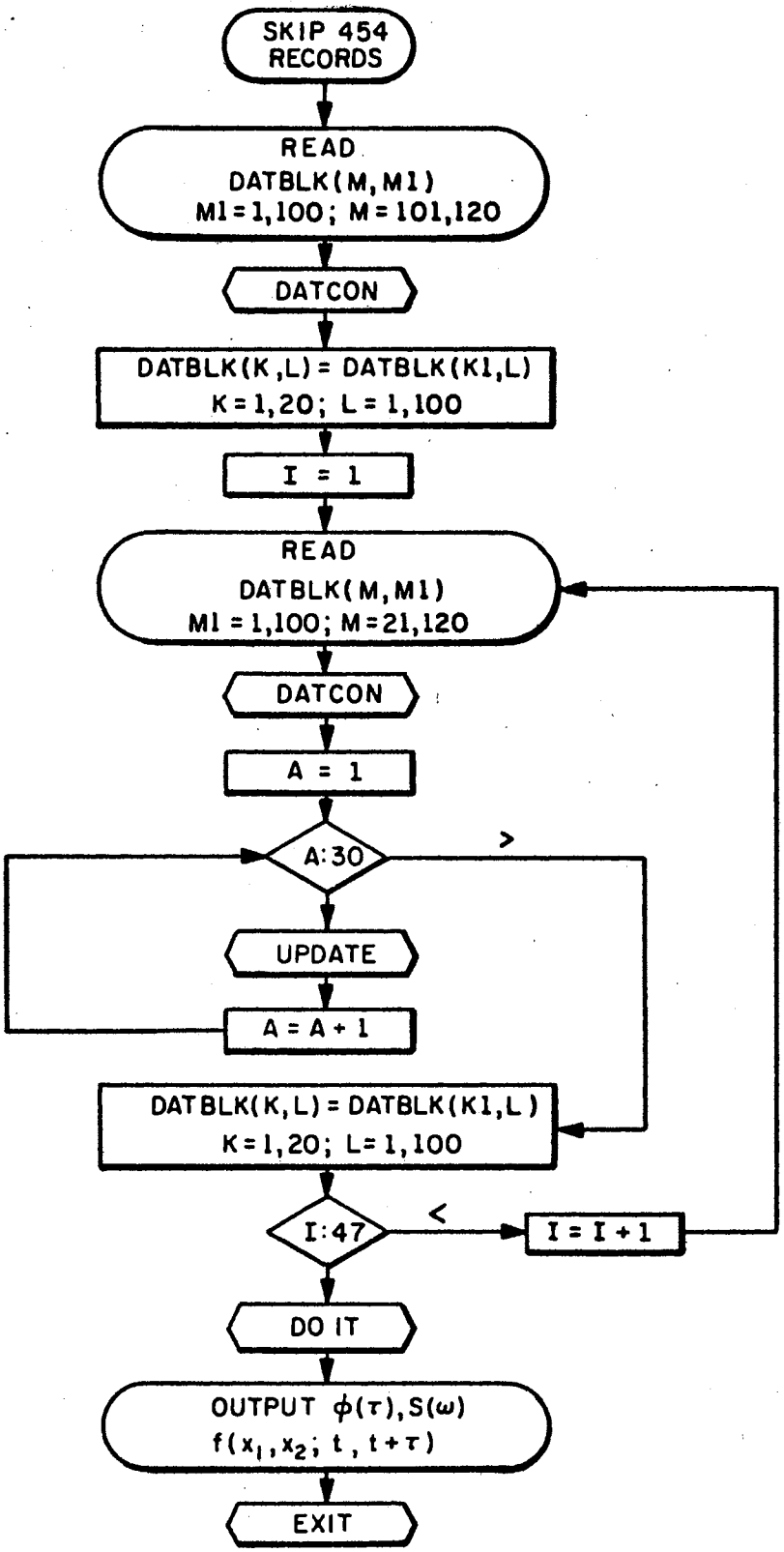


Figure 49. Flow Chart for the Sample Computer Program

then that element of DATBLK is replaced by a value of 1. Similarly, if the sample point value is in the range -6.25 to -8.75, the element of DATBLK is replaced by the number 2. The numbering scheme assigns the values 3, 4, 5, 6, 7, 8 and 9 to the windows centered at -5, -2.5, 0, 2.5, 5, 7.5 and 10, respectively. Hence, this subroutine returns DATBLK to the main program with the bottom 100 rows modified by the classification process. The main program, then, moves the bottom twenty rows to the top twenty rows of the DATBLK matrix.

The first 20 rows of DATBLK are, now, quantized and the iterative data input begins. The loop described in the following is executed 47 times for 3/4 second of thunderstorm data. Since the data blocks must be read in one block at a time, rows 21 through 40 of DATBLK are first read, followed by rows 41-60, 61-80, 81-100 and 101-120. Subroutine DATCON is called, and the input data are classified and returned to the main program. The main program is now ready to initiate the updating of the counting registers. For this particular program, thirty increments of τ were chosen. Thus, to count the number of times X(t) fell in one interval and $X(t + \tau)$ fell in another interval, thirty, 9 x 9 counting registers are required. At this point, the main program enters a loop which is executed thirty times. The loop first assigns a value to the variable A which keeps track of the particular increment in τ. Consider the 24th execution of this loop. Since τ has been chosen to increment in 5 sample steps starting at zero, this implies a shift of 115 samples. The main program sets A = 24 and calls subroutine UPDATE. To tell the subroutine where to find $X(t + \tau)$ in the DATBLK matrix, the variables DELTAT and DELTAF are used. DELTAT tells the subroutine how many columns to advance and DELTAF specifies how many rows to advance.

