THE SQUARE WAVE TRANSFORMATION OF POINT PROCESSES
A STUDY OF THE SQUARE WAVE TRANSFORMATION
OF
POINT PROCESSES

BY
PAUL DOUGLAS DOWLING, B.Sc.

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Science

McMaster University
(October) 1967
ACKNOWLEDGMENTS

The research reported in this thesis was done while the author was a graduate student in mathematics at McMaster University. It was directed by Professor I. Z. Chorneyko of the Mathematics Department.

The author acknowledges his indebtedness to Professor Chorneyko for suggesting this problem and for his invaluable advice, criticism and guidance throughout the entire period in which this research was carried out.

Finally a vote of thanks to my mother who patiently typed pages of equations and symbols Greek to her, and to Mr. E. Funke of the National Research Council for helpful conversations regarding instrumentation.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2) DEFINITIONS AND NOTATIONS</td>
<td>4</td>
</tr>
<tr>
<td>3) PROBABILITY ANALYSIS</td>
<td>8</td>
</tr>
<tr>
<td>4) SPECTRAL ANALYSIS</td>
<td>23</td>
</tr>
<tr>
<td>5) COMPUTATIONAL RESULTS AND CONCLUSIONS</td>
<td>32</td>
</tr>
<tr>
<td>APPENDIX A RESULTS OF PROBABILITY ANALYSIS</td>
<td>36</td>
</tr>
<tr>
<td>APPENDIX B RESULTS OF SPECTRAL ANALYSIS</td>
<td>61</td>
</tr>
<tr>
<td>APPENDIX C COMPUTER PROGRAMS</td>
<td>89</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

A process consisting of a series of events occurring in continuous time when interest is concentrated on the individual occurrences of the events themselves constitutes a point process. The emissions from a radioactive source or accidents occurring in time are examples of a series of point events, with the events being distinguished only by their positions in time.

In this thesis stationary processes, in which the time origin is an arbitrary point, will be dealt with, and for simplicity it is assumed that there is zero probability that two or more events occur simultaneously. There are basically two ways of looking at point processes; in terms of the number of events occurring in fixed time intervals or in terms of intervals between events.

The purpose of this thesis is to study various methods of analysing point processes by means of the square wave transformation. Figure 1-1 and Figure 1-2 are examples of a point process and the square wave transformation of it.
A point process can easily be transformed into a square wave by means of a flip-flop device at each pulse. The resulting square wave can be fed into an analyser to calculate the auto-covariances and spectral density function. The ease of instrumentation for calculation of information from the square wave makes it very practical to work with.

In chapter (2) of this thesis the basic definitions and notations are described. Probability analysis of point processes is dealt with in chapter (3) and spectral analysis in chapter (4). The square wave transformation is introduced in chapter (4) and is analysed using spectral analysis in order to extract the statistical properties of the original point process. In chapter (5) the computational results and conclusions from utilizing the methods discussed in chapters
(3) and (4) are summarized. Appendix A gives a detailed account of the experimental time series employed and the computational results from using the methods in chapter (3). Appendix B consists of the corresponding spectral analysis of the different point processes. All computations were done on the I.B.M. 7040 computer and the FORTRAN IV programs used in the time series analysis are exhibited in Appendix C.
A stochastic process is defined as a collection \( \{X(t), t \in T\} \) of random variables. The set \( T \) is called the **index set** of the process. No restriction is placed on the nature of \( T \). The two important cases are when \( T = \{0, \pm 1, \pm 2, \ldots\} \) or \( T = \{0, 1, 2, \ldots\} \) in which case the stochastic process is said to be a **discrete parameter process** or when \( T = \{t: -\infty < t < \infty\} \) or \( T = \{t: t \geq 0\} \), the **continuous parameter process**.

A special case of a stochastic process is the point process which can be defined in the following manner. Let a sequence of events occur at the instants \( t_i \) where \( i \geq 1 \), \( t_{i+1} > t_i \) and \( t_1 > 0 \). Then, this time sequence will be denoted by \( X(t), t \in \{t_1, t_2, \ldots\} \), a discrete parameter process, so that if \( X(t) \) is a stochastic process which describes the time of occurrence of events which are considered to occur for an infinitesimal duration, then \( X(t) \) is a **point process**. For example, consider events occurring periodically in time with period \( B \). Then \( X(t), t \in \{t_1, t_2, \ldots \mid t_i = i \cdot B\} \) is a point process.

In the analysis of any process only a finite record of observations is available, so that the index set of the process can be considered as the time interval \((0, T]\). The **truncated**
**sample function** is defined by $X_T(t)$ as

$$X_T(t) = \begin{cases} X(t) & 0 < t \leq T \\ 0 & \text{elsewhere} \end{cases}$$

Let $n$ be the number of events in the observation period $T$, and let $\tau$ be any fixed time interval length. Then $n(t)$ is the number of events occurring in the time interval $(0, t]$ and $n_i(\tau)$ is the number of events in the interval $((i-1)\tau, i\tau]$.

Unless otherwise stated, the time series is assumed to be stationary. This means that all statistical properties depend upon differences $X_T(t_i) - X_T(t_j)$ rather than on the time points $X_T(t_i)$ and $X_T(t_j)$ themselves. Let $A_1, A_2, \ldots$ be arbitrary sets on the real axis, and $T_hA_1, T_hA_2, \ldots$ be the sets obtained by translating through $h$. Let $N(A)$ be the number of events in $A$. The point process is stationary if the two sets of random variables

$$N(A_1), N(A_2), \ldots, N(A_k); \quad N(T_hA_1), N(T_hA_2), \ldots, N(T_hA_k)$$

have the same joint distribution for all initial sets $A_1, A_2, \ldots A_k$ and all real $h$ and all $k=1, 2, \ldots$

The time series $X_T(t)$ has the **sample mean value function**

$$\mu = \frac{1}{T} \int_0^T X_T(t) \, dt$$

$$= \frac{1}{T} \sum_{t=1}^T X_T(t)$$
for the continuous and discrete value case respectively.

The sample autocovariance function \( R_T(k) \) is defined by

\[
R_T(k) = \frac{1}{T} \int_0^T [X_T(t)-\mu] \cdot [X_T(t+k)-\mu] \, dt \quad |k| < T
\]

\[
= 0 \quad |k| \geq T
\]

and

\[
R_T(k) = \frac{1}{T} \sum_{t=1}^{T-|k|} [X_T(t)-\mu] \cdot [X_T(t+k)-\mu] \quad k=0, \pm 1, \ldots, \pm (T-1)
\]

\[
= 0 \quad k=\pm T, \pm (T+1), \ldots
\]

The sample spectral density function is defined by

\[
\int_{-T}^{T} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\omega} R_T(k) \, dk \right| = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos k\omega R_T(k) \, dk
\]

or

\[
\int_{-T}^{T} \omega = \frac{1}{2\pi} \sum_{k=-T}^{T} e^{-ik\omega} \frac{1}{2\pi} R_T(0) + \frac{1}{\pi} \sum_{k=1}^{T} \cos k\omega R_T(k)
\]

for the continuous and discrete parameter cases respectively.

Let \( \lambda(t) \), the number of events per unit time, be the (probability) rate of occurrence of an event. The Poisson process is an important stochastic process which serves as a mathematical model for empirical phenomena like the arrival of
calls at a telephone exchange, the emission of particles from a radioactive source, and the occurrence of serious coal-mining accidents. Consider point events occurring singly in time with the rate of occurrence $\lambda(t) = \lambda$, a constant. If $N(t,t+\Delta t)$ is the number of events in the interval $(t,t+\Delta t]$ then assume that, as $\Delta t \to 0^+$

$$
\begin{align*}
\text{prob} \{ N(t,t+\Delta t) = 0 \} &= 1 - \lambda \Delta t + O(\Delta t), \\
\text{prob} \{ N(t,t+\Delta t) = 1 \} &= \lambda \Delta t + O(\Delta t),
\end{align*}
$$

so that $\text{prob} \{ N(t,t+\Delta t) > 1 \} = O(\Delta t)$, where $O(\Delta t)$ denotes a function tending to zero more rapidly than $\Delta t$. Also assume that $N(t,t+\Delta t)$ is independent of occurrences in $(0,T]$. A stochastic process of point events satisfying these conditions is called a Poisson process of rate $\lambda$. 
In this chapter various methods of analyzing the intervals between events and the number of events in fixed time intervals are considered in order to determine if the point process is random, or generated by some probability mechanism or follows some pattern which can be determined from a sample of the process.

**GRAPHICAL METHODS**

1. The intervals between successive events $x_i$ are plotted as a function of $i$ or against the time at the midpoint of the interval. Trends in the interval length will be indicated. This is equivalent to plotting the time interval between pulses of the square wave. Fig. 3-1

2. If the intervals between events are independent of time then $x_i$ versus $i$ will form a scatter diagram. Then a frequency polygon or histogram can be formed to determine the probability distribution (if it exists) of the interval lengths. Fix a time interval length $x$ such that $mx = \max x_i$, where it is suggested that $12 \leq m \leq 25$, depending upon the sample size $n$. 
Then count the number \( n_j(x) \) of interval lengths \( x_i \) that belong in the interval \([(j-1)x,jx]\), \( j=1,\ldots,m \). Plot \( n_j(x) \) versus \( j \) for a frequency polygon or form a bar graph of \( n_j(x) \) as base for a histogram. The histogram approximates the density function for large \( n \). Fig. 3-2

**RANDOM SERIES**

If the series of events is random, that is the times of occurrences are independently and identically distributed from the uniform distribution, then it is called a Poisson process since the number of events in a fixed time length has the Poisson distribution. The Poisson process when graphed reveals the following,

1. \( x_i \) versus \( i \) will be a scatter diagram showing that \( x_i \) is independent of \( i \),

2. the histogram formed for the interval lengths between events will approximate an exponential density function.

This is proved as follows. Assume that \( X(t) \) is a Poisson process with rate of occurrence \( \lambda \). Take a new time origin at \( t_0 \).

If \( t_0 + Z \) is the time of the first event after \( t_0 \), the random variable \( Z \) is independent of whether an event occurs at \( t_0 \) and of occurrences before \( t_0 \).

Let \( P(x) = \text{prob}\{Z>x\} \). To determine the distribution of \( Z \), let

\[
P(x+\Delta x) = \text{prob}\{Z>x+\Delta x\} \quad \Delta x > 0
\]

\[
= \text{prob}\{Z>x\} \quad \text{and no event occurs in } (t_0+x, t_0+x+\Delta x]
\]

\[
= \text{prob}\{Z>x\} \text{prob}\{\text{no event occurs in } (t_0+x,t_0+x+\Delta x]\mid Z>x\}
\]
Fig. 3-1  Interval lengths $x_i$ versus $i$ for a point process $X(t)$

Fig. 3-2  Histogram of interval lengths of a point process $X(t)$
\[ P(x+\Delta x) = P(x)\{1-\lambda \Delta x + o(\Delta x)\} \]
\[ = P(x) - \lambda P(x) \Delta x + P(x) o(\Delta x) \]

Then
\[ \lim_{\Delta x \to 0} \frac{P(x+\Delta x) - P(x)}{\Delta x} = P'(x) = -\lambda P(x) \]
\[ \frac{dP(x)}{P(x)} = -\lambda dx \]
\[ \log_e P(x) = -\lambda x + k \]
\[ P(x) = Ke^{-\lambda x} \]

Since \( P(0) = \text{prob} \{Z > 0\} = 1 \), \( P(x) = e^{-\lambda x} \)

The distribution function of \( Z \) is \( 1 - e^{-\lambda x} \), the probability density function is \( \frac{d}{dx} (1-e^{-\lambda x}) = \lambda e^{-\lambda x} \quad (x > 0) \)

**PROPERTIES OF RANDOM PROCESSES**

If the time series is random then inferences can be made about \( \lambda \). The intervals between successive events have the exponential distribution with probability density function \( f(x) = \lambda e^{-\lambda x} \), and rate of occurrence \( \lambda \).

Let \( x_1, x_2, \ldots, x_m \) be \( m \) intervals from this distribution. Let \( s = \sum_{i=1}^{m} x_i \) and \( \bar{x} = s/m \). Then \( s \) has the probability density function

\[ f(s) = \frac{\lambda (\lambda s)^{m-1}e^{-\lambda s}}{(m-1)!} \quad s > 0 \]

a gamma distribution, and \( \bar{x} \) is distributed as
\[ g(\overline{x}) = \frac{m\lambda (m\lambda \overline{x})^{m-1} e^{-m\lambda \overline{x}}}{(m-1)!} \quad \overline{x} > 0 \]

These results are verified by finding the moment generating function of the random variables \( s \) and \( x_i \).

\( M_s(t) = E(e^{ts}) \) is the moment generating function of the random variable \( s \).

\( M_s(t) = \sum_i x_i(t) = \prod_i M_{x_i}(t) \) since \( s \) is the sum of \( m \) independent and identically distributed random variables \( x_i \).

Now

\[ M_{x}(t) = E(e^{xt}) = \int_0^\infty e^{xt} e^{-\lambda x} dx \]

\[ = \frac{\lambda}{\lambda - t} \int_0^\infty (\lambda - t) e^{-(\lambda - t)x} dx \]

\[ = \frac{\lambda}{\lambda - t} \]

Then

\[ M_s(t) = \prod_i M_{x_i}(t) = \left( \frac{\lambda}{\lambda - t} \right)^m \]

Now, if two random variables have the same moment generating function then these random variables have the same probability distribution (by the uniqueness theorem for moment generating functions).

