DIGITAL PROCESSING OF NON-STATIONARY SIGNALS

by

© JESSE FRANCISCO HUNG, B.Eng.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Engineering

McMaster University
June 1981

DIGITAL PROCESSING OF NON-STATIONARY SIGNALS

MASTER OF ENGINEERING (1981)
(Electrical & Computer Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario.

TITLE : Digital Processing of Non-Stationary Signals

AUTHOR : Jesse Francisco Hung, B.Eng. (McMaster University)

SUPERVISOR : Dr. S. K. Sarna

NUMBER OF PAGES : x, 122

ABSTRACT

Various tests had been conducted in order to examine the stationarity and the normality characteristics of electrical control signals of the digestive tract. They were done because the conventional frequency analysis, which is used extensively for the investigation of biological signals, usually assumes the signal to be stationary and normally distributed. The validity of this assumption should then be examined before any further analysis is applied. The tests are conducted by proposing a null hypothesis that the signal under investigation is stationary and normally distributed. It was found that the percentage of rejection of the hypothesis increases towards the colonic end of the tract.

Since conventional power spectral analysis does not provide any phase information on a non-stationary signal, the bispectral analysis, which is the Fourier transform of the third moment, was used in order to examine many of the still unknown characteristics of the gastrointestinal signal. The analysis mainly searches for any phase-locking between frequency components and hence identifies the generators of the signal. Two seperate analyses had been done : one was for a single channel and the other was for a double channel of signals. It was found that the same group of generators for the electrical signals on the upper

part of the tract is present most of the time. But short-lived and locally based oscillators dominate the functions in the colon. From the cross-bispectral analysis, it was found that the generators in the stomach and the duodenum usually exert driving force to the distal site but bidirectionally in the jejunum. In the colon, only independent frequency components were found to be phased-locked occasionally.

In conclusion, the analyses carried out in this study provide some alternate means to investigate many of the still largely unknown signals.

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to Dr. S. K. Sarna for his guidance and helpful suggestions and assistance throughout the course of this work.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Biomedical signals, such as EEG, gastroduodenal and colonic electrical activities are regarded as the stochastic phenomena of biological systems. In practice, statistical parameters like histogram, autocorrelation function, power spectrum and coherence function are calculated for various purposes. But these are done usually under the condition of weak stationarity of the signal, that is, the mean value, the standard deviation and the autocorrelation function have to be independent of time translation. However impulse response and evoked potential of biomedical data are generally not regarded as stationary time series. Therefore other analysis is required. So far only a few studies have been done in analyzing non-stationary signals. For instance, Priestley and Rao {19} and Priestley and Tong {20} have used the evolutionary spectra, that is, spectral functions which are time dependent, to describe the local energy distribution of a non-stationary series. Kawabata {14} has also done a similar study. He calculated the instantaneous power spectra of certain transition processes in the EEG and described them on the time-frequency plane. Pinson and Childers {17} utilized the frequency-wave number spectral analysis to investigate the high-resolution vector velocity so that the direction and speed of propagating wavefront of the EEG can be described. Additionally Praetarius, Bodenstein and Creutzfeldt {18} and Bodenstein and Praetarius {3} have used the linear prediction filtering method to divide the time series into segments and extract the transients from the data. If two channels of signal are investigated, and the transients are detected simultaneously on both channels by the method, one can say that the sharp wave detected in channel 1 is real by comparing it with its clearly discernible counterparts in channel 2. All the papers mentioned above suggest, in one way or another, some methods to examine the characteristics of a non-stationary signal. But the results obtained from these methods only guide one to scan some of the characteristics so that many of the features of a non-stationary signal are still left unknown. Hence it is difficult to draw a meaningful conclusion from this limited information.

In this project additional attempts will be made in order to extract some more information which is lost during the ordinary power spectral analysis. They are actually based on the papers {6}, {7} and {11} by Dumermuth, Gasser, Huber and Kleiner. As already mentioned above, only partial characteristic of a non-stationary signal can be obtained by the conventional spectral analysis. By definition, the power spectrum suppresses all the phase information hidden in a signal. Although it may suggest phase information by showing peaks at harmonically related frequencies, information about interrelationships between frequency bands outside the peak frequencies of the spectrum will be lost. Hence a method called the bispectral analysis is employed here to extract this particular information and provide a more in-depth observation of the signal being analyzed. A bispectrum allows one to analyze in detail about the third central moment, which is the mean cube and an order higher than that of a conventional power spectrum, of a random signal. It is essentially influenced by the phase-locking characteristic between different frequency components. This method will be utilized in this project to analyze signals recorded from the various parts of the digestive system. Not just the interrelations between different frequency components of a single channel of signal will be examined, but that of two channels of signal is also calculated in order to make attempts in analysing relationships between the two signals. Before these procedures are carried out, tests will be conducted first of all about the stationarity behaviour of the signal being used. This is done in the next chapter.

CHAPTER 2

STATIONARITY AND NORMALITY TESTS

2.1 INTRODUCTION

Statistical properties of stochastic processes, such as biomedical data, are often examined for various purposes. However most of the statistical analyses processed in the time and frequency domains are based on the assumption that the time series is weakly stationary and normally distributed. For instance, regression analysis or multivariate analysis of biomedical data is mostly based on the theory of normal distributions. Therefore normality is needed before the multivariate method is employed to analyze the data. Furthermore, data is often collected under circumstances which do not permit an assumption of stationarity based upon simple physical considerations. Hence, as the basis of the statistical analysis of the data, it is necessary to know whether it satisfies the conditions of weak stationarity and normality.

The statistics of a time series are mostly obtained by the observed data of random functions. A random function is usually specified when all its probability distribution functions are given. Hence it is said to be stationary if the probability distribution functions are independent of time. But the probability distribution functions are quite difficult to measure. It will be sufficient to say that the time series is weakly stationary if the first two moments of the probability distribution are independent of time.

In most cases, the stationarity of the data must be evaluated by studies of available sample time records. This evaluation might range from a visual inspection of the time record to detailed statistical tests of certain appropriate data parameters. It can be assumed that any non-stationarity of interest will be revealed by time trends in the mean square value of the data. This is because the mean square value Ψ^2 will usually reveal a time varying autocorrelation by the relation $R(0) = \Psi^2$. The procedures involved in investigating a single record $X(t)$ for its stationarity will be as follows {1} :

(1) divide the sample series into N equal intervals;

(2) calculate a mean square value for each interval and align these values in sequence as follows

$$\bar{X}_1^2,\ \bar{X}_2^2,\ \bar{X}_3^2,\ \ldots\ ,\ \bar{X}_N^2$$

(3) test the sequence of mean square values for the presence of underlying trends.

Step (3) in the above procedure can be accomplished in many ways. For example, one can use the Kolmogorov-Smirnov's statistic D_2 to examine the stationarity if the sampling distribution is known. But it requires a detailed knowledge of the frequency composition of the data. Such knowledge is generally not available at the time when one wants to know whether or not the data is stationary. Hence a non-parametric approach will be more appropriate in this case because it does not require a knowledge of the sampling distributions of data parameters. One such non-parametric test which is applicable to this problem is the run test. It is employed in this project to test the stationarity characteristic of a sample time series and will be explained in detail later in this chapter.

In many cases, researchers often assume that a time series is the realization of a gaussian, normal random process. A random process is said to be normal if any set of data points, $\{X(t) \mid \text{a set of } t\}$ is jointly normally distributed. Hence, for a normal random process, at least the marginal distribution of the data values, or the histogram, must be normal. In this project, the normality of X(t) will be tested by the Chi-Square Goodness-Of-Fit method which measures the deviation of the $\{X(t)\}$ distribution from the normal distribution. The general procedure involves the use of a statistic with an approximate Chi-Square distribution as a measure of the discrepancy between an observed probability density function and the theoretical density function with normal distribution. Upon studying the sampling distribution of this statistic, a hypothesis of equivalence is then examined in order to test the degree of deviation.

In addition to the Chi-Square Goodness-Of-Fit test, the skewness and kurtosis of X(t) are also calculated by utilizing the third and fourth moments. Statistics g_1 and g_2 are calculated from the Fisher's k statistics. They measure whether the distribution is symmetric and how much the symmetric distribution deviates from the normal one. This Fisher's normality test is used as an enhancement of the Chi-Square Goodness-Of-Fit test employed here first.

All these stationarity and normality tests will be implemented on the NOVA 830 minicomputer. Data files are created beforehand by sampling the electrical signal from various parts of the digestive system. The stationarity and normality of the data will be tested by the appropriate programs and the result will be outputted on the line-printer. Before going into the details of the theory and algorithms for the tests mentioned above, two statistical terms, the confidence interval and the null hypothesis test, should be explained here because they will be used as the criteria to distinguish between stationary or non-stationary (as well as normally or not normally distributed) time series.

2.2 STATISTICAL TERMS

2.2.1 CONFIDENCE INTERVAL

In practice, the estimation of parameters of random variables by the use of sample values always produces some sort of uncertainties. Different values of estimated parameters will arise with different samples and no indication is provided as to how closely a sample value estimates the parameter. Hence it would be more meaningful if one estimates the parameters of the random variables, with a known degree of uncertainty, by calculating an interval instead of a single point value. For instance, let $\bar{X}$ be the sample mean of N independent observations of a random variable X and is being used as an estimate of the true mean μ_x of the random variable. It is usually more desirable to estimate μ_x by an interval $\bar{X} \pm D$ so that there are some uncertainties that the mean μ_x will fall within that interval. Consider the equation

$$\text{prob}\left\{ Z_{1-\alpha/2} < \frac{(\bar{X} - \mu_x)\sqrt{N}}{\sigma_x} < Z_{\alpha/2} \right\} = 1 - \alpha \qquad (2.1)$$

It states that if different values of $\frac{(\bar{X}-\mu_x)\sqrt{N}}{\sigma_x}$ are computed from different samples being collected, one would expect that about $1 - \alpha$ of calculated values will fall within the indicated interval. Usually the value of α is small, say 0.05. Hence with a small degree of uncertainty, one can expect to find the computed value of $\frac{(\bar{X}-\mu_x)\sqrt{N}}{\sigma_x}$ to fall within the interval and it is called the confidence interval.