For this particular case, then, DELTAT(24) = 15 and DELTAF(24) = 1. The counting registers are given the variable name TAU and are 30 x 9 x 9 arrays. The register used on this pass is TAU(24,X,X). To advance the count in this register, DATBLK is treated column by column and row by row over the first 100 rows. For example, let D1 denote the value of an arbitrary element of DATBLK. This implies that X(t) fell in the D1 amplitude window. Let C1 denote the value of the element which is 15 columns to the right and one row down from the D1 element. This implies that X(t + τ) fell in the C1 amplitude window. Thus, for this pair of points, the element TAU(24,C1,D1) is advanced by one count. After updating a particular register, the subroutine returns control to the main program. The main program, now, shifts the bottom 20 rows of DATBLK to the top 20 rows, and the loop begins again by reading 10,000 more sample points from the tape reel.

After the data segment has been read in its entirety, then the main program uses the counting registers to generate estimates of the time averaged second-order probability density function, the autocorrelation function and the power spectral density. For this purpose subroutine DOIT is called. A discrete estimate of the probability density function is formed by dividing each element of TAU by the total number of sample points. The resulting value is $< \alpha_0(t)\alpha_1(t + \tau) >_T$ in Appendix A. The autocorrelation estimate is formed using

$$\phi_{xx}(\tau) = \sum_{x_2} \sum_{x_1} x_1 x_2 < \alpha_0(t)\alpha_1(t + \tau) >_T$$

where the x_i's are the center points of the amplitude windows. The estimated autocorrelation function is modified by the Hanning lag

window, shifted in time and the Fourier transform taken using the Fast Fourier Transform library subroutine. This, then, is the power spectral density estimate.

After these estimates have been formed, control returns to the main program which outputs the estimates on the teleprinter. The sample computer program is listed in Table II.