The random variable \( y \) with the distribution
\[ h(y) = \begin{cases} \frac{1}{\Gamma(\alpha) B^\alpha} y^{\alpha-1} e^{-y/B} & 0 < y < \infty \\ 0 & \text{elsewhere} \end{cases} \]

where \( \alpha > 0, B > 0, \Gamma(\alpha) = (\alpha-1)! \) has the gamma distribution, \( M_Y(t) = (1-Bt)^{-\alpha}, t < 1/B \).

Now \( M_s(t) = \left( \frac{1}{1-t/\lambda} \right)^m = (1-t/\lambda)^{-m} \)

and is the moment generating function of a random variable with the gamma distribution where \( \alpha = m, B = 1/\lambda \).

Then
\[
\mathcal{G}(s) = \frac{1}{(m-1)! (1/\lambda)^m} s^{m-1} e^{-\lambda s} \quad 0 < s < \infty
\]
\[
= \frac{\lambda (\lambda s)^{m-1} e^{-\lambda s}}{(m-1)!} \quad 0 < s < \infty
\]

\( M_\mathcal{X}(t) = M_s(t) = M_s(t/m) \) is the moment generating function for the random variable \( \mathcal{X} \).

\( M_s(t/m) = (1-t/m\lambda)^{-m} \), then \( \alpha = m, B = 1/m\lambda \)

and
\[
g(\mathcal{X}) = \frac{m\lambda (m\lambda \mathcal{X})^{m-1} e^{-m\lambda \mathcal{X}}}{(m-1)!} \quad \mathcal{X} > 0
\]

The following properties can be established:

1. \( g(\mathcal{X}) \) has mean \( 1/\lambda \) and variance \( 1/(m\lambda^2) \),
2. as \( m \) increases \( g(\mathcal{X}) \) becomes normally distributed,
3. \( \mathcal{X} \) is a sufficient estimator for \( 1/\lambda \),
(4) $2\lambda \bar{X}$ is distributed as $\chi^2$ with 2 degrees of freedom,
(5) $2m\lambda \bar{X}$ is distributed as $\chi^2$ with $2m$ degrees of freedom.

Using this information a confidence interval for $\lambda$ with probability $1-2\alpha$ is

$$\frac{\chi^2_{1-\alpha/2m}}{2m\bar{X}} < \lambda < \frac{\chi^2_{\alpha/2m}}{2m\bar{X}}$$

If there are two random series of events with $m_1$ and $m_2$ intervals and means $\bar{X}_1$ and $\bar{X}_2$ respectively, then the hypothesis that $\lambda_1 = \lambda_2$ can be tested by the $F$ distribution with $2m_1$ and $2m_2$ degrees of freedom. Then $F = \frac{\lambda_1 \bar{X}_1}{(\lambda_2 \bar{X}_2)}$ with $2m_1$ and $2m_2$ degrees of freedom. If there are more than two rates of occurrence then the $F$ test can not be used to test the homogeneity of the $\lambda_i$. A special application of the $\chi^2$ test, known as Bartlett's test, may be applied.

Considerable computation can be saved by applying an $F$ test to the largest and smallest variance before Bartlett's test. If the $F$ test indicates that the largest variance is not significantly different from the smallest one, then it is reasonable to assume that the variances lying in between do not differ significantly from the smallest one. The application of the $\chi^2$ and $F$ tests to the analysis of the time intervals depends on the assumption of homogeneity ($\lambda$ is not a function of time). There are several tests of homogeneity, one of which is the $g$ test, another is the previously mentioned Bartlett's test that all $\lambda_i$ are equal. These tests, which are
based on interval lengths, can be used as tests of randomness. 

TESTS OF RANDOMNESS

(1) \textit{g test}

Let \( x_m \) be the largest among \( m \) independent intervals and let \( \bar{x} \) be the mean interval length. The statistic \( g = \frac{x_m}{m\bar{x}} \) has a probability relation which has been compiled by Fisher\(^1\). To determine whether the length of an interval is significant under the hypothesis of randomness, a significance level \( \alpha \) is selected and \( g_\alpha \) is computed from tables\(^2\).

\( g \) is calculated and if \( g > g_\alpha \) then the longest interval between events is significant and the series is not random.

Since this test is based on the largest among the \( m \) independent intervals, then it is possible that measurement errors or some other conditions have created an interval length which is an outlier. If the graph of \( x_i \) versus \( i \) indicates an interval length much larger than any other, it is possible that the \( g \) test will reject the hypothesis of randomness if this value is used for \( x_m \) even though the sample is homogeneous. If the sample containing such a large deviation is not representative, or if the occurrence of such a

\begin{itemize}
\end{itemize}
large deviation is unlikely in a sample from the population in question, then it is necessary to reject this observation. However, if large deviations occur in a number of samples the presence of additional factors of intermittent character may be responsible.

(2) **Bartlett's Test**

If the sequence of intervals is divided up into $K$ sets of $m_i$ successive intervals where $v_i^2$ is the estimate of variance from sample $i$, and $\bar{V}^2$ is the estimate of the pooled variance, $\bar{V}^2 = \frac{\sum_{i=1}^{k} m_i v_i^2}{\sum_{i=1}^{k} m_i}$

then Bartlett's test requires the calculation of

$$
\chi^2 = 2.3026 \left\{ \log_{10} \bar{V}^2 \sum_{i=1}^{k} (m_i-1) - \sum_{i=1}^{k} (m_i-1) \log_{10} v_i^2 \right\} / C
$$

where $C = 1 + \left( \sum_{i=1}^{k} \frac{1}{m_i-1} - \frac{1}{m-K} \right) / 3(K-1), \quad m = \sum_{i=1}^{k} m_i$

The distribution of $\chi^2$ is approximately $\chi^2$ with $K-1$ degrees of freedom and the approximation is reasonably accurate if the $m_i-1 \geq 5$, $i=1, 2..., K$.

When all samples are of the same size $m$, then

$$
\chi^2 = 2.3026 \left\{ (m-1)K \log_{10} \bar{V}^2 - (m-1) \sum_{i=1}^{k} \log_{10} v_i^2 \right\} / C
$$

where $C = 1+(K+1)/[3K(m-1)]$

If $\chi^2 < \chi^2$ with $K-1$ degrees of freedom, then the hypothesis of homogeneity is accepted.
If a set of interval lengths are independent and exponentially distributed with parameter \( \lambda \), then the mean and variance of the interval lengths are \( 1/\lambda \) and \( 1/\lambda^2 \), respectively. Divide a sequence of interval lengths up into \( K \) sets of \( m \) successive intervals. Let \( \bar{X}_i \) be the mean interval length of the \( i \)th set, and let the intervals be independent and exponentially distributed with parameter \( \lambda_i \). The variance, \( \sigma_i^2 \), of this set can be estimated by \( \bar{X}_i^2 \), since \( \bar{X}_i \) is a sufficient estimator for \( 1/\lambda_i \). Now, under the hypothesis that all the \( \lambda_i \) are equal, the \( K \) sets will constitute a set of \( Km \) interval lengths which are exponentially distributed with parameter \( \lambda \). The variance of this pooled series, \( \bar{\sigma}^2 \), can be estimated by \( \left( \frac{1}{K} \sum_{i=1}^{K} \bar{X}_i \right)^2 \) since the parameter \( \lambda \) will have \( \frac{1}{K} \sum_{i=1}^{K} \bar{X}_i \) as an estimator of \( 1/\lambda \).

The hypothesis that all the \( \lambda_i \) are equal can be tested by computing \( \chi^2 \) and

\[
\chi^2 = 2.3026 \left[ 2(m-1) K \log \left( \frac{1}{K} \sum_{i=1}^{K} \bar{X}_i \right) - 2(m-1) \sum_{i=1}^{K} \log \bar{X}_i \right] / C
\]

\[
= 2.3026 (2m-2) K \left[ \log \left( \frac{1}{K} \sum_{i=1}^{K} \bar{X}_i \right) - \frac{1}{K} \sum_{i=1}^{K} \log \bar{X}_i \right] / C
\]

where \( C = 1 + (K+1)/[3K(m-1)] \)

DEPARTURE FROM RANDOMNESS

TRENDS AND CYCLES

There are two types of sequences in time. One is a slowly moving function of time which is often called a trend,
and is exemplified by a polynomial of fairly low degree,

\[ s(t) = a_0 + a_1 t + \ldots + a_q t^q, \quad t = t_1, t_2, \ldots, t_n. \]

Another type of sequence is cyclical, such as a finite Fourier series,

\[ s(t) = b_0 + \sum_{i=1}^{q} (b_i \cos \lambda_i t + c_i \sin \lambda_i t), \quad t = t_1, t_2, \ldots, t_n. \]

Trends or cycles in the rate of occurrence or number of occurrences can occur as functions of time and the interval lengths between successive events as a function of the interval index.

Consider a point process, where the time interval between the events occurring at time \( t_i \) and \( t_{i+1} \) is \( x_{i+1} = 2x_i \) and \( x_1 = B, \) a constant. Then the interval lengths of the process are \( x_1 = B, x_2 = 2B, \ldots, x_j = 2^{j-1}B, \ldots. \)

For events occurring at the times \( t_1, t_2, \ldots, t_j, \ldots \) then the time of the \( j \)th event at \( t_j \) is

\[ t_j = \sum_{i=1}^{j} 2^{i-1}B = B(2^j - 1) \]

The \( j \)th event occurs at time \( t_j \) so that \( n(t_j) = j. \)

Then

\[ t_j = B(2^{n(t_j)} - 1) \quad \text{or} \quad n(t) = \lfloor \log_2\{(B+t)/B\} \rfloor \]

where \( \lfloor \log_2\{(B+t)/B\} \rfloor \) denotes the largest integer less than or equal to \( \log_2\{(B+t)/B\}. \)
Then, \( n(t) = g(t) = \lfloor \log_2 \left( \frac{B+t}{B} \right) \rfloor \) and
\[ x_i = h(i) = 2^{i-1}B. \]

If a function \( s(t) \), which is a polynomial in \( t \) of specified degree \( q \), is assumed, then the problem is to estimate the coefficients \( a_0, a_1, \ldots, a_q \) on the basis of observations \( X(t_1), X(t_2), \ldots, X(t_n) \) of the sample series. In the case of the Fourier series the problem is to estimate the coefficients \( b_0, b_i \) and \( c_i, i = 1, 2, \ldots, q. \)

In either case the estimation can be done by the method of least squares, the estimates of the parameters being the values of the constants which minimize \[ \sum_{i=1}^{n} (X(t_i) - s(t_i))^2 \]

**INTERVALS WITH PROBABILITY DISTRIBUTIONS**

If the sample from a point process is independent of time then it can be specified by the intervals between events, and these intervals \( \{x_i\}_{i=1}^{n} \) can be used to construct a histogram in order to infer information about the probability distribution of the \( x_i \). A \( \chi^2 \) test can then be used to determine the goodness of fit of the theoretical distribution to the sampled data.

**SUPERPOSITION OF PERIODIC SERIES OF EVENTS**

If the series of events is periodic with period \( B \) between successive events, then \( n_i(\tau) \) versus \( i \) will approximate a straight line with constant ordinate value, and \( n(t) \) versus \( t \) will be \( n(t) = \lfloor t/B \rfloor \), where the square brackets denote the largest integer less than or equal to the argument. The
graph of $x_i$ versus $i$ will be a series of points with constant ordinate value $B$.

Now if the series is generated by the superposition of several periodic sources, then only the pooled output can be used to determine the periods $B_i$. If the series is long and the number of sources is small, it is possible to determine the $B_i$ exactly and to assign each event to its appropriate source.

This can be done by forming the histogram for the interval lengths between successive events. This will be bounded by a point concentration about $B_1$, the smallest of the $B_i$. The graph of $x_i$ versus $i$ will give the exact value of the upper bound $B_1$. Next, find an interval of length $B_1$ and from it build up the output of the first source by repeated additions and subtractions of $B_1$. Delete this set of events from the pooled series of events and analyse the remaining events to find the next smallest period.

As soon as the frequencies become very small or if two or more of the smallest frequencies are very close together, then this method is not practical. The frequency distribution of intervals is insensitive since the frequency curve is very nearly exponential except when the number of frequencies is small, or the frequencies are far apart. In order to detect whether the interval distribution is exponential (Poisson process) or merely a pooled output from periodic sources, then
the variance time curve analysis is necessary.

**VARIANCE TIME CURVE**

Let \( V(t) \) be the variance of the number of events occurring in time \((0, T]\). If the series is random with mean rate of occurrence \( \lambda \) then \( V(t) = \lambda t \). To find \( V(t) \) for the pooled output of periodic sources first consider a single source with period \( B_i \). Let \( \gamma_i = 1/B_i \). Then \( \gamma_i t = n_i + a_i \) where \( n_i \) is an integer, \( 0 \leq a_i < 1 \). Taking observations at equi­
distant intervals \( t_1, t_2, \ldots \) then an interval of length \( t_j \) contains either \( n_i \) or \( n_i + 1 \) events from this source and the limiting frequency of intervals containing \( n_i + 1 \) events is \( a_i \) since \( \gamma_i t_j = \gamma_i t_{j-1} + a_i \). Let \( Y = 1 \) be the occurrence of \( n_i + 1 \) events and \( Y = 0 \) the occurrence of \( n_i \) events. \[ \text{Prob} \{Y=1\} = a_i, \quad \text{Prob} \{Y=0\} = 1-a_i \] and \( Y \) has the bi­
nominal distribution with mean \( a_i \) and variance \( a_i(1-a_i) \).