2.2.2 HYPOTHESIS TESTS

Let $\hat{\Phi}$, calculated from a sample of N independent observations of a random variable X, be an estimator of the parameter Φ. It is also hypothesized that Φ is equal to a particular value, say Φ_o. Since different samples will give different values of $\hat{\Phi}$, one would like to know how much difference between $\hat{\Phi}$ and Φ_o should occur before the hypothesis $\Phi = \Phi_o$ becomes invalid. Consider the probability of any noted difference between $\hat{\Phi}$ and Φ_o. If the probability of a particular difference is small and it actually occurs, then it would be considered significant and the hypothesis should be rejected. On the other hand, if the probability is not small, the hypothesis $\Phi = \Phi_o$ would be accepted. The procedure mentioned above is a brief outline of the hypothesis test.

Let the probability density function of the estimator be $p(\hat{\Phi})$. Then the mean value of $p(\hat{\Phi})$ would be Φ_0 if it is hypothesized that $\Phi = \Phi_0$. The probability that $\hat{\Phi}$ would fall below the level $\Phi_{1-\alpha/2}$ is

$$\text{prob } \{ \hat{\Phi} < \Phi_{1-\alpha/2} \} = \int_{-\infty}^{\Phi_{1-\alpha}} p(\hat{\Phi})\, d\hat{\Phi} = \alpha / 2 \qquad (2.2)$$

The probability that $\hat{\Phi}$ would fall above the level $\Phi_{\alpha/2}$ is

$$\text{prob } \{ \hat{\Phi} > \Phi_{\alpha/2} \} = \int_{\Phi_{\alpha/2}}^{\infty} p(\hat{\Phi})\, d\hat{\Phi} = \alpha / 2 \qquad (2.3)$$

This is illustrated in Fig.2.1. So the probability that the estimator falls outside the indicated confidence interval is α. If α was small and the computed value of $\hat{\Phi}$ from a collected sample fell outside the interval between $1-\alpha/2$ $\alpha/2$, it would be reasonable to question the original hypothesis of $\Phi = \Phi_0$ since it is unlikely to have such a value of $\hat{\Phi}$ if the hypothesis was true. So the hypothesis should be rejected in this case. However if the calculated value of $\hat{\Phi}$ did fall within the interval, the original hypothesis would then be true and should be accepted. In the subsequent tests for stationarity and normality, either a one-sided or a two-sided null hypothesis would be used as a criterion to differentiate whether the series under test is random gaussian or not. More details about the hypothesis will be described in the stationarity and normality tests.

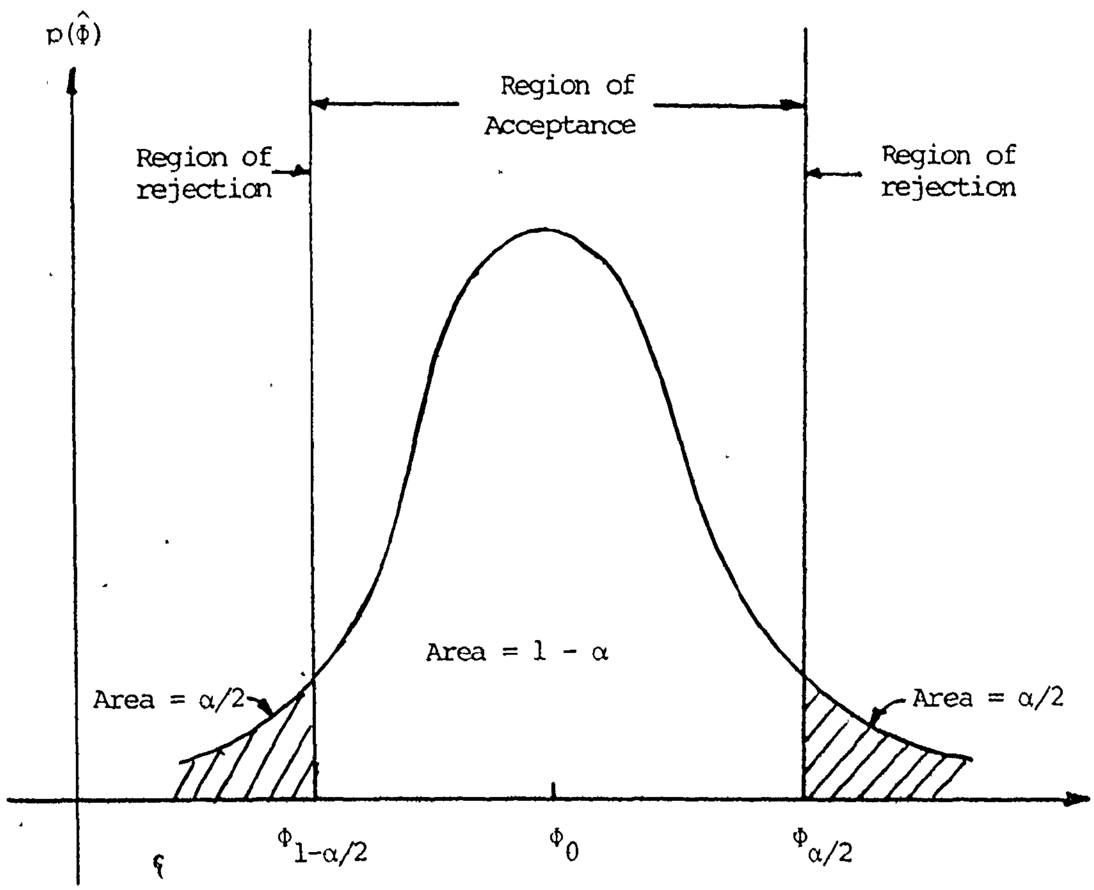


Fig. 2.1 Acceptance and rejection regions for hypothesis tests. {1}

2.3 THE RUN TEST - A STATIONARITY TEST

A weakly stationary time series has the property that the first two orders of statistics such as the mean, variance and the autocorrelation are independent of time. Hence any display of trends in the data will violate the character of stationarity. The RUN test employed in this thesis is designed to detect the existence of any trend in the data.

Consider a sequence of N independent observations of a random variable X and divide it into L = 2M segments. Assume that any non-stationarity of interest will be revealed by time trends in the mean square value of the data. The stationarity through the autocorrelations will not be considered here because if the variance of the time series is time invariant, it is very unlikely that the auto-covariance is time-dependent. Therefore, for simplicity, only the mean square value is examined. Now consider a sequence of L mean square values (S(i),i=1,2,...,L) calculated from each segment. Assign each element in the sequence into one of the two catagories by using the median T of the mean square value sequence as the differentiating criterion. If S(i) $\geq$ T, a '+' is assigned to it and a '-' is assigned if S(i) $<$ T. Then a sequence of +'s and -'s is created as follows :

++--++-+++-+--+---

A run in a sequence of symbols is then defined as a group of consecutive symbols or signs of one kind preceded and followed by symbols of another kind or by no symbols at all. In the sequence of +'s and -'s shown above, there are 5 runs of + as well as 5 runs of -. The number of runs which occurs in the sequence gives an indication as to whether the original time series has any trend or not. It is based on the fact that an unusually large or small number of total runs would suggest a lack of randomness. For instance, let the median be used as the differentiating criterion, the number of runs for + be r_a and the number of runs for - be r_b. If an upward trend is present in the mean square value sequence, the -'s will tend to come at the beginning and the +'s at the end of the sequence which results in a small number of total runs $r = r_a + r_b$. Another example is that if certain kinds of dependence are causing values in the sequence bouncing systematically back and forth from one side of the median to the other, an unusually large number of runs will result.

Consider the case mentioned above where there is a sequence of $L = 2M$ elements of mean square values and the median T is used as the differentiating criterion. Let the number of runs for the + be r_a and the - be r_b with $r = r_a + r_b$. Under the hypothesis of randomness there are (2M)! ways to arrange the sequence of M +'s and M -'s with each arrangement being distinct and equally likely. Hence the probability of a given configuration is equal to the ratio of the number of arrangements having that configuration to the total number of arrangements ($\binom{2M}{M}$). Since there are equal number of +'s and −'s, only three cases have to be considered, i.e., 1) $r_a = r_b + 1$; 2) $r_b = r_a + 1$ and 3) $r_a = r_b$.

In case 1), there are r_a distinct groups of +'s so that the −'s can be inserted into $r_a - 1$ slots. Since there are M +'s altogether, the number of ways that M −'s can be put in M − 1 slots of +'s and form $r_a - 1$ groups is ($\binom{M-1}{r_a - 1}$). Furthermore there are r_b groups of −'s, the −'s can also be arranged in ($\binom{M-1}{r_b - 1}$) ways. As a result, if $r_a = r_b + 1$, there are ($\binom{M-1}{r_a - 1}$)($\binom{M-1}{r_b - 1}$) ways to accomplish the arrangement. For case 2), one can just simply interchange r_a with r_b and the result will stay the same as that for case 1).

For case 3), the number of ways is again ($\binom{M-1}{r_a - 1}$)($\binom{M-1}{r_b - 1}$). But there are two possible arrangements: one is when the sequence starts with a + and ends with a − and the second one is the vice versa case of the first. Hence the total arrangements for case 3) is 2($\binom{M-1}{r_a - 1}$)($\binom{M-1}{r_b - 1}$).

Now if $r = r_a + r_b = 2k + 1$, two cases have to be considered, that is 1) and 2) mentioned above. The probability for the total number of runs is then

$$p(r) = \begin{cases} 2\binom{M-1}{k-1}\binom{M-1}{k-1} \Big/ \binom{2M}{M} & \text{, if } r = 2k \\ \left\{\binom{M-1}{k}\binom{M-1}{k-1} + \binom{M-1}{k-1}\binom{M-1}{k}\right\} \Big/ \binom{2M}{M} & \\ = 2\binom{M-1}{k}\binom{M-1}{k-1} \Big/ \binom{2M}{M} & \text{, if } r = 2k+1 \end{cases} \quad (2.4)$$

Now let it be hypothesized that there is no trend by assuming that the sequence of L mean square values are independent sample values of the same random variable. The hypothesis can then be tested at any significant level α by comparing the total runs to the interval $r_2 < r < r_1$. If the sequence is not stationary, r, the total number of runs, would be either too small or too large and r will lie outside the interval indicated above. So one can conclude whether the series is stationary or not by finding the values of r_1 and r_2 from equation (2.5) below :

$$\sum_{r=2}^{r_2} p(r) = \alpha / 2$$

$$\sum_{r=r_1}^{2M} p(r) = \alpha / 2 \quad . \qquad (2.5)$$

The test is implemented on the NOVA computer under the program named 'RUNTST'. Here the record length of the signal and the number of sections 'NSEG' to be divided from the whole record are set by the user at the beginning of the procedure. The computer will examine the stationarity of each section and find out how stationary the whole record is by calculating the percentage of rejection of the stationary hypothesis. Each section will be divided into 30 segments (that is, $L = 2M = 30$). The run pattern and the number of runs r (or 'NRUN' in the algorithm) will then be searched. If $r_1 < r < r_2$, the section is stationary. If r lies outside the interval between r_1 and r_2, it is non-stationary. After all the sections have been analyzed, the percentage of rejection is calculated and outputted. The detailed algorithm is illustrated in Fig.2.2.