TABLE II

THE DATA ANALYSIS PROGRAM

```
//RLJ6   JOB (10174,450-54-7505,05),'R.L.JOHNSON',MSGLEVEL=1
// EXEC FORTGCLG
//FORT.SYSIN DD *
C      THIS IS A PROGRAM TO ESTIMATE THE TIME AVERAGED SECOND-ORDER
C      PROBABILITY DENSITY FUNCTION, AUTOCORRELATION AND POWER SPECTRAL
C      DENSITY FUNCTIONS OF THUNDERSTORM NOISE DATA RECORDS CONSISTING
C      OF 480,000 SAMPLE POINTS
C
C      SYSTEM INITIALIZATION
       INTEGER A,DELTAT,DELTAF
       COMMON/SET1/A,DATBLK(120,100),DELTAT(30),DELTAF(30),NOSAMP(30),
      1TAU(30,9,9),R(30),RN(30),S(65)
C
C      INITIALIZATION OF THE DATA INPUT BLOCK
       DO 2 K=1,454
       READ (1)
2      CONTINUE
       READ (1) ((DATBLK(M,M1),M1=1,100),M=101,120)
       CALL DATCON
       DO 5 K=1,20
       K1=K+100
       DO 5 L=1,100
       DATBLK(K,L)=DATBLK(K1,L)
5      CONTINUE
C
C      ACTUAL START OF THE ITERATIVE DATA INPUT
       DO 55 I=1,47
       II=21
       IJ=40
       DO 10 K=1,5
       READ (1) ((DATBLK(M,M1),M1=1,100),M=II,IJ)
       II=II+20
       IJ=IJ+20
10     CONTINUE
       CALL DATCON
       A=1
15     IF (A.GT.30) GO TO 45
       CALL UPDATE
40     A=A+1
       GO TO 15
45     DO 55 K=1,20
       K1=K+100
       DO 55 L=1,100
       DATBLK(K,L)=DATBLK(K1,L)
55     CONTINUE
       CALL DOIT
60     FORMAT (1H1)
       WRITE (6,60)
70     FORMAT (1HO,'AUTOCORRELATION',5X,'NORMALIZED R',10X,
      1'POWER SPECTRUM')
       WRITE (6,70)
       DO 80 I=1,30
75     FORMAT (E14.5,5X,E14.5,4X,E14.5,2X,I3,1X,'KHZ')
       K=I-1
80     WRITE (6,75) R(I),RN(I),S(I),K
       X=0.0
       DO 82 I=31,65
       K=I-1
82     WRITE (6,75) X,X,S(I),K
```

TABLE II (Continued)

```
      WRITE (6,60)
      K2=4
      DO 95 K=1,30
83    FORMAT (11X,'TIME AVERAGED SECOND-ORDER PROBABILITY DENSITY ESTIMA
     1TE FOR A LAG OF',F14.9,1X,'MICROSECONDS')
      K1=K-1
      TS=(50.0/32.0)*5.0*K1
84    FORMAT (1HO)
      IF (K.LE.K2) GO TO 100
      WRITE (6,60)
      K2=K2+4
100   WRITE (6,83) TS
      DO 90 KK=1,9
85    FORMAT (9F13.7)
90    WRITE (6,85) (TAU(K,KK,KKK),KKK=1,9)
      WRITE (6,84)
      WRITE (6,84)
95    CONTINUE
      CALL EXIT
      END
C
      BLOCK DATA
      COMMON/SET1/A,DATBLK(120,100),DELTAT(30),DELTAF(30),NOSAMP(30),
     1TAU(30,9,9),R(30),RN(30),S(65)
      INTEGER A,DELTAT,DELTAF
      DATA DATBLK/12000*0.0/,DELTAT/0,5,10,15,20,25,30,35,40,45,50,55,
     160,65,70,75,80,85,90,95,0,5,10,15,20,25,30,35,40,45/,DELTAF/0,0,
     10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1/,
     1NOSAMP/30*0/,R/30*0.0/,TAU/2430*0.0/
      END
C
C     DATA CONVERSION SUBROUTINE
      SUBROUTINE DATCON
      COMMON/SET1/A,DATBLK(120,100),DELTAT(30),DELTAF(30),NOSAMP(30),
     1TAU(30,9,9),R(30),RN(30),S(65)
      INTEGER A,DELTAT,DELTAF
      DO 275 K=21,120
      DO 275 L=1,100
      IF (DATBLK(K,L))215,220,225
215   Y=0.0-DATBLK(K,L)
      W=0
      GO TO 230
225   Y=DATBLK(K,L)
      W=1
230   IF (Y.GT.125) GO TO 232
220   DATBLK(K,L)=5
      GO TO 275
232   IF (ABS(Y-250).LE.125) GO TO 235
      IF (ABS(Y-500).LE.125) GO TO 240
      IF (ABS(Y-750).LE.125) GO TO 245
      IF (ABS(Y-1000).LE.125) GO TO 250
235   DATBLK(K,L)=4+2*W
      GO TO 275
240   DATBLK(K,L)=3+4*W
      GO TO 275
245   DATBLK(K,L)=2+6*W
      GO TO 275
250   DATBLK(K,L)=1+8*W
275   CONTINUE
```