Since the different sources are independent, then for \( N \) sources \( V(t) = \sum_{i=1}^{N} a_i(1-a_i). \) If \( t \) is very large compared with \( B_i \) then \( a_i(1-a_i) \ll t\gamma_i \) so that \( V(t) \ll \sum_{i=1}^{N} t\gamma_i = t\lambda \) where \( 1/\lambda \) is the mean interval between successive events.

As \( t \) increases, \( a_i \) takes each value between 0 and 1 equally often giving \( a_i(1-a_i) \) an average value of \( 1/6 \). For large \( t \), \( V(t) \) oscillates about an average of \( N/6 \). Therefore a graph of \( V(t) \) versus \( t \) will differentiate between a random
process and a pooled series from N periodic sources.

In order to calculate $V(t)$ the series is divided into $m$ intervals of length $\tau$ ($\tau = 2T/n$ provided good results in estimating the number of sources for the examples in Appendix A) such that $n_1(\tau)$ is small.

Let $Y_i = n_i(\tau)$

Let $U_1^r = Y_1 + Y_2 + \ldots + Y_r$

Let $U_2^r = Y_1 + Y_2 + \ldots + Y_{r+1}$

Let $U_{m-r+1}^r = Y_{m-r+1} + \ldots + Y_m$

Calculate $\sum_{i=1}^{m} (U_i^r)^2$, $\sum_{i=1}^{m} U_i^r$, $(\frac{\sum_{i=1}^{m} U_i^r}{M})^2$

and $s = \sum_{i=1}^{m} (U_i^r)^2 - (\frac{\sum_{i=1}^{m} U_i^r}{M})^2$ where $M = m - r + 1$

Now calculate $\hat{V}(r\tau) = \left(\frac{3M}{3M^2 - 3M\tau + \tau^2 - 1}\right)s$

$\hat{V}(r\tau)$ is the estimate of the variance of the number of events in an interval $r\tau$. Now plot $\hat{V}(r\tau)$ versus $r\tau$ or $r$ for a number of $r$ to obtain an estimate of the variance time curve.
CHAPTER 4
SPECTRAL ANALYSIS

Since the point process $X(t)$ is defined only at the time points $t_1, t_2, \ldots$, it is necessary to transform the one-dimensional sequence into a discrete or continuous two-dimensional stochastic process so that the autocovariance function and corresponding spectral density function can be calculated.

The sequence of intervals $\{X_1, X_2, \ldots\}$ between successive events can be used to describe the point process, the time parameter being the serial number of the event. By dividing the time axis into a large number of narrow intervals of width $dt$, and counting the number of events in each interval, a new process, $dN(t)$, is obtained. Another way of studying the process is to convert the sequence of pulses or events into a square wave by means of a flip-flop device at each pulse.

**INTERVAL LENGTHS**

The sequence of interval lengths $\{X_1, X_2, \ldots\}$ has the mean value function $u(t) = \mathbb{E}(X_1) = u$ and variance $V(X_1) = \sigma^2$. This sequence is considered as a stationary
real-valued process in discrete time. With a finite sample \( \{X_1, X_2, \ldots, X_n\} \) the sample autocovariance function \( R_T(K) \) is defined as

\[
R_T(K) = \frac{1}{n-|K|} \sum_{i=1}^{n-|K|} (X_i - u) \cdot (X_{i+|K|} - u)
\]

for \( K = 0, \pm 1, \ldots, \pm (n-1) \) and the corresponding spectral density function \( f_T(\omega) \) is defined as

\[
f_T(\omega) = \frac{1}{2\pi} \sum_{K=-n+1}^{n-1} e^{-iK\omega} R_T(K)
\]

where \( R_T(0) = \sigma^2 \)

**OCCURRENCE RATE**

Consider the process \( \{dN(t)\} \) where \( dN(t) \) is the number of events in the interval \( (t, t+dt] \).

Let \( E\left\{ \frac{dN(t)}{dt} \right\} = u(t) = u \).

The covariance function for this stationary stochastic process is defined by \( R(K) = E\left( \frac{dN(t) \cdot dN(t+K)}{(dt)^2} \right) - u^2, \quad K > 0 \).

For \( K < 0 \) \( R(K) = R(-K) \) and for \( K = 0 \), \( R(0) = \sigma^2 \)

where

\[
\sigma^2 = E\left( \left[ \frac{dN(t)}{dt} \right]^2 \right) - u^2. \quad \text{For all } K, \text{ the complete covariance function is } \overline{R}(K) = \sigma^2 \delta(K) + R(K) \text{ where } \delta(K) \text{ is the Dirac delta function, the probability density function of a probability distribution located entirely at the point}
K=0. The spectral density function for \( dN(t) \) is defined by

\[
\tilde{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iK\omega} \tilde{R}(K) dK .
\]

However, for any time interval \((0,T]\), only a finite number of intervals of length \( dt \) are used, transforming \( \{dN(t)\} \) into a discrete valued process. The autocovariance function and spectral density function would be calculated as described for the sequence of interval lengths.

The point process \( X(t) \) defined at \( t_1, t_2, \ldots, t_n \) in the interval \((0,T]\) can be transformed into a continuous process \( Y(t) \) by a square wave. Then \( Y(t) \) takes the values \( a \) and \( b \) alternately at each occurrence for the length of time until the next occurrence (Figure 4-1). Now the problem is to derive information about the original process \( X(t) \) by correlation and spectral analysis of the square wave \( Y(t) \). In order to determine if spectral analysis can yield as much information as probability analysis, a probability distribution for the occurrence of interval lengths or sign changes cannot be assumed. Then, it is necessary to calculate the sample...
mean value function and autocovariance function and then the spectral density function from the observed time series.

Without loss of generality, it can be assumed that the transformation of the time series to a square wave begins with \( Y(t) \) assigned the value \( a \). For the process \( Y(t) \), the sample mean value function is easily calculated,

\[
u = \frac{1}{T} \int_0^T Y(t) dt = \frac{1}{T} \left( \sum_{i=1}^{[n/2]} aX_{2i-1} + \sum_{i=1}^{[n/2]} bX_{2i} \right)
\]

where \([n/2]\) denotes the largest integer less than or equal to \( n/2 \). The sample autocovariance function \( R_T(K) \) is

\[
R_T(K) = \frac{1}{T} \int_0^T \left( Y(t) - \mu \right) \left( Y(t+K) - \mu \right) dt
\]

\[= \frac{1}{T} \int_0^{T-|K|} Y(t) \cdot Y(t+|K|) dt - \mu^2 \]

Since \( Y(t) \) and \( Y(t+|K|) \) only assume the values \( a \) and \( b \) then

\[
Y(t) \cdot Y(t+|K|) = \begin{cases} a^2 & \text{if } |K| < T \\ b^2 & \text{if } |K| < T \\ a \cdot b & \text{if } |K| < T \\
\end{cases}
\]

and

\[
\frac{1}{T} \int_0^{T-|K|} Y(t) \cdot Y(t+|K|) dt = \frac{1}{T} \left( r_1(K) \cdot a^2 + r_2(K) \cdot b^2 + \right.
\]

\[
\left. [T-|K|-r_1(K)-r_2(K)] \cdot ab \right)
\]

where \( r_1(K) \) and \( r_2(K) \) are to be determined.
The only problem in determining $R_T(K)$ is that of calculating $r_1(K)$ and $r_2(K)$, which is computationally difficult. $r_1(K)$ and $r_2(K)$ can be calculated for any particular value of $K$ or at equidistant intervals $\Delta K$, where $\Delta K$ can be taken as small as required. Then $R_T(K)$ is not restricted as in the discrete valued case of interval lengths, to a maximum number of estimated values equal to the total number of point events.

$$R_T(0) = \frac{1}{T} \int_0^T Y(t) \cdot Y(t) dt - u^2$$

$$= \frac{1}{T} \left( \sum_{i=1}^{n-[n/2]} a^2 X_{2i-1} + \sum_{i=n/2}^{[n/2]} b^2 X_{2i} \right) - u^2$$

where $[n/2]$ denotes the largest integer less than or equal to the argument $n/2$.

It is necessary to calculate the spectral density function from the sample autocovariances without any assumptions about the distribution of interval lengths. Once a method for estimating the spectral density function is determined, its effectiveness can be determined by comparing these results to the spectrum obtained by taking the Fourier transform of the theoretical autocovariance function.

In order to estimate the autocovariances of the square wave the following method is suggested. For $0 < K < T$ fix a value of $K$ and determine the minimum value of the integers $m_1$
and $m_2$ such that $K< \sum_{i=1}^{m_1} x_i$ and $K< \sum_{i=1}^{m_2} x_{n-i+1}$.

Let $\sum_{i=1}^{m_1} x_i - K = W_m$, and $\sum_{i=1}^{m_2} x_{n-i+1} - K = V_{n-m_2+1}$ so that

$$V = (x_1, x_2, \ldots, x_{n-m_1}, v_{n-m_2+1}) = (v_1, v_2, \ldots, v_{n-m_2+1})$$

$$W = (W_m, x_{m+1}, \ldots, x_n) = (W_m, W_{m+1}, \ldots, W_n)$$

where $V_i = x_i$ $i = 1, \ldots, n-m$

$W_j = x_j$ $j = m + 1, \ldots, n$

Now arrange the $n-m_1+1$ and $n-m_2+1$ elements in columns

<table>
<thead>
<tr>
<th>$V_1$</th>
<th>$W_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_2$</td>
<td>$W_{m+1}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{n-m_2+1}$</td>
<td>$W_n$</td>
</tr>
</tbody>
</table>

The graphic representation of $V$ and $W$ is given by Figure 4-2.

---

Fig. 4-2 The square wave as denoted by $V$ and $W$ respectively.
Take the two elements of the first row, $V_1$ and $W_m$, pick the smallest element and record it and the indices 1 and $m_i$ of these two interval lengths. Subtract the smaller element i.e. $\min\{V_1, W_m\}$ from the two interval lengths in the first row, $V_1$ and $W_m$. Either one element is now zero, or both are if $V_1 = W_m$. Any interval length in the first row reduced to zero by subtracting the smaller element from it is replaced by the next interval length in that column. If $V_1 - \min\{V_1, W_m\}$ is zero it is replaced by $V_2$, and if $W_m - \min\{V_1, W_m\}$ is zero it is replaced by $W_{m+1}$ and the new interval index or indices are recorded. If $V_1 - \min\{V_1, W_m\}$ or $W_m - \min\{V_1, W_m\}$ is not zero then the index of this reduced interval length remains the same. For example

$V = (2.0, 2.0, 1.0, 2.0, 2.0, 1.0, 0.5)$

$W = (0.5, 2.0, 1.0, 2.0, 2.0, 1.0, 1.0, 1.0)$

are obtained from a point process with eight events where $X_1 = 2.0$, $X_2 = 2.0$, $X_3 = 1.0$, $X_4 = 2.0$, $X_5 = 2.0$, $X_6 = 1.0$, $X_7 = 1.0$, $X_8 = 1.0$ and $K = 1.5$.

Then

<table>
<thead>
<tr>
<th></th>
<th>V</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.0</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>2.0</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

$\min\{V_1, W_m\} = \min\{2.0, 0.5\} = 0.5$ with indices $(1,1)$

$V_1 - \min\{V_1, W_m\} = 2.0 - 0.5 = 1.5$
\( W_{m_i} - \min \{ V_1, W_{m_i} \} = 0.5, -0.5 = 0 \), so that \( W_{m+1} = 2.0 \)

replaces \( W_{m_i} = 0.5 \).

The new array is

<table>
<thead>
<tr>
<th>V</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Now the procedure is repeated taking 1.5 and 2.0 as the new entries in the first row.

\[
\min \{ 1.5, 2.0 \} = 1.5 \text{ with indices (1,2)},
\]

\[
1.5 - \min \{ 1.5, 2.0 \} = 0, \quad \text{and element } V_2 = 2.0 \text{ is moved into the first row.}
\]

\[
2.0 - \min \{ 1.5, 2.0 \} = 0.5, \quad \text{and this is the reduced interval length for the second column.}
\]

The new array is

<table>
<thead>
<tr>
<th>V</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>0.5</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

and the same procedure is repeated on the elements 2.0 and 0.5. At each step the values of the smallest element and the interval indices are recorded and are used to calculate \( r_1(K) \) and \( r_2(K) \). If both indices are odd, then both intervals were from \( Y(t) = a \), and if both are even they are from \( Y(t) = b \).

The sum of these minimum lengths with both indices odd is the
value of $r_1(K)$ and the sum with both indices even is the value of $r_2(K)$. For a series with a large number of occurrences this method involves long tedious computations. With the aid of a computer $R_T(K)$ can be easily calculated for various values of $K$. A computer program for this is given in Appendix C.

From the preceding example the following was obtained,

<table>
<thead>
<tr>
<th>minimum value</th>
<th>indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1 1</td>
</tr>
<tr>
<td>1.5</td>
<td>1 2</td>
</tr>
<tr>
<td>0.5</td>
<td>2 2</td>
</tr>
<tr>
<td>1.0</td>
<td>2 3</td>
</tr>
<tr>
<td>0.5</td>
<td>2 4</td>
</tr>
<tr>
<td>1.0</td>
<td>3 4</td>
</tr>
<tr>
<td>0.5</td>
<td>4 4</td>
</tr>
<tr>
<td>1.5</td>
<td>4 5</td>
</tr>
<tr>
<td>0.5</td>
<td>5 5</td>
</tr>
<tr>
<td>1.0</td>
<td>5 6</td>
</tr>
<tr>
<td>0.5</td>
<td>5 7</td>
</tr>
<tr>
<td>0.5</td>
<td>6 7</td>
</tr>
<tr>
<td>0.5</td>
<td>6 8</td>
</tr>
<tr>
<td>0.5</td>
<td>7 8</td>
</tr>
</tbody>
</table>

and $r_1(K) = 1.5, \quad r_2(K) = 2.0$.