Fig.2.3 shows a typical colonic signal recorded from a dog. The pictorial illustration explains how the record is divided into NSEG sections and L (=30) segments. Fig.2.4 presents the output from the analysis of the colonic signal shown in Fig.2.3. From the large value of the percentage of rejection one can conclude that the signal under investigation is by no means stationary and hence some informations will be lost if conventional power spectral analysis is used to examine the signal. Before the bispectral analysis is used later in this project, each signal will be investigated by 'RUNTST' and the result will be presented in the appropriate location of this thesis.

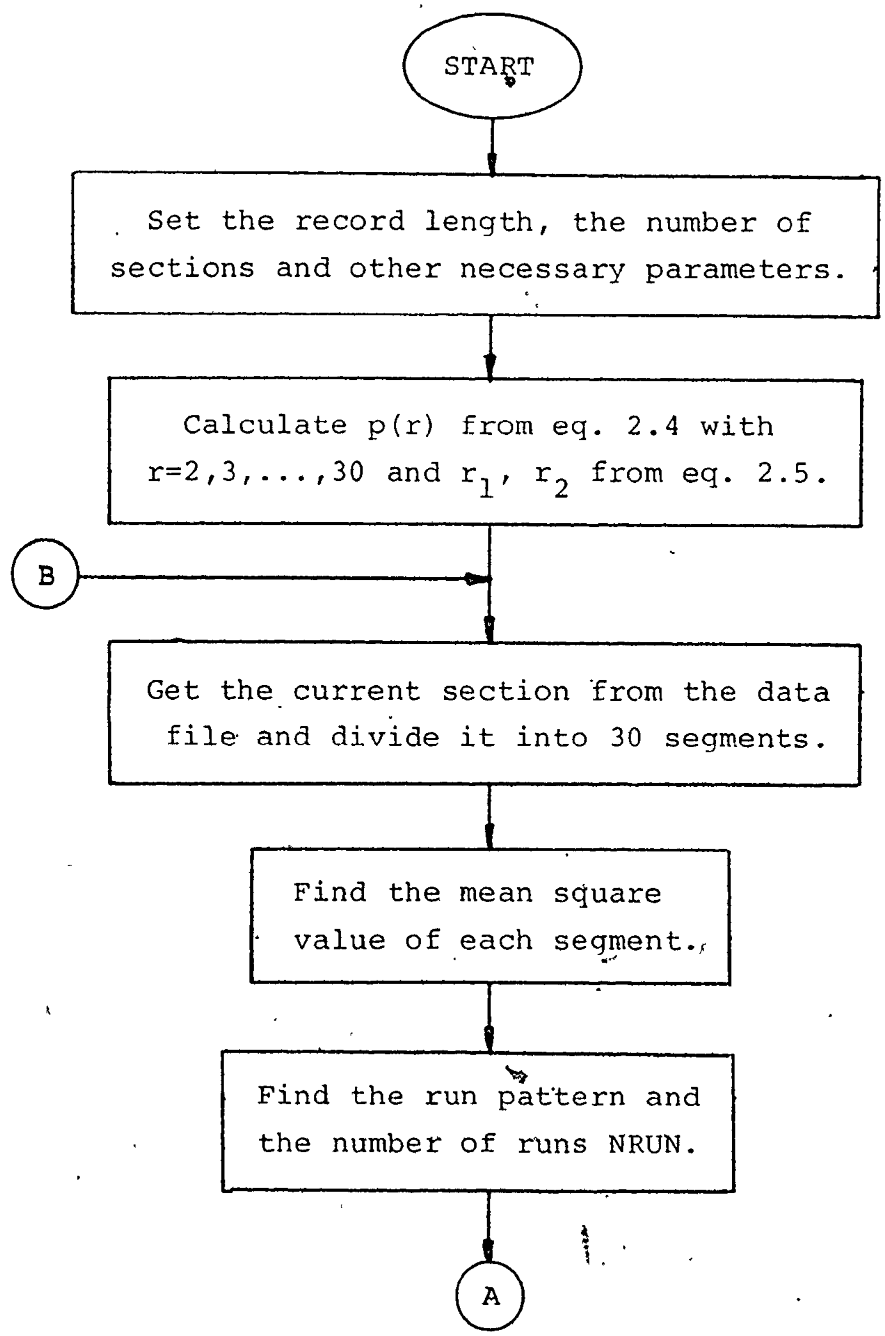


Fig. 2.2 The flow diagram of the RUN test.

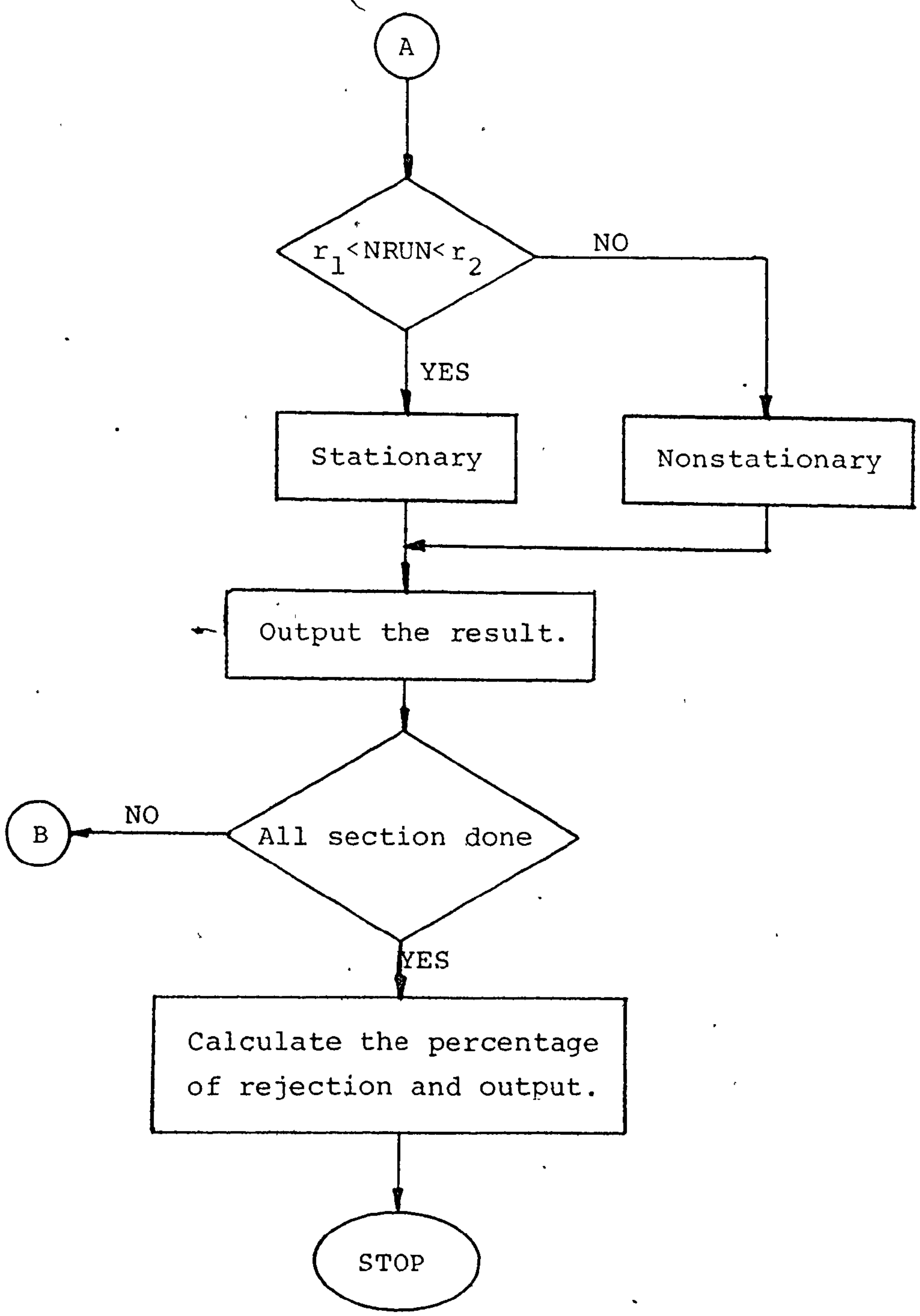


Fig. 2.2 (continued)