TABLE II (Continued)

```
      RETURN
      END
C
C     THIS SUBROUTINE UPDATES THE COUNTING REGISTERS
      SUBROUTINE UPDATE
      COMMON/SET1/A,DATBLK(120,100),DELTAT(30),DELTAF(30),NOSAMP(30),
     1TAU(30,9,9),R(30),RN(30),S(65)
      INTEGER A,DELTAT,DELTAF
      INTEGER C1,D1
      DO 415 M=1,100
      DO 415 N=1,100
      MO=M+DELTAF(A)
      NO=N+DELTAT(A)
      IF (NO.GT.100) GO TO 410
401   C1=DATBLK(MO,NO)
      D1=DATBLK(M,N)
      TAU(A,C1,D1)=TAU(A,C1,D1)+1
      GO TO 415
410   MO=MO+1
      NO=NO-100
      GO TO 401
415   CONTINUE
      NOSAMP(A)=NOSAMP(A)+1
      RETURN
      END
C
C     THIS SUBROUTINE COMPUTES THE ESTIMATE OF THE TIME AVERAGED
C     SECOND-ORDER STATISTICS
      SUBROUTINE DOIT
      DIMENSION DATA(1024),TRAN(2,513),WORK(2,1024),N(1),RO(30)
      COMMON/SET1/A,DATBLK(120,100),DELTAT(30),DELTAF(30),NOSAMP(30),
     1TAU(30,9,9),R(30),RN(30),S(65)
      EQUIVALENCE (DATA,TRAN)
      INTEGER B,C,Z
      DATA N(1)/1024/
C
C     ESTIMATE THE TIME AVERAGED SECOND-ORDER DENSITY FUNCTION
      DO 500 Z=1,30
      DO 500 B=1,9
      DO 500 C=1,9
      TAU(Z,B,C)=TAU(Z,B,C)/(NOSAMP(Z)*10000)
500   CONTINUE
C
C     ESTIMATE THE TIME AVERAGED AUTOCORRELATION FUNCTION
      DO 511 Z=1,30
      E=-12.5
      R(Z)=0.0
      DO 511 B=1,9
      E=E+2.5
      D=-12.5
      DO 511 C=1,9
      D=D+2.5
      EXP=E*D*TAU(Z,B,C)
      R(Z)=R(Z)+EXP
511   CONTINUE
C
C     MODIFY THE AUTOCORRELATION ESTIMATE BY THE HANNING LAG WINDOW,
C     PERFORM A TIME SHIFT AND ESTIMATE THE POWER SPECTRAL DENSITY
C     FUNCTION
```