Then

$$R_T(K) = \frac{1}{T} \{r_1(K)a^2 + r_2(K)b^2 + [T-|K| - r_1(K) - r_2(K)] ab\}$$

$$R_T(1.5) = \frac{1}{12} [1.5 + 2 + (12-3.5)(-1)]$$

$$= -\frac{5}{12}$$

where $a = -b = 1$. 
Probability analysis, using the methods in Chapter 2, reveals that no single test for randomness is effective, but that a combination such as Bartlett's test and the variance time curve analysis is necessary. If the intervals between events have a probability distribution, this can be effectively determined by a histogram and a $\chi^2$ test of fit. The pooled output from several periodic sources can be distinguished from a Poisson process by means of the variance time curve, and the number of periodic sources can be determined. Unless the number of periods is small and the periods far apart, the histogram analysis of the pooled output will not reveal any information about the constituent periods. The computational results for several examples appear in Appendix A.

Spectral analysis of the interval lengths and occurrence rate yielded very little information. However, a spectral analysis of the square wave produced some very interesting results. While spectral analysis of the square wave did not determine if a process was random or not, it did distinguish between the pooled output of several periodic sources and a random or "near random" point process. The "near
random" process was the result of splicing two random processes with rates of occurrence \( \lambda_1 \) over \((0,t)\) and \( \lambda_2 \) over \((t, 2t)\) to form a single point process over \((0, 2t)\). While spectral analysis of the square wave could not distinguish between the random and "near random" series, no study was made to find out if they could be distinguished, and if so, under what conditions.

For the examples with the intervals having a probability distribution, spectral analysis of the square wave revealed a function which closely approximated the density function of the interval lengths. The autocovariance function can be calculated in terms of the inverse Laplace transform of a function involving the Laplace transform of the interval lengths [3]. However, some examples dealt with in Appendix B have truncated distributions, since the interval lengths have to be positive, and the Laplace transforms of these functions, as well as most density functions, have not been tabulated. The spectral density functions for the square wave of time series with the different density functions can be calculated and graphed, and the resulting graphs used to identify the type of point process.

The only available information from probability analysis of a pooled series from several periodic sources is the number of constituent periodic series. The method of spectral analysis to determine constituent frequencies was then applied
to the pooled series in order to see if this would locate the periods. Spectral analysis of the interval lengths and occurrence rate of the pooled series revealed little information, but analysis of the square wave revealed peak frequencies $\omega_j$, located in the range $(0, \pi)$. Furthermore, it was observed that the peak frequencies $\omega_j/\pi$ corresponded to a linear combination of the occurrence rates of the periodic sources and that the number of peaks was related to the number of periods. After examining the spectrum for many pooled series (almost 20, with some examples given in Appendix B) and encountering the same results in all of them, the following conjecture was made. It appears that for $N$ periodic sources with periods $B_i$ and rates of occurrence $\lambda_i$ ($i=1, 2, \ldots, N; N$ small), spectral analysis of the square wave of the pooled output results in the occurrence of $2^{N-1}$ peaks of $f_T(\omega)$ at frequencies $\omega_j$, $j=1, 2, \ldots, 2^{N-1}$. If $\hat{\lambda}_j = \omega_j/\pi$, then each $\hat{\lambda}_j$ is a function of $\lambda_1, \lambda_2, \ldots, \lambda_N$ and is equal to the absolute value of a linear combination of the $\lambda_i$,

$$\hat{\lambda}_j = |a_1\lambda_1 + a_2\lambda_2 + \ldots + a_N\lambda_N| \quad j=1, 2, \ldots, 2^{N-1}$$

where $a_1 = 1$ and $a_K = \begin{cases} +1 & K=2, \ldots, N \\ -1 & \end{cases}$

The $2^{N-1}$ possible sign changes for the $a_K$, $K=2, \ldots, N$ account for the $2^{N-1}$ peak frequencies.

The proof of this conjecture is hindered by the fact that no one has produced a model to describe the superposition series of several periodic sources. A model to describe the
pooled output from two periodic sources with periods B_1 and B_2, B_1<B_2, is X(t) = Y(t)+R(t) where Y(t) is a periodic point process with period (B_1+B_2)/(B_2-B_1) and R(t) is a point process with the intervals between events having a uniform distribution over (0, B_1). This was obtained by analysis of the frequency of the interval length B_1 from the pooled output. Further analysis in this direction may prove rewarding, but it is possible that no closed expression exists to describe the pooled series since the interval lengths tend to an exponential distribution as the number of sources increases [4]. The effect of increasing the number of sources has not been investigated yet, examples with N=2, 3, and 4 periodic sources are given in Appendix B, and all computational results pertaining to Chapter 4.
APPENDIX A

In order to evaluate the information obtainable from probability and spectral analysis, several types of point processes were generated and analysed.

TABLE A-1

<table>
<thead>
<tr>
<th>EXAMPLE</th>
<th>TYPE OF PROCESS</th>
<th>NUMBER OF EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Poisson process with mean rate of occurrence $\lambda = 0.167$</td>
<td>250</td>
</tr>
<tr>
<td>P2</td>
<td>Point process consisting of a Poisson process with $\lambda_1 = 0.203$ over the time period $(0, 600]$ and a Poisson process with $\lambda_2 = 0.297$ over $(600, 1200]$</td>
<td>300</td>
</tr>
<tr>
<td>P3</td>
<td>Poisson process with mean rate of occurrence $\lambda = 0.208$</td>
<td>250</td>
</tr>
<tr>
<td>S1</td>
<td>Superposition of 4 periodic point processes with periods 2.00, 2.27, 5.15, 8.23</td>
<td>250</td>
</tr>
<tr>
<td>S2</td>
<td>Superposition of 4 periodic point processes with periods 1.93, 2.89, 3.83, 3.96</td>
<td>250</td>
</tr>
</tbody>
</table>
Examples P1, P2, P3, were generated using a table of random numbers. In order to obtain a point process with Gaussian distributed interval lengths, a computer subroutine was used to obtain a series of random Gaussian distributed numbers with mean zero and any standard deviation, $\sigma$. The subroutine generates two uniformly distributed numbers on $(0,1)$.  

<table>
<thead>
<tr>
<th>EXAMPLE</th>
<th>TYPE OF PROCESS</th>
<th>NUMBER OF EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>Point process with the intervals between events Gaussian distributed with density function $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp{-1/2(x-u)^2/\sigma^2}$ where $u = 1.50$ and $\sigma = 0.50$</td>
<td>250</td>
</tr>
<tr>
<td>G2</td>
<td>Point process with the intervals between events Gaussian distributed with $u = 6.0$ and $\sigma = 2.0$</td>
<td>250</td>
</tr>
<tr>
<td>R1</td>
<td>Point process with the intervals between events having a Rayleigh distribution with density function $f(x) = \frac{x}{r^2} \exp{-1/2(x/r)^2}$ where $r^2 = 2.0$</td>
<td>250</td>
</tr>
<tr>
<td>R2</td>
<td>Point process with the intervals between events having a Rayleigh distribution with $r^2 = 5.0$</td>
<td>250</td>
</tr>
</tbody>
</table>
x and y. These are used to give two elements of the list as
\((-2 \log x)^{1/2} \cos 2\pi y\) and \((-2 \log x)^{1/2} \sin 2\pi y\). By a translation of the origin to \(-3a\), negative numbers were eliminated and a mean \(u = 3\sigma\) was obtained. The interval lengths with a Rayleigh distribution were generated by using the probability integral transform \(y = \int f(t)dt\)
where \(f(t) = \frac{t}{r^2} e^{-1/2(t/r)^2}\). This transformation maps the positive real line into the interval from zero to one, and \(y\) has the uniform distribution. Then \(y = F(x) = \int f(t)dt\) and \(y = 1-e^{-1/2(x/r)^2}\) yields \(x = 2r^2 \cdot \log y\). Random numbers on the interval \((0, 1)\) were generated by a subroutine which computed a string of pseudorandom numbers \(z_i\) by the relation \(z_i = az_{i-1} \mod m\). Values of \(x\) were calculated by solving the equation \(x = \sqrt[2]{2r^2} \cdot \log y\) (\(y\) takes on the values \(z_i\)).

GRAPHICAL METHODS

For each example the interval lengths \(x_i\) were plotted versus \(i\), but only for the first 40 or 50 interval lengths since these were considered as a representative sample from each point process. If the graph of \(x_i\) versus \(i\) indicated that the interval lengths were independent of time, then the histogram of interval lengths was formed, (Figure A-1 to Figure A-9).

Examination of the graphs for each example reveals several examples which appear to be Poisson processes, \(P1\), \(P2\), \(P3\), \(S1\) and \(S2\). For each one of these examples \(x_i\) versus \(i\) is
Fig. A-1(a) Interval lengths versus index for example Pl.

Fig. A-1(b) Histogram of interval lengths for example Pl.
Fig. A-2(a) Interval lengths versus index for example P2.

Fig. A-2(b) Histogram of interval lengths for example P2.
Fig. A-3(a) Interval lengths versus index for example P3.

Fig. A-3(b) Histogram of interval lengths for example P3.
Fig. A-4(a) Interval lengths versus index for example Sl.

Fig. A-4(b) Histogram of interval lengths for example Sl.
Fig. A-5(a) Interval lengths versus index for example S2.

Fig. A-5(b) Histogram of interval lengths for example S2.
Fig. A-6(a) Interval lengths versus index for example Gl.

Fig. A-6(b) Histogram of interval lengths for example Gl.
Fig. A-7(a) Interval lengths versus index for example G2.

Fig. A-7(b) Histogram of interval lengths for example G2.
Fig. A-8(a) Interval lengths versus index for example Rl.

Fig. A-8(b) Histogram of interval lengths for example Rl.
Fig. A-9(a) Interval lengths versus index for example R2.

Fig. A-9(b) Histogram of interval lengths for example R2.
a scatter diagram and the histogram of interval lengths approximates the probability density function of the exponential distribution. Tests for randomness were then tried in an attempt to identify the type of process as Poisson or not. The test based on interval lengths, Bartlett's test, was tried first and then compared with other tests based on the number of occurrences in an interval length.

**TESTS FOR RANDOMNESS**

**BARTLETT'S TEST**

Each example was divided up into $K = 10$ equal series of $m = n/K$ interval lengths, and

$$X_i = \sum_{j=1}^{m} \frac{x_j}{m}$$

$i = 1, \ldots, K$ was calculated for the $K$ sets of interval lengths. Then $\chi^2$ was calculated and the hypothesis that all the $\lambda_i$ are equal was tested at the 5% level by calculating $\chi^2_{.05, K-1}$ with $K-1$ degrees of freedom.

$$\chi^2 = 2.3026 \left[ 2(m-1) \log \left( \frac{1}{K} \sum_{i=1}^{K} X_i \right) - 2(m-1) \sum_{i=1}^{K} \log X_i \right] / C$$

$$C = 1 + (K+1)/[3K(m-1)]$$

The results are tabulated along with the results from other tests of randomness (TABLE A-3).

**KOLMOGOROV-SMIRNOV TEST**

This test for randomness is a measure of the maximum deviation of $n(t)$ from the straight line joining the points $(0, 0)$ and $(T, n)$. For a sample of size $n$ in a continuum $(0, T)$
\[ n \cdot D_n = \max_{0 \leq t \leq T} | n(t) - n \cdot t/T | \] is calculated.

To test the hypothesis of randomness a significance level \( \alpha \) is selected, and a critical value of \( E_{n\alpha} \), which is tabulated for particular \( n \) and \( \alpha \), is selected. If \( n \cdot D_n > E_{n\alpha} \), the hypothesis that the series is random is rejected. Instead of finding \( E_{n\alpha} \) for a particular \( n \) and \( \alpha \) from tables the approximations
\[
E_{n.95} \approx 1.358 \sqrt{n} \\
E_{n.99} \approx 1.628 \sqrt{n}
\]
were used.

**RECTANGULAR DISTRIBUTION TEST**

The time interval \( (0,T] \) was divided up into \( m = 25 \) equal time intervals of length \( \tau \) for each example. \( O_i \), the number of events observed in the interval \( (i-1)\tau, i\tau] \) \( i=1, \ldots, m \) and \( E = n \cdot \tau/T \) were calculated. Then

\[
\chi^2 = \sum_{i=1}^{m} \frac{(O_i - E)^2}{E}
\]

The upper and lower limits of \( \chi^2 \) for a 95\% confidence interval and \( m-1 = 24 \) degrees of freedom are \( \chi^2_{.975, 24} = 39.4 \) and \( \chi^2_{.025, 24} = 12.4 \) respectively.

**VARIANCE TIME CURVE**

In order to distinguish between a Poisson process and the pooled output of several periodic sources the variance time curve was calculated for the examples whose histogram of interval lengths appeared exponentially distributed. The series was divided up into small intervals of length \( \tau \), and \( Y_i = n_i(\tau) \) was calculated.
\( U_i^r = Y_i + Y_{i+1} + \ldots + Y_{i+r-1} \) was calculated for \( i = 1, 2, \ldots, M \) and \( r = 1, 2, \ldots, 15 \) where \( M = 100 \).