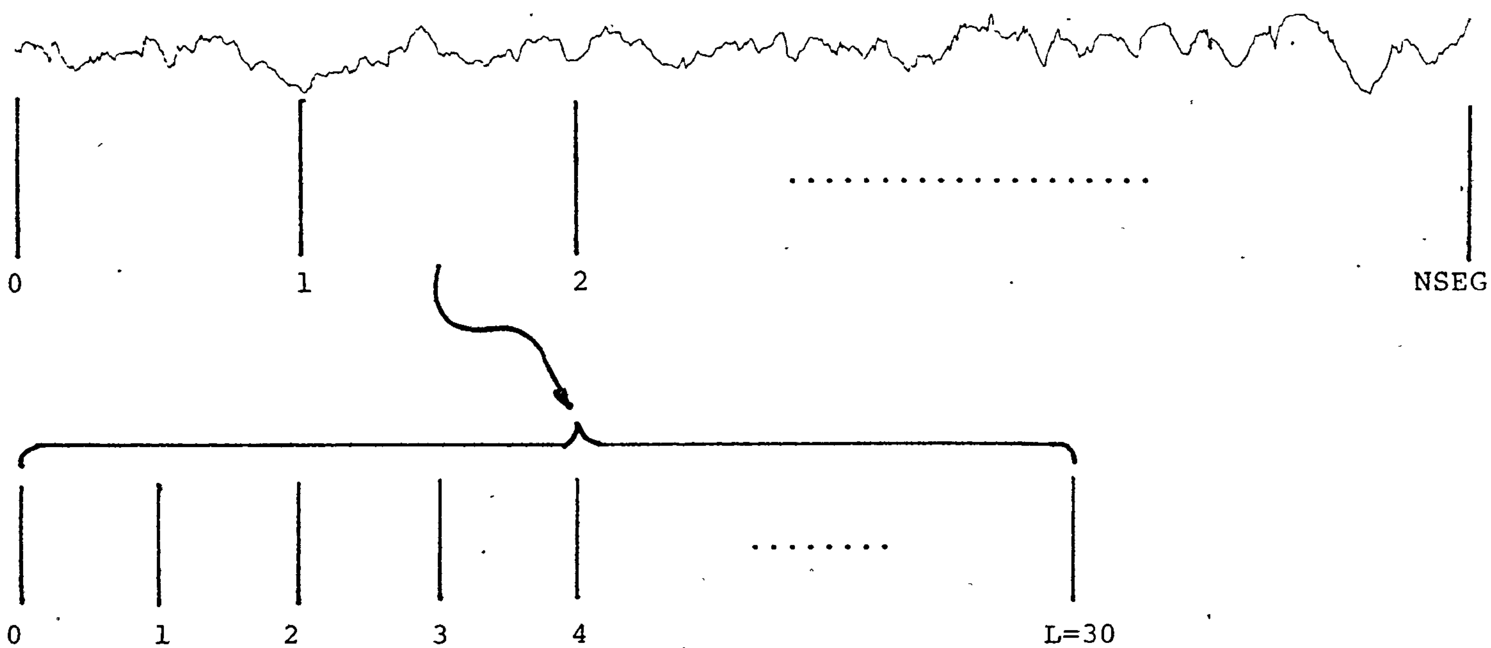


Fig. 2.3 A typical section of electrical signal taken from the colon. Here the method of division used in the RUN test is also shown.

```
CHANNEL NUMBER =  3
INITIAL RECORD NUMBER =    1
NUMBER OF DATA POINTS PER SECTION  =  120

THE NULL HYPOTHESIS BOUNDARY IS FOR
THE NUMBER OF RUNS IN THE RANGE OF 11 - 20
```

SECTION NUMBER	RUN PATTERN	NUMBER OF RUNS	STATIONARITY
1	--++++---++++++--------+-+-+++	10	NON-STATIONARY
2	++-----++++++---++---+++---+-+	11	STATIONARY
3	----+-+++---------+++++++++++-	7	NON-STATIONARY
4	+++++----+++++++++-----------+	5	NON-STATIONARY
5	----+++++++-+--+++++++--------	7	NON-STATIONARY

```
TOTAL NUMBER OF SECTIONS =    5
DURATION OF EACH SECTION = 12 00 SECONDS
TOTAL RECORD TIME =   1.00 MINUTES
PERCENTAGE OF REJECTION =  80 00 %
```

Fig. 2.4 The RUN test output for the colonic signal illustrated in Fig. 2.3.

2.4 CHI-SQUARE GOODNESS-OF-FIT METHOD FOR THE TESTING OF NORMALITY

The normality of a set of data points $X(t)$ can be tested by using the chi-square goodness-of-fit test which measures any discrepancy between the distribution of $X(t)$ and the normal distribution. The general procedure involves the use of a statistic with an approximate chi-square distribution as a measure of the difference between the sample and the normal density function. Then a hypothesis of equivalence can be tested by studying the sampling distribution of this statistic.

Suppose there are N sample data points to be analyzed. Let the points be grouped among k class intervals such that the number of expected points (or frequencies) in each interval will be the same. Note the class width for each interval will be different in this case. The probability for a particular point to fall in any one of the intervals is then $1/k = 1 - \alpha$. For instance, for interval #1, $\alpha = 1 - 1/k$. For interval #2, $\alpha = 1 - 2(1/k)$ etc.. The appropriate interval limits for the normal distribution hypothesis are established by finding the required value of z for each interval limit from the table of 'area under standardized normal density function', say Table A2 of { 1 }. These standardized interval limits are then transformed into boundary values in terms of the observed sample. This is done by the equation

$$Y = z_{\alpha} S + \bar{X} \tag{2.6}$$

where $\bar{X}$ is the mean and S is the standard deviation of the sample. Now we can count the number of points that fall within each interval. Let F_i be the observed frequency of interval i. The frequency of each class interval is also selected to be identical and equal to N/k. The value $\Sigma(F_i - N/k)$ will measure the discrepancy between the observed sample and the normal density function. Let the sample statistics β to be established by summing the squares of the differences in each interval as follows :

$$\beta = \sum_{i=1}^{k} \frac{(F_i - N/k)^2}{(N/k)} \tag{2.7}$$

Let it be hypothesized that the sample is normally distributed. Any deviation from the hypothesized normal density function will cause β to increase. Hence a one-sided test is used and the region of acceptance is

$$\beta \leq \chi^2_{n:\alpha} \tag{2.8}$$

where $\chi^2_{n:\alpha}$ is the percentage point of the chi-square distribution and can be found from tables like Table A3 of { 1}. Here α is the level of significance for the hypothesis and N is the degree of freedom. N will depend on the number of different independent linear restrictions imposed on the observed sample. In this case the first restriction is that the frequency of the last class interval is determined once the frequencies of the first k - 1 classes are known. Furthermore, two parameters, the mean and the variance, must be calculated from the sample in order to fit the normal density function because they are not known from the expected theoretical function. Thus the degree of freedom n for the test is k - 3. One additional note that has to be mentioned here is that the optimal number of classes k can be computed as follows {25} :

$$k = 4 \left(\frac{2\,(N-1)^2}{z_\alpha^2} \right)^{1/5} \tag{2.9}$$

For example, if N = 150 and $\alpha = 0.05$, z_α can be found to be 1.645 from Table A2 of { 1}. So from (2.9), k = 27 and the degree of freedom n = k - 3 = 24. From table A3 of $\chi^2_{n:\alpha} = \chi^2_{24:0.05} = 36.42$. One can then use equations (2.7) and (2.8) to test whether the hypothesis is true or not.

The algorithm of this test implemented in the computer is shown in Fig. 2.5. As mentioned before, when establishing the interval limits, values of z_α have to be found from the α / z_α table, say Table A2 of { 1}. Due to the limited amount of memory space available, only a

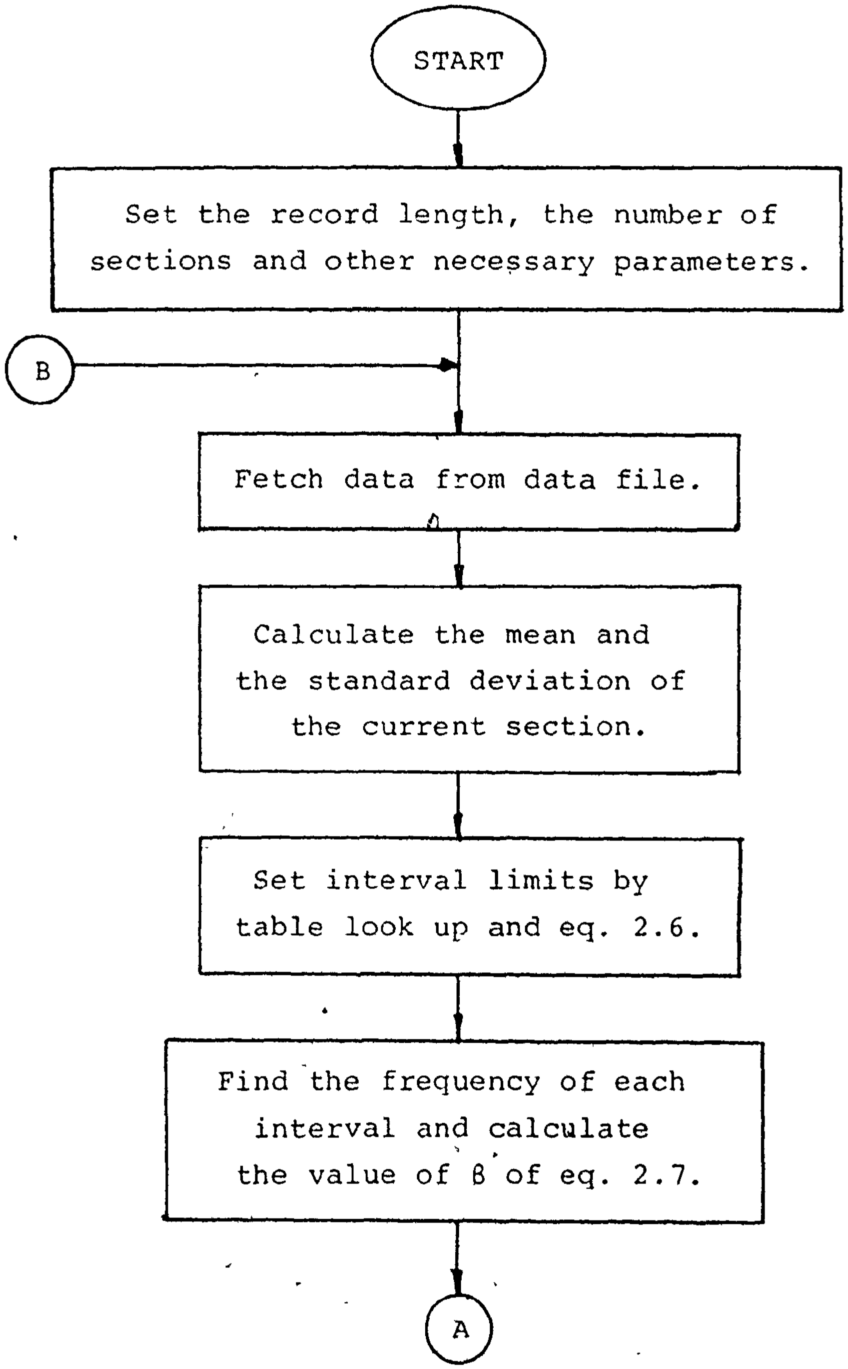


Fig. 2.5 The flow diagram of the chi-square goodness-of-fit normality test.