TABLE II (Continued)

```
      DO 515 Z=2,30
515   RN(Z)=R(Z)/R(1)
      RN(1)=1.0
      DO 520 Z=1,30
      RA=3.1415926536*(Z-1)/29.0
520   RO(Z)=(1.0+COS(RA))*0.5*RN(Z)
      DATA(512)=1.0
      DO 525 I=2,30
      J2=I+511
      J3=513-I
      CON=RO(I)
      DATA(J2)=CON
525   DATA(J3)=CON
      DO 530 I=1,482
530   DATA(I)=0.0
      DO 535 I=542,1024
535   DATA(I)=0.0
      CALL FOURT(DATA,N,1,-1,0,WORK,1024)
      J3=1
      DO 545 J2=1,65
      S(J2)=SQRT(TRAN(1,J3)**2+TRAN(2,J3)**2)*(.250/32000.0)
545   J3=J3+8
      RETURN
      END
//GO.FT01F001 DD UNIT=(TAPE9,,DEFER),
//             VOL=SER=RJOH06,DSNAME=WBDAFL10,
//             DISP=OLD,DCB=(RECFM=V,BLKSIZE=8008,LRECL=8004)
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
//
```

APPENDIX D

CIRCUIT DIAGRAMS OF THE DATA GATHERING ELECTRONICS

D.1 Introduction. The purpose of this appendix is to document the circuit diagrams of the data gathering electronics referred to in Chapter II. Included with the diagrams is a generalized discussion of the considerations involved in the design of the electronics and selection of the instrumentation.

D.2 Details of the Measurement Instrumentation. The vertically polarized electromagnetic field is sensed through the use of a five foot vertical whip antenna mounted atop the fuselage of a D-18 twin engine Beechcraft airplane. The signal at the antenna terminals is then fed through the preamp/cathode follower circuit shown in Figure 50. The first stage VI, is a cathode follower acting as a buffer between the antenna and preamplifier stage, V2. Owing to the low input impedance of the commercially developed broadband amplifier shown in Figure 51, a second cathode follower stage, V3, is included between the preamplifier and the main amplifier. The main amplifier consists of three amplifier stages cascaded on an integrated circuit chip. The amplified output signal is fed through two emitter follower stages and then recorded on magnetic tape. The motivation for two emitter follower stages is to enhance the versatility of the circuit board, allowing for example, a detector stage to be inserted between Q1 and Q2, if desired. Commenting on the broader aspects of the measurement electronics, it should be