Then

\[
\hat{V}(r^\tau) = \left[ \sum_{i=1}^{M} (U_i^r)^2 - \left( \sum_{i=1}^{M} U_i^r \right)^2 \right] \cdot \left[ \frac{3M^2}{(3M^2 - 3Mr + r^2 - 1)} \right]
\]

was tabulated and graphed (TABLE A-2, Figure A-10(a) to A-10(d))

<table>
<thead>
<tr>
<th>( r )</th>
<th>( \hat{V}(r^\tau) ) EX</th>
<th>( \hat{V}(r^\tau) ) EX</th>
<th>( \hat{V}(r^\tau) ) EX</th>
<th>( \hat{V}(r^\tau) ) EX</th>
<th>( \hat{V}(r^\tau) ) EX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.11</td>
<td>2.61</td>
<td>2.43</td>
<td>0.56</td>
<td>0.89</td>
</tr>
<tr>
<td>2</td>
<td>3.68</td>
<td>6.47</td>
<td>5.46</td>
<td>0.68</td>
<td>0.59</td>
</tr>
<tr>
<td>3</td>
<td>5.43</td>
<td>10.3</td>
<td>8.17</td>
<td>0.63</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>6.61</td>
<td>14.9</td>
<td>10.9</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>5</td>
<td>7.88</td>
<td>20.9</td>
<td>13.8</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>6</td>
<td>9.19</td>
<td>27.9</td>
<td>17.1</td>
<td>0.78</td>
<td>0.69</td>
</tr>
<tr>
<td>7</td>
<td>11.3</td>
<td>35.3</td>
<td>20.0</td>
<td>0.67</td>
<td>0.90</td>
</tr>
<tr>
<td>8</td>
<td>13.4</td>
<td>43.9</td>
<td>23.1</td>
<td>0.24</td>
<td>0.37</td>
</tr>
<tr>
<td>9</td>
<td>15.9</td>
<td>53.0</td>
<td>25.4</td>
<td>0.58</td>
<td>0.80</td>
</tr>
<tr>
<td>10</td>
<td>18.2</td>
<td>61.4</td>
<td>26.8</td>
<td>0.67</td>
<td>0.58</td>
</tr>
<tr>
<td>11</td>
<td>20.0</td>
<td>70.2</td>
<td>28.1</td>
<td>0.81</td>
<td>0.77</td>
</tr>
<tr>
<td>12</td>
<td>22.2</td>
<td>80.4</td>
<td>29.4</td>
<td>0.75</td>
<td>0.88</td>
</tr>
<tr>
<td>13</td>
<td>23.2</td>
<td>91.5</td>
<td>30.3</td>
<td>0.52</td>
<td>0.46</td>
</tr>
<tr>
<td>14</td>
<td>25.0</td>
<td>102.</td>
<td>30.4</td>
<td>0.88</td>
<td>0.83</td>
</tr>
<tr>
<td>15</td>
<td>26.8</td>
<td>113.</td>
<td>31.5</td>
<td>0.76</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Fig. A-10(a) Variance time curve for example P1.

Fig. A-10(b) Variance time curve for example P2.
Fig. A-10(c) Variance time curve for example Sl

Fig. A-10(d) Variance time curve for example S2.
The results for the various tests for randomness are tabulated in order to compare their accuracy. (TABLE A-3)

<table>
<thead>
<tr>
<th>Example</th>
<th>Bartlett's Test $\chi^2$</th>
<th>Kolmogorov-Smirnov Test $n^*D_n$</th>
<th>Rect. Dist. Test $\chi^2$</th>
<th>Variance Time Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>6.80</td>
<td>13.2 21.5 25.7</td>
<td>14.8</td>
<td>Linear</td>
</tr>
<tr>
<td>P2</td>
<td>19.58</td>
<td>35.0 23.5 28.2</td>
<td>29.2</td>
<td>Linear</td>
</tr>
<tr>
<td>P3</td>
<td>5.96</td>
<td>15.7 21.5 25.7</td>
<td>20.8</td>
<td>Linear</td>
</tr>
<tr>
<td>S1</td>
<td>0.26</td>
<td>2.31 21.5 25.7</td>
<td>1.40</td>
<td>Oscillates</td>
</tr>
<tr>
<td>S2</td>
<td>0.22</td>
<td>2.70 21.5 25.7</td>
<td>2.40</td>
<td>Oscillates</td>
</tr>
</tbody>
</table>

The .95 upper bound for Bartlett's Test is 16.92 and the .95 bounds for the rectangular distribution test are [12.4, 39.4]. It is quite apparent that there is no single test of randomness that is effective for all point processes, and that a combination of tests such as the variance time curve and Bartlett's test is necessary.

INTERVALS WITH PROBABILITY DISTRIBUTION

The histogram of the following examples indicate that the intervals may have a probability distribution.
<table>
<thead>
<tr>
<th>Example</th>
<th>Probability Distribution</th>
<th>Density Function</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>EXPONENTIAL</td>
<td>( f(x) = \lambda e^{-\lambda x} ) ( x &gt; 0 ) ( = 0 ) ( \text{otherwise} )</td>
<td>( 1/\lambda )</td>
<td>( 1/\lambda^2 )</td>
</tr>
<tr>
<td>P2</td>
<td>EXPONENTIAL on ( (0, 600] )</td>
<td>( f(x) ) where ( \lambda = \lambda_1 ) ( 1/\lambda_1 ) ( 1/\lambda_1^2 )</td>
<td>( 1/\lambda_2 )</td>
<td>( 1/\lambda_2^2 )</td>
</tr>
<tr>
<td></td>
<td>EXPONENTIAL on ( (600, 1200] )</td>
<td>( f(x) ) where ( \lambda = \lambda_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>EXPONENTIAL</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G1</td>
<td>GAUSSIAN</td>
<td>( g(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} ) ( x &gt; 0 ) ( = 0 ) ( x \leq 0 )</td>
<td>( \mu )</td>
<td>( \sigma^2 )</td>
</tr>
<tr>
<td>G2</td>
<td>GAUSSIAN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R1</td>
<td>RAYLEIGH</td>
<td>( h(x) = \frac{x}{r^2} e^{-x^2/r^2} ) ( x &gt; 0 ) ( = 0 ) ( x \leq 0 )</td>
<td>( r\sqrt{\pi/2} )</td>
<td>( r^2 (2-\pi/2) )</td>
</tr>
<tr>
<td>R2</td>
<td>RAYLEIGH</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
The sample mean $\bar{X}$ and sample variance $S_x^2$ were calculated for the sequence of interval lengths $\{x_i\}_1^n$ for each example where

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{X})^2.$$

The statistics $\bar{X}$ and $S_x^2$ were used as estimates of the mean and variance respectively of the theoretical distribution. A $\chi^2$ test of fit was then used to determine the goodness of fit of the histogram to the theoretical distribution.

Example P1 is worked out in detail.

Ex. P1 $\bar{X} = 5.996$, $\lambda = 1/\bar{X} = 0.167$, sample size $n = 250$

The exponential curve to fit the histogram would be $f(x) = 0.167e^{-0.167x}$. For the interval length $L = [a, b)$ let $P(L)$ equal the area under $f(x)$ for $a \leq x < b$.

Then

$$P(L) = \int_a^b f(x) \, dx = e^{-\lambda a} - e^{-\lambda b}$$

$$= e^{-0.167a} - e^{-0.167b}$$

The expected number of interval lengths $L$ is calculated as $nP(L)$, and the observed number of interval lengths of length $L$ are recorded. (TABLE A-5)
<table>
<thead>
<tr>
<th>Interval Length</th>
<th>Probability of L ( P(L) )</th>
<th>Expected Number ( E )</th>
<th>Observed Number ( O )</th>
</tr>
</thead>
<tbody>
<tr>
<td>([0.0,1.0))</td>
<td>0.154</td>
<td>38.5</td>
<td>44</td>
</tr>
<tr>
<td>([1.0,2.0))</td>
<td>0.130</td>
<td>32.5</td>
<td>29</td>
</tr>
<tr>
<td>([2.0,3.0))</td>
<td>0.110</td>
<td>27.5</td>
<td>29</td>
</tr>
<tr>
<td>([3.0,4.0))</td>
<td>0.093</td>
<td>23.3</td>
<td>26</td>
</tr>
<tr>
<td>([4.0,5.0))</td>
<td>0.079</td>
<td>19.3</td>
<td>18</td>
</tr>
<tr>
<td>([5.0,6.0))</td>
<td>0.067</td>
<td>16.8</td>
<td>21</td>
</tr>
<tr>
<td>([6.0,7.0))</td>
<td>0.056</td>
<td>14.0</td>
<td>12</td>
</tr>
<tr>
<td>([7.0,8.0))</td>
<td>0.048</td>
<td>12.0</td>
<td>16</td>
</tr>
<tr>
<td>([8.0,9.0))</td>
<td>0.041</td>
<td>10.3</td>
<td>5</td>
</tr>
<tr>
<td>([9.0,10.0))</td>
<td>0.034</td>
<td>8.5</td>
<td>11</td>
</tr>
<tr>
<td>([10.0,11.0))</td>
<td>0.029</td>
<td>7.3</td>
<td>4</td>
</tr>
<tr>
<td>([11.0,12.0))</td>
<td>0.024</td>
<td>6.0</td>
<td>2</td>
</tr>
<tr>
<td>([12.0,13.0))</td>
<td>0.021</td>
<td>5.3</td>
<td>3</td>
</tr>
<tr>
<td>([13.0,14.0))</td>
<td>0.017</td>
<td>4.3</td>
<td>7</td>
</tr>
<tr>
<td>([14.0,15.0))</td>
<td>0.015</td>
<td>3.8</td>
<td>4</td>
</tr>
<tr>
<td>([15.0,16.0))</td>
<td>0.013</td>
<td>3.3</td>
<td>3</td>
</tr>
<tr>
<td>([16.0,17.0))</td>
<td>0.011</td>
<td>2.8</td>
<td>2</td>
</tr>
<tr>
<td>([17.0,18.0))</td>
<td>0.009</td>
<td>2.3</td>
<td>3</td>
</tr>
<tr>
<td>([18.0,19.0))</td>
<td>0.007</td>
<td>1.8</td>
<td>1</td>
</tr>
<tr>
<td>([19.0,20.0))</td>
<td>0.007</td>
<td>1.8</td>
<td>3</td>
</tr>
<tr>
<td>([20.0,21.0))</td>
<td>0.005</td>
<td>1.3</td>
<td>0</td>
</tr>
</tbody>
</table>

\(O, E\) and \(|O-E|\) are tabulated in \(K\) cells. When the class frequency \(E\) is less than five the adjoining cells are added to ensure a class frequency of five or more. (TABLE A-6)
TABLE A-6

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>44</td>
<td>29</td>
<td>29</td>
<td>26</td>
<td>18</td>
<td>21</td>
<td>12</td>
<td>16</td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Expected</td>
<td>39</td>
<td>33</td>
<td>28</td>
<td>23</td>
<td>19</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D/E</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

K = 16 cells  Calculate  \( \chi^2 = \sum \frac{(O-E)^2}{E} = 16.5 \)

If \( \chi^2 > \chi^2_{K-p-1,.95} \), the hypothesis that the interval lengths are exponentially distributed is rejected at the 5% level. The number of degrees of freedom, \( K-p-1 \), consists of the number of cells \( K \), minus the number \( p \) of parameters estimated, minus one.

\[ \chi^2_{14,.95} = 23.7 \]

The results for all examples were tabulated. (TABLE A-7).

Under the hypothesis that the point processes have a probability distribution, the \( \chi^2 \) test of fit accepts the hypothesis that the histogram of the interval lengths is a good fit to the theoretical distribution.
TABLE A-7

<table>
<thead>
<tr>
<th>Example</th>
<th>Mean</th>
<th>Variance</th>
<th>Sample Mean</th>
<th>Sample Variance</th>
<th>Test of fit $\chi^2$ .95 lim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>6.00</td>
<td>36.0</td>
<td>6.00</td>
<td>29.35</td>
<td>16.5 23.7</td>
</tr>
<tr>
<td>P2</td>
<td>4.93</td>
<td>24.3</td>
<td>4.00</td>
<td>13.86</td>
<td>18.7 21.0</td>
</tr>
<tr>
<td></td>
<td>over (0.600]</td>
<td>3.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>over (600,1200]</td>
<td>11.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>4.81</td>
<td>23.1</td>
<td>4.79</td>
<td>22.69</td>
<td>13.6 22.4</td>
</tr>
<tr>
<td>G1</td>
<td>1.50</td>
<td>0.25</td>
<td>1.48</td>
<td>0.22</td>
<td>8.8  15.5</td>
</tr>
<tr>
<td>G2</td>
<td>6.00</td>
<td>4.0</td>
<td>5.93</td>
<td>3.46</td>
<td>9.3  22.4</td>
</tr>
<tr>
<td>R1</td>
<td>1.80</td>
<td>0.86</td>
<td>1.74</td>
<td>0.84</td>
<td>16.2 26.3</td>
</tr>
<tr>
<td>R2</td>
<td>2.80</td>
<td>2.15</td>
<td>2.93</td>
<td>2.34</td>
<td>12.6 19.7</td>
</tr>
</tbody>
</table>

SUPERPOSITION OF EVENTS

Consider the point process constructed by pooling three periodic point processes with periods 1.40, 3.63 and 7.81. The graph of $x_i$ versus $i$ is bounded at $x_i = 1.40$ and the histogram of interval lengths reveals a bound at one end, maximum of the $x_i$ (Fig. A-11). The series of events with period $B = 1.40$ can be deleted from the pooled series. Then the graph of $x_i$ versus $i$ is bounded above by $x_i = 3.63$ and this series of events with period $B = 3.63$ can be deleted from the pooled series. The constituent periods can be determined and each point can be assigned to its proper source,
Fig. A-11(a) Interval lengths versus index for a pooled series.