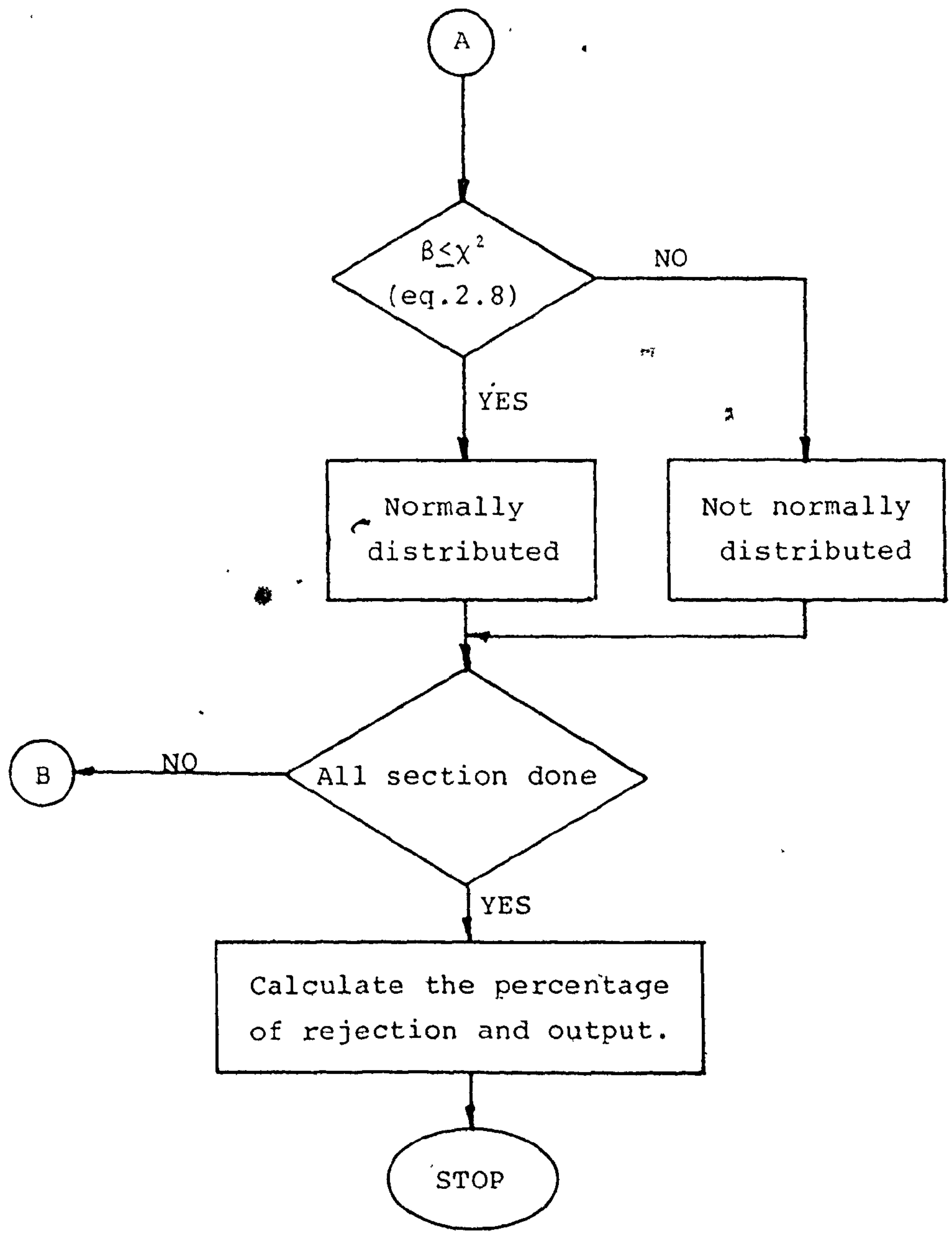


Fig. 2.5 (continued)

portion of the table is entered into the computer and the number of sample points N is restricted to either 60 or 150. This is because for each class interval, a distinct value of α (=1-i(1/k) and k is calculated from (2.9)) is found and hence, k different values of z_α have to be entered for each value of N. This test is a rather simple test and the algorithm in Fig. 2.5 is self explanatory. Fig. 2.6 illustrates a sample output of the test after the signal shown in Fig. 2.3 is processed. From there we can see that the signal under study is not normally distributed.

2.5 FISHER'S NORMALITY TEST

The normality test by using the Fisher's k statistics is an enhancement of the chi-square goodness-of-fit test just mentioned on section 2.4. Two parameters g_1 and g_2 are introduced from the k statistics. g_1 is the coefficient of skewness which measures whether the sample distribution is symmetric about the mean or not. A positive g_1 indicates that the low-valued numbers of the observed sample are close to the mean, and the high-valued numbers extend far above the mean. On the other hand, a negative g_1 will have an opposite effect. However if the sample is normally distributed, g_1 is also normally distributed with zero mean and variance :

$$\mathrm{var}(g_1) = \frac{6N(N-1)}{(N-2)(N+1)(N+3)} \tag{2.10}$$

```
CHANNEL NUMBER =  3
INITIAL RECORD NUMBER =   1
NUMBER OF DATA POINTS PER SECTION =  150
THE SIGNIGICANT LEVEL VALUE =   5 00 %
```

SECTION NUMBER	BETA	NORMALITY
1	35 920000	NOT NORMALLY DISTRIBUTED
2	12 860000	NORMALLY DISTRIBUTED
3	129 470000	NOT NORMALLY DISTRIBUTED
4	230 120000	NOT NORMALLY DISTRIBUTED

```
TOTAL NUMBER OF SECTIONS =   4
DURATION OF EACH SECTION = 15 00 SECONDS
TOTAL RECORD TIME =  1.00 MINUTES
THE % POINT OF THE CHI-SQUARE DIST  = 26.3
PERCENTAGE OF REJECTION =  75 00 %
```

Fig. 2.6 The output of the chi-square goodness-of-fit normality test for the colonic signal illustrated in Fig. 2.3.

where N is the number of sample points analyzed.

g_2, on the other hand, is the coefficient of kurtosis. It measures the peakedness of a distribution and shows how much the symmetric distribution (if any) may deviate from the normal distribution. A positive g_2 indicates the peak is sharp and the tails are thick, that is, the shape is more pronounced. Nevertheless, the region between the peak and the tails becomes lower than that in the normal distribution. On the other hand, a negative valued g_2 means the peak is thicker, the tails are thinner and the region between them is thicker. If the sample is normallv distributed, g_2 is also normally distributed with zero mean and variance :

$$\mathrm{var}(g_2) = \frac{24N(N-1)^2}{(N-3)(N-2)(N+3)(N+5)} \tag{2.11}$$

As a summary, g_1 and g_2 measure the degree of deviation of the sample distribution from the normal one. As a matter of fact, the Fisher's normality test can be used alone as a normality test of any time series.

Now we can show the calculations of g_1 and g_2 and demonstrate the hypothesis test involved. g_1 and g_2 are computed from the Fisher's k statistics as follows :

$$g_1 = k_3 \,/\, (k_2\sqrt{k_2}) \tag{2.12}$$

$$g_2 = k_4 / (k_2^2) \tag{2.13}$$

where

$$k_2 = \frac{1}{(N-1)} m_2$$

$$k_3 = \frac{N}{(N-1)(N-2)} m_3 \tag{2.14}$$

$$k_4 = \frac{N}{(N-1)(N-2)(N-3)} \{ (N+1) m_4 - \frac{3(N-1)}{N} m_2^2 \}$$

and $$m_\nu = \sum_i (x(i) - \bar{X})^\nu , \qquad \nu = 2,3,4 \tag{2.15}$$

On a level of significance α, the sample distribution is said to be normal if the following conditions are satisfied :

$$\hat{z}_{g_1} = | g_1 | / \sqrt{var(g_1)} \leq z_{\alpha/2} \tag{2.16}$$

$$z_{g_2} = | g_2 | / \sqrt{var(g_2)} \leq z_{\alpha/2} \tag{2.17}$$

where z_{g_1} and z_{g_2} are the normalized statistics and $z_{\alpha/2}$ is the standardized variable of the normal distribution at the $\alpha/2$ level. $z_{\alpha/2}$ can be found from the table of 'area under standardized normal density function',e.g. Table A2 of { 1 }.

The algorithm for the Fisher's normality test is shown in Fig. 2.7. Here sections of data are analyzed