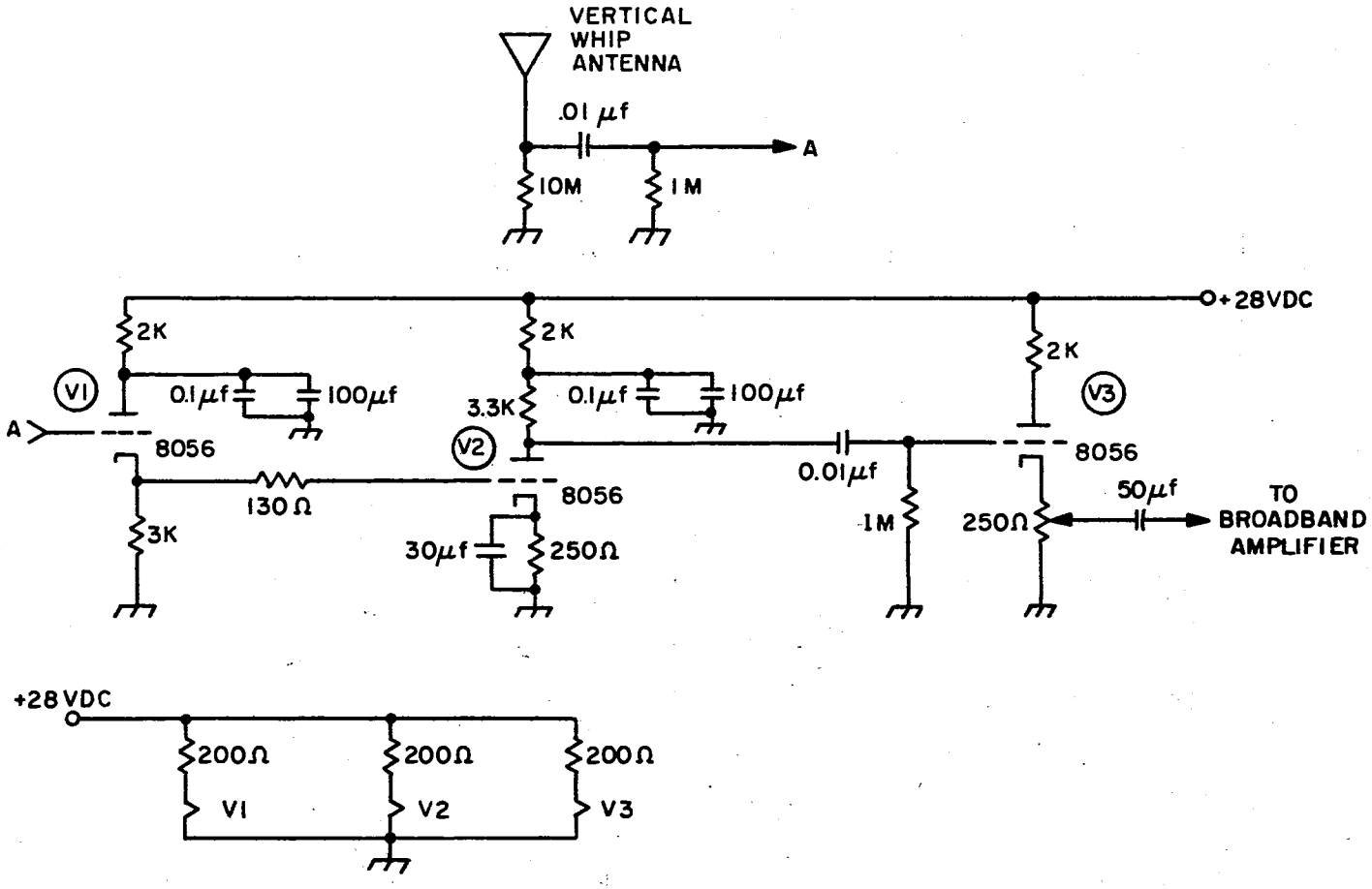


Figure 50. Circuit Diagram of the Preamplifier

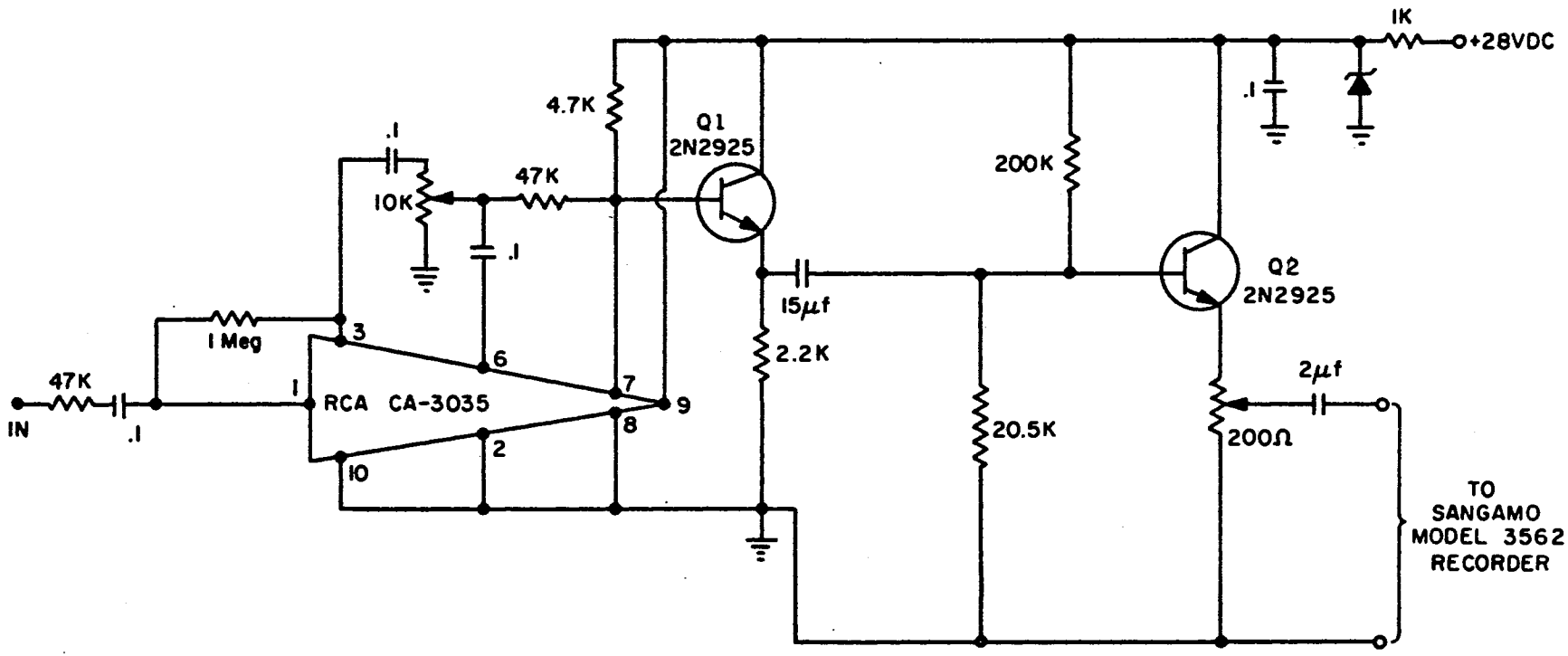


Figure 51. Circuit Diagram of the Main Amplifier

pointed out that because the instrumentation is subject to continuing vibration while measurements are being made, solid state devices were utilized as opposed to vacuum tubes in an effort to reduce internally generated noise and to conserve space and weight. This approach was extended to the selection of commercial equipment as well as to the electronics designed and fabricated locally. One exception is the use of nuvistors in the preamp/cathode follower circuit owing to their high gain and low noise figure.

The saturation effects observed in the analog data and described in Chapter V, are attributed to saturation of the preamplifier. A review of the analog data leads to the belief that this effect is caused by near return strokes. Treating the antenna as a voltage source and assuming that it can be modeled as a low impedance source, compared with the 10 megohm input of the measurement system, it appears that the .01 μfarad capacitor and the 1 megohm resistor compose a differentiator circuit. This network, then, plays a large role in determining the recovery time of the preamplifier.

VITA

Richard Leon Johnson

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STUDY OF THE SECOND-ORDER STATISTICAL PROPERTIES OF THUNDERSTORM NOISE

Major Field: Electrical Engineering

Biographical:

Personal Data: Born on June 12, 1938, in Enid, Oklahoma, the son of Clarence W. and Wilma A. Johnson.

Education: Graduated from Arlington High School, Arlington, Texas in May, 1956; received the Bachelor of Science degree in Electrical Engineering from the University of Texas at Arlington in August, 1963; received the Master of Science degree in Electrical Engineering from Southern Methodist University in August, 1966; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1971.

Professional Experience: U. S. Air Force from June, 1956 to March, 1960; employed by Southwest States Telephone Company during Summer, 1960; employed by Mountain States Telephone and Telegraph Co. for September, 1963 to May, 1964; Aerosystems Engineer, General Dynamics Fort Worth Division from May, 1964 to August, 1966; Graduate Research Assistant, School of Electrical Engineering, Oklahoma State University from September, 1966 to October, 1970.

Professional Organizations: Member of Eta Kappa Nu, Tau Beta Pi, and the Institute of Electrical and Electronic Engineers.