Fig. A-11(b) Histogram of interval lengths for the pooled series.
but only if the periods are not close together.

Series $S_1$ and $S_2$ are not bounded by a particular value of $x_i$, so that the periods can not be found in the above manner. However, the number of periodic sources can be found by employing the variance time curve. $S_1$ oscillates about 0.65 and $S_2$ about 0.68. Since $N/6 = 0.65$, $S_1$ is the pooled output of $N = 4$ periodic sources and $S_2$ is also the pooled output of 4 periodic sources. This is the only information that can be derived from these processes by this method of analysis.
APPENDIX B
SPECTRAL ANALYSIS

INTERVAL LENGTHS

The sequence of interval lengths \( \{X_i\}_{i=1}^n \) with mean \( \mu \) and variance \( \sigma^2 = R_T(0) \) has the sample autocovariance function

\[
R_T(K) = \frac{1}{n-|K|} \sum_{i=1}^{n-|K|} (X_i - \mu)(X_{i+|K|} - \mu) \quad |K| = 0, 1 \ldots, n-1.
\]

\( R_T(K) \) was calculated using the statistics \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \) and

\[
S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]
as estimates of the mean \( \mu \) and variance \( \sigma^2 \) respectively. Then the spectral density function

\[
\hat{f}_T(\omega) = \frac{1}{2\pi} R_T(0) + \frac{1}{\pi} \sum_{K=1}^{n/2} \cos K\omega R_T(K)
\]
or

\[
\left( \frac{\hat{f}_T(\omega)}{\sigma^2} \right) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{K=1}^{n/2} \cos K\omega \frac{R_T(K)}{R_T(0)} \right\}
\]

was estimated at equidistant intervals of \( \omega \) for \( 0 \leq \omega \leq \pi \), so that values of \( \hat{f}_T(\omega_j) \) were recorded for \( \omega_j = \frac{2\pi j}{n}, j=0,1,\ldots,n/2 \). \( R_T(K) \) was estimated for \( K=1, 2,\ldots,n/5 \) and \( K=1, 2,\ldots,n-1 \).

Using \( n-1 \) autocovariances resulted in negative values of \( \hat{f}_T(\omega_j) \) for values of \( \omega_j \) near \( \pi \), while using only twenty per
cent of the autocovariances reduced this variability of $\int T(\omega_j)$. The estimate of $\int T(\omega)/\sigma^2$ versus $\omega$ was graphed for several of the examples. (Fig. B-1 to B-6)

Spectral analysis of the interval lengths of the examples revealed very little information about the frequency distribution of the interval lengths. For the superposition of several periodic sources one or more peak frequencies was located. For example Fig. B-5, the superposition of two periodic sources with periods $B_1 = 2.0$, $B_2 = 2.27$, and $\lambda_1 = 0.500$, $\lambda_2 = 0.441$ where $\lambda_i = 1/B_i$ $i=1,2$ had one peak frequency at $\omega_j$ where $j=117$, $n=250$. Then $\omega_j/\pi = 0.936$ and $\lambda_1 + \lambda_2 = 0.941$. However, no relationship between the peak frequencies and rates of occurrence of the constituent periodic sources could be determined for the pooled output of three and four periodic sources. For example, Fig. B-6 records two peak frequencies and $\omega_j/\pi = \begin{cases} 0.704 & j=88 \\ 0.952 & j=119 \end{cases}$ (n=250).

This example is the superposition of three periodic sources with $B_1 = 1.93$, $B_2 = 2.89$, $B_3 = 8.27$ and $\lambda_1 + \lambda_2 + \lambda_3 = 0.985$, $\lambda_1 + \lambda_2 - \lambda_3 = 0.743$. Example S1 (Fig. B-2) with four periodic sources has only two peak frequencies revealed and $\omega_j/\pi = \begin{cases} 0.704 & j=88 \\ 0.800 & j=100 \end{cases}$ (n=250).

However, no linear combination $\alpha_1 \lambda_1 + \alpha_2 \lambda_2 + \ldots + \alpha_N \lambda_N$, where $\alpha_i = \begin{cases} +1 \\ -1 \end{cases}$ and $N$ is the number of periodic sources, was close to the values of $\omega_j/\pi$. 
Fig. B-1 Spectral density function of the interval lengths for example Pl.

Fig. B-2 Spectral density function of the interval lengths for example Sl.
Fig. B-3 Spectral density function of the interval lengths for example Gl.

Fig. B-4 Spectral density function of the interval lengths for example R2.
Fig. B-5 Spectral density function of the interval lengths for a pooled series from 2 periodic sources.

Fig. B-6 Spectral density function of the interval lengths for a pooled series from 3 periodic sources.
OCCURRENCE RATE

The process \( \{dN(t)\} \) was obtained by dividing up the point process \( X(t) \) into \( n \) intervals of equal length \( \tau \) and counting the number of occurrences \( n_i(\tau) \) in the \( i \)th interval \( i=1, 2, \ldots, n \). Then

\[
\left\{ \frac{dN(t)}{dt} \right\} = \left\{ \frac{n_i(\tau)}{\tau} \right\}_{i=1}^n
\]

\[
\mu = \sum_{i=1}^{n} \frac{n_i(\tau)}{n \tau}
\]

\[
\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{n_i(\tau)}{\tau} - \mu \right)^2 = R_T(0)
\]

\[
R_T(K) = \frac{1}{n-|K|} \sum_{i=1}^{n-|K|} \left( \frac{n_i(\tau)}{\tau} - \mu \right) \left( \frac{n_{i+|K|}(\tau)}{\tau} - \mu \right)
\]

were calculated. The estimate of \( \int_T(\omega) / \sigma^2 \) versus \( \omega \) was graphed for several of the examples (Fig. B-7 to B-9). Very little information was obtained by this method of analysis other than the observation that the spectrum of the occurrence rate resembled the spectrum of the interval lengths for the superposition of several periodic sources.
Fig. B-7 Spectral density function of the occurrence rate for example Pl.

Fig. B-8 Spectral density function of the occurrence rate for example Sl.
Fig. B-9 Spectral density function of the occurrence rate for example Gl.

*SQUARE WAVE*

The spectral density function

\[ f_T(\omega_j) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{K} \frac{R_T(K) \cos K\omega}{\sigma^2} \right\} \quad \omega_j = 2\pi j/n, \quad j=0,1,2,\ldots,n/2 \]

was calculated using \( n=250 \) values of \( R_T(K) \) at equidistant intervals of length \( \Delta K = T/5n \). This was done for square wave values \( a=1, b=-1 \) and \( a=1, b=0 \) for all examples. There was no significant difference in the spectral density estimates obtained except about very low frequencies. This was attributed to the mean \( \mu \) which dominates the contribution to \( f(\omega) \) at zero frequency as well as frequencies close to zero. The
effect is reduced by subtracting the sample mean from the autocovariances but some effect still persists. When \( a=1, b=-1 \) the sample mean tends to be closer to zero than when \( a=1 \) and \( b=0 \). For this reason only, the results of spectral analysis are given for \( a=1, b=-1 \).

\( \frac{\sum T(\omega_d)}{\sigma^2} \) was calculated using the first 16 autocovariances only, (Fig. B-10 to B-16)

Spectral analysis of the square wave transformation appears to yield more information about the point process and the frequency distribution of its interval lengths than the other two methods of spectral analysis. Some of the experimental results can be compared with the theoretical results. The autocovariance function \( R(K) \) can be written as

\[
R(K) = \lim_{M \to \infty} \frac{1}{2\pi} \int_{-M}^{M} T(x)T(x+K)dx
\]

where \( T(x) \) is the square wave. It has been shown by F. Brooks and N. Diamantides that

\[
R(K) = \frac{1}{2} - \frac{1}{2u} L^{-1}\left\{ \frac{1-\mathcal{F}(s)}{s^2(1+\mathcal{F}(s))} \right\}
\]

where \( \mathcal{F}(s) \) denotes the Laplace transform of the density function \( f(x) \) of the interval lengths, \( u \) denotes the mean of \( f(x) \), and \( L^{-1} \) denotes the inverse Laplace transform.

**SUPERPOSITION OF PERIODIC SERIES OF EVENTS**

For the examples of the superposition of periodic sources, peaks were observed in the square wave spectral
Fig. B-10(a) Spectral density function of the square wave for example Pl.

Fig. B-10(b) Spectral density function of the square wave for example Pl (16 autocovariances).
Fig. B-11(a) Spectral density function of the square wave for example Gl.

Fig. B-11(b) Spectral density function of the square wave for example Gl (16 autocovariances).
Fig. B-12(a) Spectral density function of the square wave for example R2.

Fig. B-12(b) Spectral density function of the square wave for example R2 (16 autocovariances).
Fig. B-13  Spectral density function of the square wave for a periodic process.

Fig. B-14  Spectral density function of the square wave for the pooled series from 2 periodic sources.
Fig. B-15 Spectral density function of the square wave for a pooled series from 3 periodic sources.

Fig. B-16 Spectral density function of the square wave for example Sl.
estimates as was the case with interval and occurrence rate analysis, but more peaks were obtained. In order to discover the relationship (if any) of these peaks to the constituent periods, the superposition process was studied for examples with one, two, three and then four periodic sources.

For a single periodic source with period $B$ and rate of occurrence $\lambda = 1/B$, the square wave can be considered as an approximation of the sine curve $Z(t) = \sin \omega t$ where $\omega = \pi/\Delta t$ radians per unit time. Then the occurrences at the time instants $B, 2B, \ldots$ are the zeroes of $Z(t) = \sin \omega t, t > 0$ where $\omega = \pi/\Delta t = \pi/B = \pi \lambda$. (Fig. B-17)

![Fig. B-17 Square wave approximation of a sine wave.](image)

For a single periodic source of period $B$ the spectrum $f(\omega)$ consists of a single peak at $\omega = \pi/B$, and information about the rate of occurrence and period is directly available. Fig. B-13 is the spectrum of the periodic point process with $B=1.83$, calculated for values of $\omega$ where $\omega = \frac{2\pi j}{n}$, $j=0,1,\ldots,n/2$ and $n=250$. The peak occurred at $j=68$ and since $\omega = \pi \lambda$ \[ \lambda = \frac{2.68}{0.544} = 4.92. \] Then $B = 1/0.544 = 1.84$
and the period can accurately be estimated.

For two periodic sources with periods $B_1 = 2.00$ and $B_2 = 2.27$ and mean rates of occurrence $\lambda_1 = 0.500$ and $\lambda_2 = 0.441$, spectral analysis of the pooled output revealed two peaks at $\omega_j$ where $j=7$ and $j=118$. (Fig. B-14)

Let $\lambda_1 = \omega_j/\pi$, $j=7$, and $\lambda_2 = \omega_j/\pi$, $j=118$. Then $\lambda_1 = 0.056$ and $\lambda_2 = 0.944$ and $\lambda_1 = \lambda_1 - \lambda_2 = 0.059$

and $\lambda_2 = \lambda_1 + \lambda_2 = 0.941$

since $\omega_j = \frac{2\pi j}{n}$, $j=0, 1, \ldots, n/2$, $n=250$.

Fig. B-15 is an example of three periodic sources pooled together: $B_1 = 1.93$, $\lambda_1 = 0.518$; $B_2 = 2.89$, $\lambda_2 = 0.346$; $B_3 = 8.27$, $\lambda_3 = 0.121$. Spectral analysis of the square wave of the pooled output revealed peaks at frequencies $\omega_j$ where $j=6, 37, 93$ and $123$ ($n=250$). Then $\lambda_1 = 0.048$, $\lambda_2 = 0.296$, $\lambda_3 = 0.744$ and $\lambda_4 = 0.984$. $\lambda_1 + \lambda_2 + \lambda_3 = 0.985 = \lambda_4$

$\lambda_1 + \lambda_2 - \lambda_3 = 0.743 = \lambda_3$

$\lambda_1 - \lambda_2 + \lambda_3 = 0.293 = \lambda_2$

$\lambda_1 - \lambda_2 - \lambda_3 = 0.051 = \lambda_1$

Since $\omega = \pi \lambda$ and $\omega$ can be measured by $\int f(\omega)$ only in the range $[0, \pi]$ then $\lambda$ can be measured for $0 \leq \lambda \leq 1$. Since values of $\lambda > 1$ are possible, then the original time series $X(t)$ of length $T$ can be converted to time $cT$, where $c$ is a positive constant. Then each periodic series of period $B_i$ is transformed into a
periodic series of period $cB_i$ with rate of occurrence $\lambda_i = 1/cB_i$. For $c>1$, there exists a number $c_0$ such that $\lambda_i < 1$ for all $c>c_0$, so that any rate of occurrence can be decreased in magnitude by lengthening the period of the original time series. Similarly, any rate of occurrence can be increased for $c<1$. Since $\pi \sum \lambda_i$ is the largest frequency obtained in spectral analysis of the square wave of the pooled series, then if $\sum \lambda_i > 1$ this frequency is outside the measurable range $[0, \pi]$. Then, it is necessary to convert the pooled output in time $T$ to a pooled output in time $cT$. Then each constituent periodic series is converted to time $cT$, so that $c$ can be taken sufficiently large to ensure that $c^{-1} \pi \sum \lambda_i < \pi$.