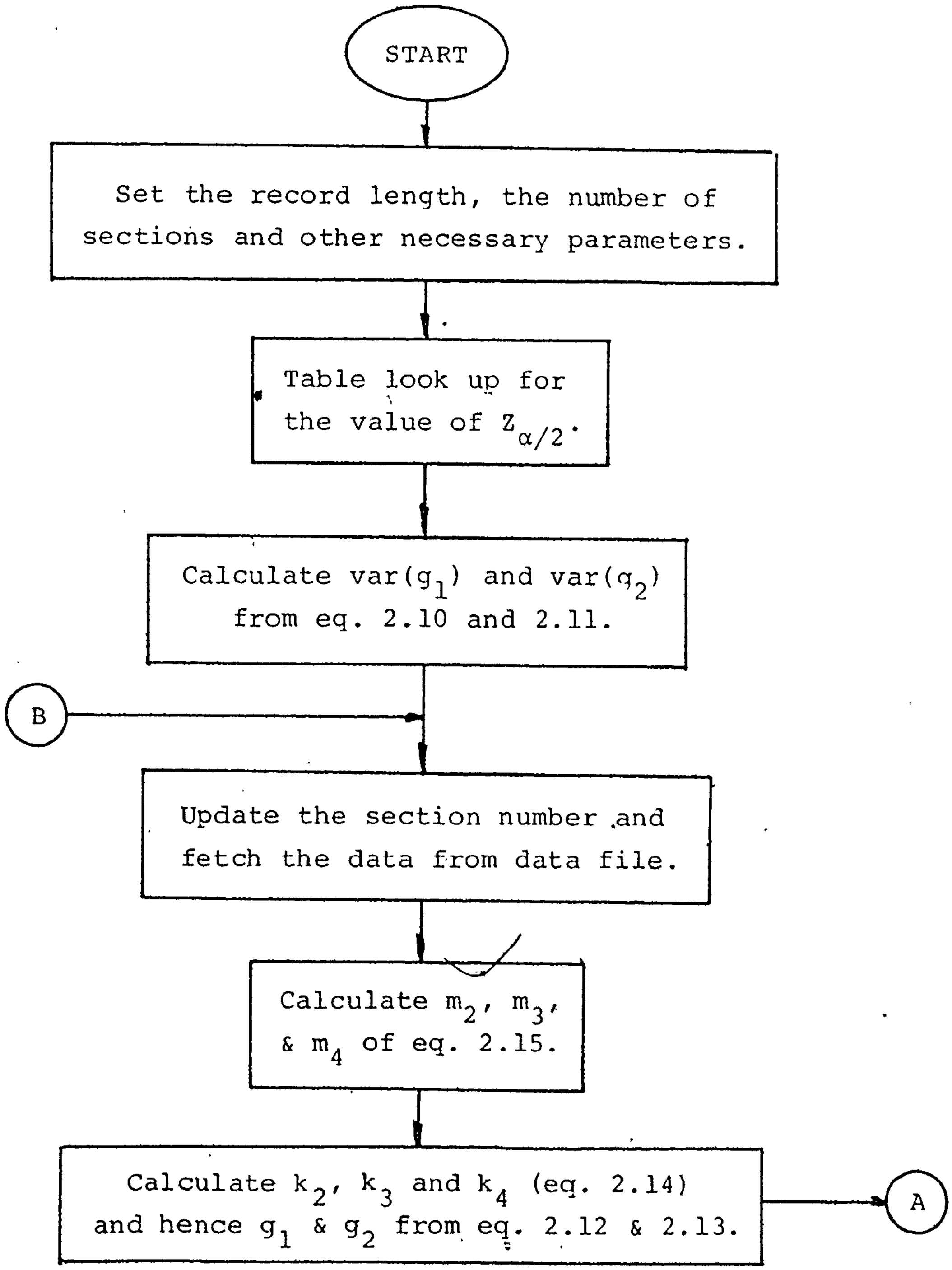


Fig. 2.7 Flow diagram of the Fisher's normality test.

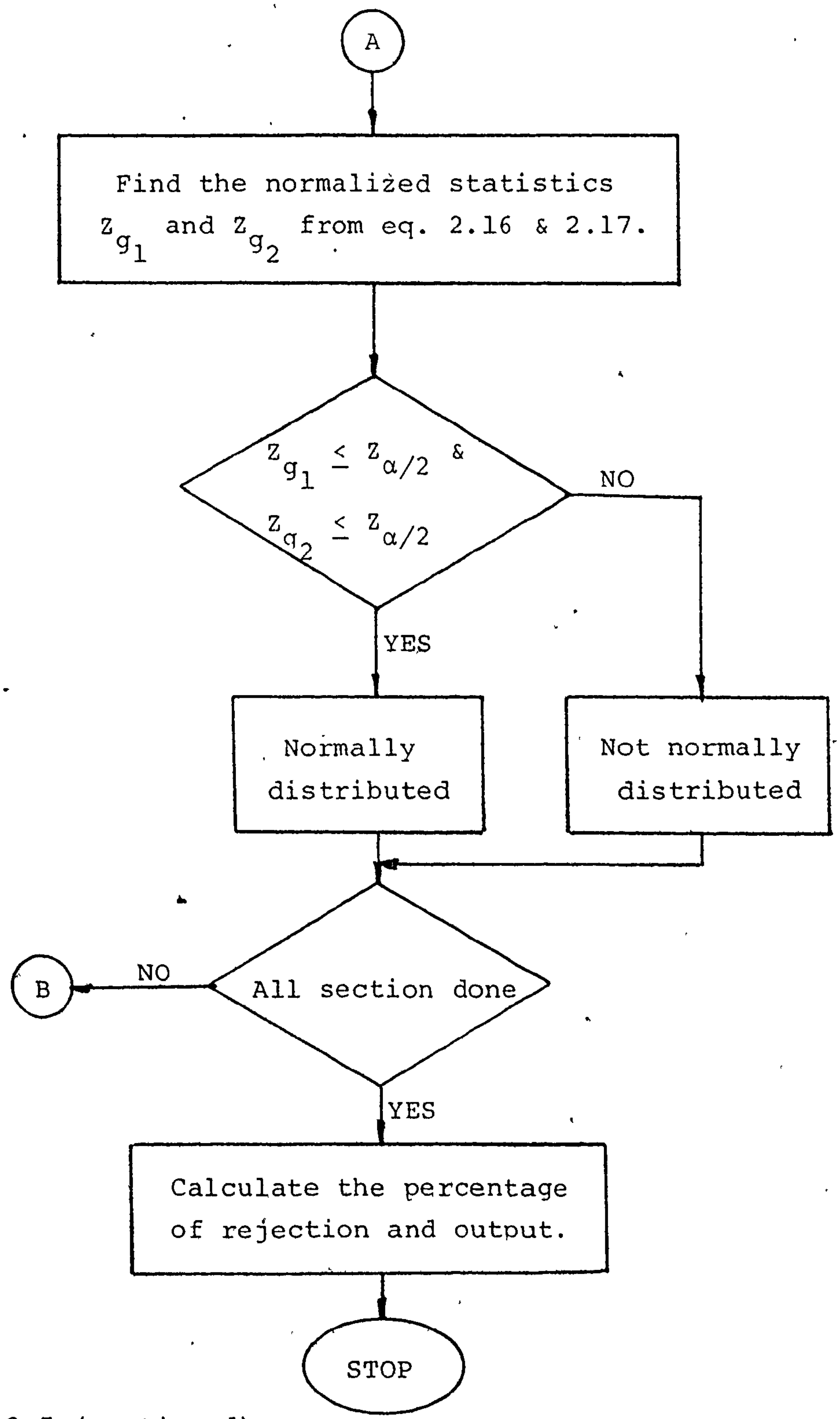


Fig. 2.7 (continued)

continuously as indicated in Fig. 2.3. The procedure is rather straight forward and the algorithm is self explanatory. Fig. 2.8 shows a sample output of the test when the signal of Fig. 2.3 is studied. The percentage of rejection here coincides with the result obtained from the chi-square goodness-of-fit test.

CHANNEL NUMBER = 3
INITIAL RECORD NUMBER = 1
NUMBER OF DATA POINTS PER SECTION = 150
THE SIGNIFICANT LEVEL VALUE = 5 00 %
THE STANDARIZED VARIABLE Z(ALPHA/2) = 1 960

SECTION NUMBER	ZG1	ZG2	NORMALITY
1	3.143	1 647	NOT NORMALLY DISTRIBUTED
2	0.264	0 532	NORMALLY DISTRIBUTED
3	3 221	1 152	NOT NORMALLY DISTRIBUTED
4	2.048	1 354	NOT NORMALLY DISTRIBUTED

TOTAL NUMBER OF SECTIONS = 4
DURATION OF EACH SECTION = 15 00 SECONDS
TOTAL RECORD TIME = 1.00 MINUTES
PERCENTAGE OF REJECTION = 75.00 %

Fig. 2.8 The output of the Fisher's normality test for the colonic signal illustrated in Fig. 2.3.

CHAPTER 3

THE BISPECTRAL ANALYSIS

3.1 INTRODUCTION

Conventional power spectrum of a stationary gaussian stochastic process displays the energy distribution of the process in the frequency domain. It essentially contains all the statistical information about the process under investigation. Since the process is gaussian and stationary, different frequency components in the spectrum are expected to be mutually uncorrelated, that is, they do not affect each other in the generating process. Hence the spectrum itself is merely a linear superposition of these frequency components. On the other hand, when harmonically related peaks exist in the spectrum, wave components that are not sine-shaped may be present. This means that some components in the Fourier series of the signal are somehow phase related to each other and they sum up to give the non-sinusoidal components. Let the Fourier transform of a random process x(t) be

$$X(f) = A(f)\exp\{jF(f)\} \tag{3.1}$$

where A(f) is a positive even function for the amplitude and F(f) is an odd function which represents the phase. If some terms in the Fourier series are related to each other as mentioned above, the corresponding terms in the function F(f) will then be phase-locked to each other in some degree. Furthermore, even if the frequency components are not integer multiples of each other, they could be phase-locked and this information will not be revealed by the power spectral analysis. Hence a quantitative measurement of this kind of phase-locking characteristics may be of interest.

By definition, the power spectrum of a random process x(t) is

$$\begin{aligned} P(f) &= \frac{1}{T} \{ X(f)X^{*}(f) \} \\ &= \frac{1}{T} \{ A(f)^{2} \} \end{aligned} \tag{3.2}$$

where X(f) is defined by equation (3.1). It is obvious that the power spectrum only shows the amplitude distribution of the process and no phase information of any kind can be observed. One may argue that the phase relations can be obtained if the peaks are presented at harmonically related frequencies. But are these frequency components really true harmonics of each other? Additionally, existence of any phase-locking characteristics between frequency bands outside the peak frequencies of the spectrum cannot be observed from the spectrum. In some cases, observed data may be generated by some non-linear mechanisms. Again the conventional power spectrum do not suggest any information on this aspect. All these problems thus suggest that higher order theory should be employed in order to gain some insights on the observed data if it does not satisfy the definition of gaussianity and stationarity.

The third and fourth moments are very common in statistics. They measure the skewness and the kurtosis, respectively, of the sample distribution (see details on these moments in section 2.5). But they are only estimates of the deviation of the observed sample from a gaussian distribution and reveal nothing about the non-linear interaction between different frequency components. Hence a more differentiated means, called the bispectral analysis, is employed in this study. A bispectrum is the Fourier transform, or the frequency decomposition, of the third order moment function of a stochastic process. Since the third moment is influenced either by the non-stationarity of the signal or the interrelations between frequency components, the bispectrum does provide a means to explore the phase relations and the degree of phase-locking among components in different frequency bands, It not only identifies the harmonically or sub-harmonically related frequency components, it also detects the coupling between two or more generators of the signal. In summarizing all these, the bispectral analysis investigates the phase-locking character and gives some new insights into the non-linear aspects of the generating process of the signal. We will first discuss the theory and the derivation of the bispectrum. Then the estimation, that is, the method, procedures and the algorithm, of the bispectrum and the bicoherence (which is just the normalized bispectrum) will be presented. Finally in this chapter the algorithm will be tested by artificial data.