For example, consider the two periodic series with periods $B_1=1.21, B_2=1.38$ and mean rates of occurrence $\lambda_1=0.826$ and $\lambda_2=0.725$ respectively. Then $\lambda_1 + \lambda_2 = 1.541$ and $\lambda_1 - \lambda_2 = 0.101$. Since $\sum \lambda_i > 1$, then only one peak frequency should be discovered initially. Spectral analysis of the square wave revealed a single peak at $\omega_j, j=13$, so that $\hat{\lambda}_1 = \frac{13}{250} = 0.104$. However, the variance time curve indicated two sources so that the other peak must be at a near zero frequency or greater than $\omega=\pi$. Converting the original series to time $cT$ where $c=2.0$ revealed two peaks at $\omega_j, j=6$ and 97. Then $c^{-1}\hat{\lambda}_1 = 0.048$ and $c^{-1}\hat{\lambda}_2 = 0.776$ so that $\hat{\lambda}_1 = 0.096$ and $\hat{\lambda}_2 = 1.552$. Once again $\hat{\lambda}_1 + \hat{\lambda}_2 = \hat{\lambda}_2$ and $\hat{\lambda}_1 - \hat{\lambda}_2 = \hat{\lambda}_1$. 
Since the frequency $\omega_j$; $j=6$ is very close to zero, then $c^{-1}\hat{\lambda}_1$ will not give a value of $\hat{\lambda}_1$ that is as accurate as that obtained in time $T$. To determine $\lambda_1$ and $\lambda_2$ the equations

\[
\lambda_1 - \lambda_2 = 0.104 \\
\lambda_1 + \lambda_2 = 1.552
\]

should be used. The constituent periods $B_1$ and $B_2$ can then be determined.

Spectral analysis of the square wave of example Sl in time $T$ revealed peaks at $\omega_j$ where $j=17, 32, 47, 78, 109,$ (Fig. B-16). Peaks were found at $j=8, 16, 24, 39, 54, 63, 79$ for the series in time $cT$ with $c=2.0$ (Fig. B-18(a)) and at $j=7, 66$ for time $cT$ with $c=0.25$ (Fig. B-18(b)). The corresponding values of $\hat{\lambda} = \omega_j/\pi$ (n=250) were

<table>
<thead>
<tr>
<th>$j$</th>
<th>$c^{-1}\hat{\lambda}$</th>
<th>$\hat{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.056</td>
<td>0.014</td>
</tr>
<tr>
<td>66</td>
<td>0.528</td>
<td>0.132</td>
</tr>
<tr>
<td>32</td>
<td>-</td>
<td>0.256</td>
</tr>
<tr>
<td>47</td>
<td>-</td>
<td>0.376</td>
</tr>
<tr>
<td>78</td>
<td>-</td>
<td>0.624</td>
</tr>
<tr>
<td>109</td>
<td>-</td>
<td>0.872</td>
</tr>
<tr>
<td>63</td>
<td>0.504</td>
<td>1.008</td>
</tr>
<tr>
<td>79</td>
<td>0.632</td>
<td>1.264</td>
</tr>
</tbody>
</table>
Fig. B-18(a) Spectral density function of the square wave for example S1 in time 2T.

Fig. B-18(b) Spectral density function of the square wave for example S1 in time T/4.
Then

\[ \hat{\lambda}_1 = 0.014 \approx \left| \lambda_1 - \lambda_2 - \lambda_3 + \lambda_4 \right| = 0.015 \]
\[ \hat{\lambda}_2 = 0.132 \approx \lambda_1 - \lambda_2 + \lambda_3 - \lambda_4 = 0.131 \]
\[ \hat{\lambda}_3 = 0.256 \approx \left| \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 \right| = 0.255 \]
\[ \hat{\lambda}_4 = 0.376 \approx \lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 = 0.375 \]
\[ \hat{\lambda}_5 = 0.624 \approx \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 = 0.625 \]
\[ \hat{\lambda}_6 = 0.872 \approx \lambda_1 + \lambda_2 - \lambda_3 + \lambda_4 = 0.869 \]
\[ \hat{\lambda}_7 = 1.008 \approx \lambda_1 + \lambda_2 + \lambda_3 - \lambda_4 = 1.013 \]
\[ \hat{\lambda}_8 = 1.264 \approx \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1.257 \]

It appears that for \( N \) periodic sources, \( N \) small, with periods \( B_1 < B_2 < \ldots < B_N \) and rates of occurrence \( \lambda_1, \lambda_2, \ldots, \lambda_N \) respectively, then spectral analysis of the square wave of the pooled output results in the occurrence of peaks of \( f(\omega) \) at frequencies \( \omega_1, \omega_2, \ldots \). If \( \hat{\lambda}_i = \omega_i / \pi \) then each \( \hat{\lambda}_i \) is a function of \( \lambda_1, \lambda_2, \ldots, \lambda_N \) and \( \hat{\lambda}_i \) is equal to the absolute value of a linear combination of the \( \lambda_j, j=1, \ldots, N; \)

\[ \hat{\lambda}_i = |a_1 \lambda_1 + a_2 \lambda_2 + \ldots + a_N \lambda_N| \]

where \( a_1 = 1 \) and \( a_k = \begin{cases} +1 & \text{for } k=2, \ldots, N, \\ -1 & \end{cases} \)

There are \( 2^{N-1} \) possible combinations of \( a_1, \ldots, a_N \) which accounts for the occurrence of \( 2^{N-1} \) peaks of \( f(\omega) \) at \( \omega_1, \omega_2, \omega_3, \ldots, \omega_{2^{N-1}} \); each frequency corresponding to a \( \hat{\lambda}_i \), \( i=1, \ldots, 2^{N-1} \).

Using the variance time curve, the number \( N \) of periodic sources can be determined and spectral analysis of
the square wave could be used to locate the $2^{N-1}$ peak frequencies. Then a series of equations would have to be solved for by ordering the $\hat{\lambda}_i, i=1, \ldots, 2^{N-1}$. With $N=4$ and 8 peak frequencies, then by ordering and relabelling $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \ldots \leq \hat{\lambda}_8$ where

\[
\hat{\lambda}_8 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \\
\hat{\lambda}_7 = \lambda_1 + \lambda_2 + \lambda_3 - \lambda_4 \\
\hat{\lambda}_6 = \lambda_1 + \lambda_2 - \lambda_3 + \lambda_4 \\
\hat{\lambda}_5 = \begin{cases} 
\lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 \\
\text{or } \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 
\end{cases}
\]

This yields two sets of possible equations, one of which is linearly dependent so that another equation has to be substituted for $\hat{\lambda}_5 = \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4$. As $N$ increases the number of possible sets of equations also increases, and the autocovariance function tends to an exponential function.

If the $\hat{\lambda}_i, i=1, \ldots, 2^{N-1}$ are ordered and relabelled (if necessary) so that $\hat{\lambda}_1 < \hat{\lambda}_2 < \ldots < \hat{\lambda}_{2^{N-1}} < \hat{\lambda}_{2^N}$, then

\[
\lambda_N = \frac{\hat{\lambda}_{2^{N-1}} - \hat{\lambda}_{2^{N-1}-1}}{2}
\]

and

\[
\lambda_{N-1} = \frac{\hat{\lambda}_{2^{N-1}} - \hat{\lambda}_{2^{N-1}-2}}{2}
\]

All that is necessary is to determine the three largest frequencies and to use these to calculate estimates of $\lambda_N$ and $\lambda_{N-1}$, and then $B_N$ and $B_{N-1}$. Then by calculating the jth order differences in interval lengths $X_{ij}$ where
\[ X_{ij} = X(t_{i+j}) - X(t_i) \quad i=1, 2, \ldots, n-j \quad j<n \]
\[ j=1, 2, \ldots \]
and comparing these with the estimates of \( B_N \) and \( B_{N-1} \) the exact periods and occurrences corresponding to the time series with these periods can be determined. Considering example 51

\[
\lambda_N = \frac{1.264 - 1.008}{2} = 0.128
\]

\[
\lambda_{N-1} = \frac{1.264 - 0.872}{2} = 0.196
\]

\[
B_N = \frac{1}{\lambda_N} = 7.81
\]

\[
B_{N-1} = \frac{1}{\lambda_{N-1}} = 5.10
\]
and the differences \( X_{ij} \) were calculated to determine the periods exactly. The 6th order differences \( X(t_{i+6}) - X(t_i) \) located a period of length 5.15, part of the difference table is given in TABLE B-1.

<table>
<thead>
<tr>
<th>( X_{ij} ) ( j=6 )</th>
<th>OCCURRENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.85</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>( X(t_{57}) - X(t_{51}) )</td>
</tr>
<tr>
<td>5.67</td>
<td></td>
</tr>
<tr>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>( X(t_{63}) - X(t_{57}) )</td>
</tr>
<tr>
<td>4.33</td>
<td></td>
</tr>
<tr>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>( X(t_{69}) - X(t_{63}) )</td>
</tr>
<tr>
<td>4.75</td>
<td></td>
</tr>
</tbody>
</table>
The periodic series with period 5.15 can then be eliminated from the pooled process. Although the estimate of $B_4 = 8.23$ is not very good this could be improved. Since $\hat{\lambda}_8 = 1.264$ the original series could be converted to time $cT$ where $c = 1.5$, yielding a more accurate estimate of $\hat{\lambda}_8$. Even this is not necessary since one periodic source has been eliminated from the pooled series. Spectral analysis of the deleted pooled series will only locate $2^{N-2}$ peak frequencies now, and the remaining periods can be found by working with successive deleted series.

It appears that a model to describe the pooled output of several periodic processes would be helpful, but as yet no such model has been found in the available journals. A model to describe the special case of two pooled periodic processes is now given. Future analysis in this direction might prove rewarding.

Let $X_1(t)$ be a periodic point process with period $B_1$ and $X_2(t)$ a periodic point process with period $B_2$, where $B_1 < B_2$ and $B_1$, $B_2$ are mutually irrational. Let $X(t)$ be the point process obtained by pooling $X_1(t)$ and $X_2(t)$. In the case of a pooled output of two periodic point processes, it has been observed that there is a pattern to the occurrence of the interval length equal to the smallest period $B_1$.

For example consider $X_1(t)$, $X_2(t)$ defined at $t_1$, $t_2$, ... such that

$$X_1(t_j) = j \cdot B_1$$

$$X_2(t_j) = j \cdot B_2$$

for $j = 1, 2, ....$. 
When $B_1 = 2.00$, $B_2 = 2.27$ and $X_i = X(t_{i+1}) - X(t_i)$, $i = 1, 2, \ldots$

$X_{15} = 2.00$
$X_{30} = 2.00$
$X_{47} = 2.00$
$X_{62} = 2.00$
$X_{79} = 2.00$
$X_{94} = 2.00$
$X_{109} = 2.00$

The first occurrence of $X_i = B_1$ is $X_{15}$, the next $X_{30}$ etc. so that if $n_1$ and $n_2$ are the number of occurrences of $X_1(t)$ and $X_2(t)$ between occurrences of $X_i = B_1$ then the following is observed:

<table>
<thead>
<tr>
<th>OCCURRENCE OF $B_1$</th>
<th>INTERVAL</th>
<th>$n_1$</th>
<th>$n_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_{15}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>$X_{30}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>$X_{47}$</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>$X_{62}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>$X_{79}$</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>$X_{94}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>$X_{109}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>$X_{126}$</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>$X_{141}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>$X_{158}$</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>
The Kth occurrence of B₁ is the interval length X_K(n₁+n₂)+m
where m is a non-negative integer. It appears that some form
of periodicity exists which determines the index i=K(n₁+n₂)+m
of the Kth interval length with Xᵢ=B₁

X₁(tⱼ) = j·B₁ is the time of the jth occurrence of B₁ and
X₂(tⱼ) = j·B₂ is the time of the jth occurrence of B₂.

Consider \( \bar{X}_i = X_2(t_i) - X_1(t_i) \)
Then \( \bar{X}_i = i·B₂-i·B₁ = i·(B₂-B₁) \) so that \( \{\bar{X}_i\} \) is an increasing
sequence of interval lengths, which increases by a factor of
B₂-B₁.

When \( \bar{X}_i>B₁ \) then \( i(B₂-B₁)>B₁ \)
\( i\cdot B₂>(i+1)\cdot B₁ \)
\( X_2(t_i)>X_1(t_{i+1}) \)

Since B₁<B₂, then \( X_1(t_i)<X_2(t_i) \)

For \( \bar{X}_i>B₁, \)
\( X_1(t_i)<X_1(t_{i+1})<X_2(t_i) \)
and for \( \bar{X}_i<B₁, \)
\( X_2(t_{i-1})<X_1(t_i). \)

Then the first occurrence of an interval length equal to B₁
occurs when \( \bar{X}_i>B₁, \)
\( i(B₂-B₁)>B₁ \)
\( i>B₁/(B₂-B₁). \)
Let $\alpha_2 = B_1/(B_2-B_1)$. Then $i_{1(2)} = [\alpha_2]$ is the smallest integer which denotes the number of interval lengths from $X_2(t)$ before the first interval of length $B_1$. The number of interval lengths from $X_1(t)$ for an interval of length $B_1$ will be the largest integer $j$ satisfying $j \cdot B_1 < i \cdot B_2$.

Then

$$j \cdot B_1 < \left( \frac{B_1}{B_2-B_1} \right) \cdot B_1$$

$$j < \left( \frac{B_1}{B_2-B_1} \right) \cdot \frac{B_2}{B_1}$$

$$j < \frac{B_2}{B_2-B_1}$$

Let $\alpha_1 = B_2/(B_2-B_1)$. Then $i_{1(1)} = [\alpha_1]$ is the integer which denotes the number of interval lengths from $X_1(t)$ to get the first interval of length $B_1$.

Then $i_{1(1)} + i_{1(2)} = [\alpha_1] + [\alpha_2]$ is the number of interval lengths necessary to obtain the first interval length $X_1 = B_1$, $i = i_{1(1)} + i_{1(2)}$.

Since $\alpha_1 = B_2/(B_2-B_1) = 1 + B_1/B_2 + B_1^2/B_2(B_2-B_1)$ and $\alpha_2 = B_1/B_2 + B_1^2/B_2(B_2-B_1)$, then $\alpha_1 = \alpha_2 + 1$, and $[\alpha_1] = [\alpha_2 + 1] = [\alpha_2] + 1$.

The first occurrence of $X_1 = B_1$ occurs when $\bar{X}_i > B_1$ and $\bar{X}_i < 2B_1$. If $\bar{X}_i > 2B_1$ then $X_2(t_1) > X_1(t_{i+2})$ and the interval of length $B_1$ has occurred at least twice.

$$i(B_2-B_1) > 2B_1$$

$$i > \frac{2B_1}{B_2-B_1} = 2\alpha_1$$

so that $i_{2(2)} = [2\alpha_2]$ is the number of interval lengths from
the process \( X_2(t) \) before the second occurrence of \( X_i = B_1 \).

In general then the Kth occurrence of an interval of length \( B_1 \) will occur only when \( \bar{X}_i > K B_1 \) and \( i > \frac{K B_1}{B_2 - B_1} = K \alpha_2 \).

Then \( i^{(2)}_K = [K \alpha_2] \) denotes the number of occurrences of interval lengths from the process \( X_2(t) \) before the Kth occurrence of \( X_i = B_1 \) and \( i^{(1)}_K = [K \alpha_1] \) the number of interval lengths for \( X_i = B_1 \). The occurrence of the Kth interval length \( B_1 \) is

\[
X_i^{(1)}(n) + i^{(2)}_K = X[K \alpha_1] + [K \alpha_2].
\]

Since \( [K \alpha_1] + [K \alpha_2] \neq [K(\alpha_1 + \alpha_2)] \) in general then the occurrence of the interval of length \( B_1 \) is not periodic with period \( \alpha_1 + \alpha_2 = (B_1 + B_2)/(B_2 - B_1) \). However \( \alpha_1 \) and \( \alpha_2 \) can be written as \( \alpha_1 = n_1 + s \) and \( \alpha_2 = n_2 + s \) where \( n_1, n_2 \) are non-negative integers and \( 0 \leq s < 1 \). Then

\[
[K \alpha_1] + [K \alpha_2] = [K(n_1 + s)] + [K(n_2 + s)] = K(n_1 + n_2) + 2[Ks]
\]

\[
[K(\alpha_1 + \alpha_2)] = [K(n_1 + n_2 + 2s)] = K(n_1 + n_2) + [2Ks].
\]

Writing \( Ks = n + r \) where \( n \) is a non-negative integer and \( 0 \leq r < 1 \) then \( [2Ks] = 2n \) and \( [2Ks] = [2n + 2r] = 2n + [2r] \).

For all intents and purposes then, we can consider

\[
[K \alpha_1] + [K \alpha_2] = [K(\alpha_1 + \alpha_2)]
\]

so that the occurrence of \( B_1 \) is periodic with period \( (B_1 + B_2)/(B_2 - B_1) \).

Assume that the periods \( B_i, (i=1,\ldots,N) \) from periodic sources
are positive numbers and are mutually irrational in the sense that there exists no set of positive or negative integers \( n_i \), not all zero, such that \( \sum_{i=1}^{N} n_i B_i = 0 \). The generalized form of Weyl's theorem\(^3\) states that if \( \{a\} \) denotes the fractional part of \( a \) and if \( a_1, a_2, \ldots, a_K \) are irrational numbers themselves mutually irrational, then the sequences \( \{na_1\}, \ldots, \{na_K\} \) are independently uniformly distributed over \((0, 1)\) for \( n=1, 2, \ldots \).

For any pooled output of periodic sources the largest interval length possible is equal to the smallest period \( B_1 \). If this upper bound \( B_1 \) is removed then the sequence of interval lengths is uniformly distributed over the interval \((0, B_1)\).

For the case of the superposition of two periodic sources then the pooled series \( X(t) \) can be expressed as

\[
X(t) = Y(t) + Z(t)
\]

where \( Y(t) \) is a periodic point process with period \( (B_1 + B_2)/(B_2 - B_1) \) and \( Z(t) \) is a point process with the intervals between events having the uniform distribution over the interval \((0, B_1)\).

APPENDIX C

TO CALCULATE THE AUTOCOVARIANCES AND SPECTRAL DENSITY FUNCTION
OF THE SQUARE WAVE WITH VALUES W1 AND W2
TIME IS THE LENGTH OF THE TIME SERIES X(T), AND AUTOCOVARIANCES
WILL BE CALCULATED FROM 0 TO HMAX AT INTERVALS OF HK

TIME=370.43
HMAX=74.0
HK=0.30
W1=1.0
W2=-1.0

DIMENSION X(500), A(1000), B(1000), C(1000), D(1000), SDF(500,15),
V(1500), T(600)

READ IN THE N INTERVAL LENGTHS X(I).

N=250
READ(5,2) (X(I), I=1,N)
2 FORMAT(F8.3)
II=0
DUM=0.
DO 10 I=1,N+2
10 DUM=X(I)+DUM
SMEAN = (W1 * DUM + W2 * (TIME - DUM)) / TIME
VAR = (W1 * W1 * DUM + W2 * W2 * (TIME - DUM)) / TIME - SMEAN * SMEAN

WRITE (6, 15) SMEAN, VAR

15 FORMAT (1H-, 20X, 7H MEAN =, F10.5, 10X, 11H VARIANCE =, F10.5)

WRITE (6, 18)

18 FORMAT (1H-, 32X, 3H K =, 10X, 7H COV(K))

C
CALCULATE M1 = KP1 AND M2 = KP2

HH = HK

320 L2 = 0
RK = 0.
DO 20 I = 1, N
RK = X(I) + RK
IF (HH.GT.HMAX) GO TO 400
IF (HH.LT.RK) GO TO 30
IF (L2.EQ.0) GO TO 25
L2 = 0
GO TO 20
25 L2 = 1
20 CONTINUE

30 KP1 = I
KMB = N - KP1 + 1
B(I) = RK - HH
JJ = KP1 + 1
DO 40 J = JJ, N
K = J - KP1 + 1

40 B(K) = X(J)

SK = 0.

DO 80 I = 1, N
M = N - I + 1
SK = X(M) + SK
IF(HH .LE. SK) GO TO 90

80 CONTINUE

90 KP2 = M
A(KP2) = SK - HH
JJ = KP2 - 1
DO 100 J = 1, JJ

100 A(J) = X(J)

C     CALCULATE THE AUTOCOVARIANCES AND STORE IN D(II)

L1 = 0
J = 1
K = 1
Y1 = 0.
Y2 = 0.
Y1Y2 = 0.
R = A(J)
S = B(K)

200 Z = AMIN1(R, S)
IF(L1 .EQ. L2) GO TO 110
Y1Y2 = Y1Y2 + Z
GO TO 130
110 IF(L1.EQ.0) GO TO 120
    Y2=Y2+Z
    GO TO 130
120 Y1=Y1+Z
130 R=R-Z
    S=S-Z
    IF(R.EQ.0.) GO TO 140
    K=K+1
    IF(K.GT.KMB) GO TO 300
    S=B(K)
    IF(L2.EQ.0) GO TO 150
150 L2=1
    GO TO 200
140 J=J+1
    IF(J.GT.KP2) GO TO 300
    R=A(J)
    IF(L1.EQ.0) GO TO 160
160 L1=1
    GO TO 200
300 COV=Y1*W1*W1+Y2*W2*W2+Y1Y2*W1*W2
    COV=COV/(TIME-HH)-SMEAN*SMEAN*(1.0-HH/TIME)
TO COMPUTE ESTIMATES OF THE SPECTRAL DENSITY FUNCTION

IF WEIGHTING IS TO BE USED, TO COMPARE WITH THE FIRST ESTIMATE

USING THE WEIGHTS 1 - K/M WHERE M IS THE NUMBER OF LAGS, K=1,M

THEN PUT MAX=1. LAGS FROM 2 TO LMAX WILL BE DONE IN STEPS OF 2

MAX=1
LMAX=64
L=1
PIE=3.14159265
G=HMAX/HK
M=INT(G)
NN=126

DO 410 J=1,NN
   SUM=0.
   DO 420 I=1,M
      R=(FLOAT(J-1)*C(I)*2.0*PIE)/250.0
      S=COS(R)
      SUM=(D(I)*S)/VAR + SUM
   END DO

410 CONTINUE
420 CONTINUE
SDF(J,L) = (1.0+2.0*SUM)/(2.0*PIE)

410 CONTINUE
IF(MAX.LT.1) GO TO 440
IF(L.EQ.1) GO TO 490
IF(M.GE.LMAX) GO TO 440

490 M=8*L
L=L+1
GO TO 430

440 WRITE(6,450)

450 FORMAT(1H-,15X,4H J = ,20X,27H SPECTRAL DENSITY ESTIMATES)
WRITE(6,460)

460 FORMAT(1H-,30X,13H NO WEIGHTING,5X,15H WITH WEIGHTING)
DO 470 J=1,NN
JJ=J-1
470 WRITE(6,480)JJ, (SDF(J,I), I=1,L)

480 FORMAT(1H-,15X,14,12X,9F10.5)
STOP
END
TO COMPUTE THE MEAN, VARIANCE AND AUTO-COVARIANCES OF A PROCESS \( X(T) \)

\( N \) IS THE TOTAL NUMBER OF OBSERVATIONS OF \( X(T) \)

\( M \) IS THE NUMBER OF AUTO-COVARIANCES TO BE CALCULATED, \( M = N - 1 \) IS THE MAXIMUM NUMBER. \( \text{COV}(N) = \text{VARIANCE IN THIS PROGRAM} \)

DIMENSION \( X(1000) \), \( \text{COV}(1000) \), \( \text{SDF}(400,20) \), \( V(2000) \), \( T(500) \)

\( N = 250 \)

\( M = N / 5 \)

READ IN THE \( N \) INTERVAL LENGTHS OR OCCURRENCE RATES \( X(I) \).

READ(5,2) \((X(I), I=1,N)\)

2 FORMAT(F20.8)

SMEAN = 0.

DO 10 \( J = 1,N \)

10 SMEAN = SMEAN + X(J)

SMEAN = SMEAN / FLOAT(N)

DO 20 \( K = 1,M \)

NP = N - K

DUM = 0.

DO 30 \( J = 1,NP \)

I = J + K

30 DUM = DUM + (X(J) - SMEAN)*(X(I) - SMEAN)

20 COV(K) = DUM / FLOAT(NP)

VAR = 0.
DO 40 J=1,N
DUM=X(J)-SMEAN
DUM=DUM*DUM
40 VAR=DUM+VAR
COV(N)=VAR/FLOAT(N)
WRITE(6,50)
50 FORMAT(1H-,20X,12H THE MEAN IS,20X,16H THE VARIANCE IS,20X,20H THE
1 COVARIANCES ARE)
WRITE(6,60) SMEAN,COV(N)
60 FORMAT(1H-,10X,F20.8,10X,F20.8,30X,3H K=)
DO 65 K=1,M
65 WRITE(6,70) K,COV(K)
70 FORMAT(1H-,90X,I5,3X,F20.8)
C TO COMPUTE ESTIMATES OF THE SPECTRAL DENSITY FUNCTION
C IF WEIGHTING IS TO BE USED, TO COMPARE WITH THE FIRST ESTIMATE
C USING THE WEIGHTS 1 - K/M WHERE M IS THE NUMBER OF LAGS, K=1,M
C THEN PUT MAX=1 , LAGS FROM 2 TO LMAX WILL BE DONE IN STEPS
C OF 2
MAX=1
LMAX=16
L=1
PIE=3.14159265
NN=N/2+1
130 DO 110 J=1,NN
SUM = 0.
DO 120 K=1,M
R = (FLOAT(J-1)*FLOAT(K)*2.0*PIE)/FLOAT(N)
S = COS(R)
SUM = (COV(K)*S)/COV(N) + SUM
IF(L.LT.2) GO TO 120
SUM = (1.0-FLOAT(K)/FLOAT(M))*SUM
120 CONTINUE
SDF(J,L) = (1.0+2.0*SUM)/(2.0*PIE)
110 CONTINUE
IF(MAX.LT.1) GO TO 140
IF(L.EQ.1) GO TO 190
IF(M.GE.LMAX) GO TO 140
190 M=2*L
L=L+1
GO TO 130
140 WRITE(6,150)
150 FORMAT(1H-,15X,4H J= ,20X,27H SPECTRAL DENSITY ESTIMATES)
WRITE(6,160)
160 FORMAT(1H-,30X,13H NO WEIGHTING,5X,15H WITH WEIGHTING)
DO 170 J=1,NN
JJ=J-1
170 WRITE(6,180)JJ, (SDF(J,I), I=1,L)
180 FORMAT(1H-,15X,14,12X,9F10.5)
BIBLIOGRAPHY