3.2 THEORY AND DERIVATION OF THE BISPECTRUM

In this project the theory of the bispectrum will be derived through the analogous expression of the conventional spectral density function. One can refer to { 4 } if the formal derivation is of interest. In this reference a general k-order polyspectrum is introduced but the theory behind this is quite involved. Hence the detailed formal procedure in deriving the bispectral expression is omitted here. By definition, an ordinary power spectrum distributes the energy (or the mean square $\bar{x}^2$) of a time series $x(t)$ in the frequency domain. It is usually obtained by the Fourier transform of the first order autocovariance $E\{x(t)x(t+\tau)\}$ of $x(t)$ or more efficiently, the product of two Fourier components of $x(t)$ with the frequencies of the components adding to zero. By contrast, the bispectrum is a decomposition of the mean cube $\bar{x}^3$ of $x(t)$ into its frequency counterpart. It can be calculated from the two dimensional Fourier transform of the second order autocovariance $E\{x(t)x(t+\tau_1)x(t+\tau_2)\}$ (which is a function of two delays) of the time series $x(t)$. Analogous to the case of power spectrum, the bispectrum can also be obtained from the product of three Fourier components of $x(t)$ with the resultant frequency sum equal to zero. But usually the relations between different time series are of interest in time series analysis as well, e.g., through the cross-spectrum and the coherence function, there is no reason to confine the bispectral expression into a single time series. Let $x_1(t)$, $x_2(t)$ and $x_3(t)$ be three independent realizations of a random variable X for the time being. The bispectrum can then be defined as the Fourier transform of the second order cross-covariance function $E\{x_1(t)x_2(t+\tau_1)x_3(t+\tau_2)\}$. Note the order of arrangement of x_1, x_2 and x_3 in the expected value expression is completely general and irrevelent. More details about this point will be discussed later in the cross-bispectrum section.

Consider the estimate of the cross-covariance $E\{x_1(t+\tau_1)x_2(t+\tau_2)x_3(t+\tau_3)\}$ of x_1, x_2 and x_3 be

$$R_{123}(\tau_1,\tau_2,\tau_3,)$$

$$= 1/N \sum_{t=1-\min(\tau_1,\tau_2,\tau_3)}^{N-\max(\tau_1,\tau_2,\tau_3)} x_1(t+\tau_1)x_2(t+\tau_2)x_3(t+\tau_3) \quad (3.3)$$

where $\tau_1 = 0, \pm 1, \pm 2, \ldots, \pm m$

$\tau_2 = 0, \pm 1, \pm 2, \ldots, \pm m$

$\tau_3 = 0$

The estimate is centered at $\tau_3 = 0$ since the product moment does not change with time. With $\tau_3 = 0$,

$$R_{123}(\tau_1,\tau_2,\tau_3) = R_{123}(\tau_1,\tau_2) \tag{3.4}$$

Taking the 2-dimensional Fourier transform of equation (3.3), we obtain the expression for the bispectrum as

$$B_{123}(\omega_1,\omega_2)$$

$$= 1/(2\pi)^2 \sum_{\tau_1=-m}^{m} \sum_{\tau_2=-m}^{m} \exp\{-j(\tau_1\omega_1+\tau_2\omega_2)\} R_{123}(\tau_1,\tau_2) \tag{3.5}$$

If we seperate the double summation and let

$$G_{123}(\omega_2;\tau_1)$$

$$= 1/(2\pi) \sum_{\tau_2=-m}^{m} \exp(-j\tau_2\omega_2) R_{123}(\tau_1,\tau_2) \tag{3.6}$$

then

$$B_{123}(\omega_1,\omega_2)$$

$$= 1/(2\pi) \sum_{\tau_1=-m}^{m} \exp(-j\tau_1\omega_1) G_{123}(\omega_2;\tau_1) \tag{3.7}$$

On the other hand, the bispectrum can also be obtained from the smoothed product of three Fourier components with the sum of the frequencies equal to zero as mentioned before. Let $X(\omega)$ be the Fourier transform of the time series $x(t)$. The bispectrum is then

$$B_{123}(\omega_1,\omega_2,\omega_3)$$

$$= 1/P \sum_{j=1}^{P} X_{1j}(\omega_1)X_{2j}(\omega_2)X_{3j}(\omega_3) \tag{3.8}$$

with $\omega_1 + \omega_2 + \omega_3 = 0$.

In this expression, the bispectrum is obtained by averaging P sections of the signal. With $\omega_3 = -(\omega_1 + \omega_2)$, equation (3.8) becomes

$$B_{123}(\omega_1,\omega_2)$$

$$= 1/P \sum_{j=1}^{P} X_{1j}(\omega_1)X_{2j}(\omega_2)X^{*}_{3j}(\omega_1+\omega_2) \tag{3.9}$$

where the * is designated as the conjugate of a complex number. If $x_1(t)$, $x_2(t)$ and $x_3(t)$ are all stationary gaussian processes, any moment higher than the second order will be zero and as a result, the bispectrum vanishes theoretically. However, since the output of a Fourier transform is in a complex numbered form, the bispectral expression in equation (3.9) will also yield a complex value.

Sometimes the complex numbers may cancel each other during the addition operation, especially when the amplitudes are symmetrically distributed. A zero value of the mean cube does not necessarily imply that the bispectrum vanishes. Hence, a normalized bispectrum, or the bicoherence, may be more desirable. The bicoherence is defined as

$$C_{123}(\omega_1,\omega_2) = \left\{ \frac{P\;|B_{123}(\omega_1,\omega_2)|^2}{\sum_{j=1}^{P}|X_1(\omega_1)|^2 \sum_{j=1}^{P}|X_2(\omega_2)|^2 \sum_{j=1}^{P}|X_3(\omega_1+\omega_2)|^2} \right\} \tag{3.10}$$

Notice that the expression for $C_{123}(\omega_1,\omega_2)$ has the following relation :

$$C_{123}(\omega_1,\omega_2) \propto \frac{X_1(\omega_1)}{|X_1(\omega_1)|} \cdot \frac{X_2(\omega_2)}{|X_2(\omega_2)|} \cdot \frac{X_3(\omega_1+\omega_2)}{|X_3(\omega_1+\omega_2)|} \tag{3.11}$$

If $X(\omega) = |X(\omega)|\exp\{j(\omega+\phi)\}$,

$$C_{123}(\omega_1,\omega_2) \propto \exp\{j(\phi_1+\phi_2-\phi_3)\} \tag{3.12}$$

Through equation (3.12) the bicoherence investigates the phase-relationship of the frequency triple ω_1, ω_2 and $\omega_1+\omega_2$.

Usually the bicoherence is represented by a contour map with the abscissa being the axis for ω_2 and the ordinate as the axis for ω_1. The value proportional to exp $j(\phi_1+\phi_2-\phi_3)$ is then plotted on the map. For a random process the phase value of each individual frequency component should be independent of each other. Hence no peak in the contour map should appear. On the other hand, for a non-stationary process, sometimes it is possible for a peak to appear at a point (ω_1,ω_2) when a particular section of signal is being analyzed. But this may not be true for some other sections. So the smoothing (or averaging) procedure should be included in the computational process in order to find out the phase-locked components in the signal.

Usually after the smoothing procedure, a peak will show up at a particular point when $\phi_1 + \phi_2 - \phi_3$ is constant for all the sections of the signal being partitioned during the procedure. However for a non-stationary signal (the kind that is of interest in this project), the probability that $\phi_1 + \phi_2 - \phi_3$ being constant everywhere should be very small. The only possible case for an occurance of a peak would be when $\phi_1 = \phi_2 = \phi_3$, i.e., they are all phase-locked. Hence the main purpose of the bispectral analysis is to investigate the degree of phase-locking between individual frequency components. Many applications can be obtained through the bispectral analysis. For example, it can easily identify the harmonics of a fundamental in a time series since all the harmonics are supposed to be phase-locked. However, if an independent time series is composed of some frequency components which happen to be the harmonics of another time series and if the two series are placed side by side in order to form a single series, the result will produce a power spectrum which may be misleading. It will show a fundamental and several of its harmonics and one may think that they sum up to form a particular non-sinusoidal wave component in the signal. But since the two original series are independent of each other, their phase values will exhibit no relations whatsoever and hence the integer multiple related components will not add up together. Therefore the bispectral analysis can easily identify the two independent oscillators which together generate the signal. More details about this example will be presented later in the chapter.

At this point perhaps it is worthwhile to present another example which will prove the validity of the equation (3.12). Let x(t) be a periodic signal composed of the fundamental ω_0 and its two harmonics $2\omega_0$ and $3\omega_0$, i.e.

$$x(t) = \cos(\omega_0 t+\phi_1) + \cos(2\omega_0+\phi_2) + \cos(3\omega_0+\phi_3) \quad (3.13)$$

The Fourier transform of x(t) will then be

$$\begin{aligned} X(\omega) = 1/2\{ & \exp(j\phi_1(\delta(\omega-\omega_0) + \delta(\omega+\omega_0))) \\ & + \exp(j\phi_2(\delta(\omega-2\omega_0) + \delta(\omega+2\omega_0))) \\ & + \exp(\cancel{j\phi_3(\delta(\omega-3\omega_0)} + \delta(\omega+3\omega_0))) \} \end{aligned} \quad (3.14)$$

$$\text{With} \quad \int_{-\infty}^{\infty} \delta(t - \tau)\, d\tau = 1, \qquad t = \tau$$
$$= 0, \qquad t \neq \tau \tag{3.15}$$

the bicoherence of x(t) at $(\omega, 2\omega_0, 3\omega_0)$

$$\begin{aligned} & C_x(\omega_0, 2\omega_0) \\ & \propto 1/k \; \{ \; X(\omega_0)X(2\omega_0)X^*(3\omega_0) \; \} \\ & = 1/k \; \{ \; 1/2 \;\; \exp\,(j(\phi_1+\phi_2-\phi_3)) \; \} \\ & \equiv \text{equation (3.12)} \end{aligned}$$

So if x(t) is a random process, a peak will only appear in the bicoherence map when $\phi_1 = \phi_2 = \phi_3$, i.e., $X(\omega_0)$, $X(2\omega_0)$ and $X(3\omega_0)$ are phase-locked.

Let λ_0 be the highest frequency value present in all of the three time series $x_1(t)$, $x_2(t)$ and $x_3(t)$ of a multivariate process. The bispectral and the bicoherence functions of these series will be defined on a square two-dimensional area in the frequency plane centered on $\omega_1 = \omega_2 = \omega_3 = 0$ with sides equal to λ_0. However if any two or all of the three series are identical, the symmetric property will be effected and the amount of computations of the estimates can be reduced. If any two of the series are identical, one needs only to calculate the bispectral values on a 90^0 triangular area on the frequency plane with sides equal to λ_0. In the case that all the three series are the same, the bispectrum needs only be computed for the

the octant $0 \leq \omega_1 \leq \omega_2$, $\omega_1 + \omega_2 \leq \lambda_0$.

Another interesting parameter worth discussing briefly here is as follows. Since the bispectrum is a complex value, it can be expressed by its modulus and its phase angle. The modulus can be expressed by the bicoherence function just mentioned. The phase value, named biphase, is defined as

$$\Phi(\omega_1,\omega_2) = \tan^{-1} \frac{\mathrm{img}\{B(\omega_1,\omega_2)\}}{\mathrm{real}\{B(\omega_1,\omega_2)\}} \tag{3.16}$$

It is a measure of the mutual phase-shift between different frequency components. If the harmonically related frequency components are under consideration, the biphase will exhibit the mutual time shifting relations between the fundamental and the higher harmonics. But in this project the biphase values of any time series will not be computed due to the lack of practical usefulness at the present state of the research. But this could be a good topic for further exploration of any largely unknown non-stationary signal.

Now let us consider the case where two of the three series mentioned above are identical. From equation (3.5), the cross-bispectrum of $x_1(t)$ and $x_2(t)$ will be

$$B_{12}(\omega_1,\omega_2)$$

$$= 1/(2\pi)^2 \sum_{\tau_1=-m}^{m} \sum_{\tau_2=-m}^{m} \exp\{-j(\tau_1\omega_1+\tau_2\omega_2)\} R_{12}(\tau_1,\tau_2) \tag{3.17}$$

with $$R_{12}(\tau_1,\tau_2) = E\{x_1(t)x_2(t+\tau_1)x_2(t+\tau_2)\} \tag{3.18}$$

On the other hand, if the triple product method of different frequency components of the two channels of signals is used, a similar form of expression as that of equation (3.9) will be obtained. For the moment, perhaps it is better to omit the averaging part of the equation. Since there is no restriction in the arrangement of the terms, the following combinations will be valid :

$$B_{12}(\omega_1,\omega_2) = X_1(\omega_1)X_1(\omega_2)X_2^*(\omega_1+\omega_2) \tag{3.19}$$

$$B_{12}(\omega_1,\omega_2) = X_2(\omega_1)X_2(\omega_2)X_1^*(\omega_1+\omega_2) \tag{3.20}$$

$$B_{12}(\omega_1,\omega_2) = X_1(\omega_1)X_2(\omega_2)X_1^*(\omega_1+\omega_2) \tag{3.21}$$

$$B_{12}(\omega_1,\omega_2) = X_2(\omega_1)X_1(\omega_2)X_1^*(\omega_1+\omega_2) \tag{3.22}$$

$$B_{12}(\omega_1,\omega_2) = X_1(\omega_1)X_2(\omega_2)X_2^*(\omega_1+\omega_2) \tag{3.23}$$

$$B_{12}(\omega_1,\omega_2) = X_2(\omega_1)X_1(\omega_2)X_2^*(\omega_1+\omega_2) \tag{3.24}$$

Due to the symmetry condition, equations (3.21) and (3.22) are equivalent and so is the case for equations (3.23) and (2.24). Thus four distinct combinations are involved, i.e. equations (3.19), (3.20), (3.21) and (3.23). So far no one has done any interpretation on the cross-bispectrum or used it in any application. After some considerations, the author has arrived at the following meanings about the cross-bispectral value in equations (3.19), (3.20), (3.21) and (3.23) : in each equation a peak in the cross-bicoherence contour map indicates that the three frequency components are phase-locked to each other. If it happens that ω_2 and $\omega_1+\omega_2$ are integer multiple of ω_1, one would suspect that these components may be generated by the same source and it influences the behaviour of the 2 signals under investigation. It does not matter which channel an individual frequency component in the expression belongs to. For instance, consider equation (3.21). If a high bicoherence is observed at ω_1 and ω_2 with ω_2 being an integer multiple of ω_1, the frequency components at ω_1, $\omega_1+\omega_2$ in channel 1 as well as at ω_2 in channel 2 will come from the same generator. Hence the cross-bispectral values will give one a more in-depth understanding about the relationship between 2 different channels of signal. It can also be treated as an enhancement of the conventional cross-spectral analysis. Further interpretation about the cross-bispectrum will be presented when practical signal is tested.

There is one final remark before we move on to the next section. A high bicoherence usually indicates the existence of a non-linear or non-stationary effect. Theoretically the bicoherence can have an infinite value (as opposed to the conventional coherence function) if complete phase-locking occurs. The value will usually depend on the sample size and several other computational details. But this should not be a problem if many estimates, computed by the same method, are to be compared simultaneously. In this project, the bicoherence and the cross-bicoherence methods will be used to examine the gastrointestinal signal which fails the stationarity criteria. In the next section the method, procedures and algorithms about the computation of the required parameters will be discussed in more detail.

3.3 METHODS AND PROCEDURES OF THE ESTIMATION

Two sep rate programs have to be established in this project for the bispectral analysis. The first is designed for the processing of a single channel of signal or time series. The second is for two independent series so that the cross-bispectrum can be found. However, due to the similarity between the natures of the two programs, the method and procedures described below will be based on the processing of a single time series. Additional requirements for the computation of the cross-bispectrum will then be described later.

Three methods have been suggested so far {11} in calculating the bispectrum of a time series : (1) complex demodulation, that is, narrow bandpass filtering the output components of the Fast Fourier Transform of the time series, transform them back to the time domain and do the averaging there {7,11}; (2) averaging in the frequency domain, that is, averaging the neighbouring components of the FFT output; and (3) averaging over successive records. One of these methods has to be carried out during the smoothing procedure of the computation. One would choose method (1) if the complex demodulates have to be computed anyway. Method (2) will be quite handy if only some subsets of the entire frequency range are of interest, e.g. on the diagonal where $\omega_1 = \omega_2$. But for a laboratory computer with a limited amount of memory, say the NOVA 830 where this project was carried out, method (3), sectioning the records and averaging over the individual biperiodogram, becomes the obvious choice. This smoothing method will be used in all the programs involved in this project (including those that calculate the power spectrum).

Due to the finite duration of the record, leakage will be found through the spreading of the main lobe and the addition of an infinite number of smaller side lobes of of the power spectral density function (for detailed explanation, see { 1 }p.315-317). Hence at the beginning the record is tapered by a full cosine bell in order to minimize the leakage. However this tapering will increase the variance, so the 50% overlapped sectioning technique is used in order to compensate this resulting deficit. The following paragraphs will show the computing procedure in greater detail.

First of all the record x(t) will be divided into P 50% overlapped sections with m discrete points in each section, that is,

$$x_j(t+m/2) = x_{j+1}(t), \qquad t = 1,2,\ldots,m/2 \quad (3.25)$$
$$j = 1,2,\ldots,P-1$$

For each section the mean is subtracted from each data point. Then the whole section will be tapered by a full cosine bell :

$$TP(i) = 0.5 * \{1 - \cos(2\pi(i-1)/m\} \qquad (3.26)$$

Zeros will then be added to the end of the section in order to make a total of NF data points which are convenient for a Fast Fourier Transform. (NF is to be specified by the user, e.g. 1024). Performing the FFT on the tapered data yields

$$Y_{jq} = 1/NF \sum_{t=0}^{NF-1} x_j(t)\exp\{-j2\pi(qt/NF)\},$$
$$0 \leq q \leq NF/2 \qquad (3.27)$$

Here q is related to the actual frequency value ω of any particular component by

$$\omega = q(\lambda_0/NF), \quad \lambda_0 = \text{sampling frequency} \qquad (3.28)$$

After the FFT components are obtained, the raw estimate of the power spectrum as well as the bispectrum are computed for that section through

$$\hat{f}_j(\omega) = |Y_{jq}|^2, \qquad 0 \leq \omega \leq \lambda_N \tag{3.29}$$

and $$\hat{B}_j(\omega_1,\omega_2) = Y_{jq_1} Y_{jq_2} Y^*_{jq_3}$$

where $$q_1=\omega_1(NF/\lambda_0), \quad q_2=\omega_2(NF/\lambda_0), \quad q_3=(\omega_1+\omega_2)(NF/\lambda_0)$$

$$0 \leq \omega_1 \leq \omega_2, \quad \omega_1 + \omega_2 \leq \lambda_N \tag{3.30}$$

Here λ_N is a frequency value which is big enough to reveal the information on the interested frequency range but $\leq \lambda_0/2$, the sampling frequency. Finally after all the sections have been processed through the above procedures, the individual periodogram and biperiodogram will be averaged by

$$\hat{f}(\omega) = 1/P \sum_{j=1}^{P} \hat{f}_j(\omega) \tag{3.31}$$

$$\hat{B}(\omega_1,\omega_2) = 1/P \sum_{j=1}^{P} \hat{B}_j(\omega_1,\omega_2) \tag{3.32}$$

where the bicoherence can then be found :

$$C(\omega_1,\omega_2)$$

$$= \left\{ \frac{|\hat{B}(\omega_1,\omega_2)|^2}{\hat{f}(\omega_1)\ \hat{f}(\omega_2)\ \hat{f}(\omega_1+\omega_2)} \right\}^{1/2} \tag{3.33}$$

All these procedures can be summarized by the flow charts shown in Fig. 3.1. Special attention has to be paid to the part where the bispectral components of each section are calculated. Because of the special symmetric characteristic, the bispectrum is only to be calculated on the triangular plane with the base length $q_N = (\lambda_N * NF)/\lambda_0$ (from equation (3.30)) and the height of $q_N/2$. Fig. 3.2 shows the flow chart of the section where the biperiodogram will be calculated. Note that a total of $(q_N * q_N/2)/2$ bispectral elements are calculated here. Hence a storage array of that size is needed.

After the procedures in computing the bispectrum of a single series have been described, we should take a scan at the additional requirements for the calculation of the cross-bispectrum of two independent time series. The program used is based on the four differently combined equations (3.19), (3.20), (3.21) and (3.23). It has essentially the same procedures as described before except that the program here has to handle the two series simultaneously. It will divide both series into the required number of sections, subtract the means on both series, tapering, finding the power spectral and bispectral components on both of them etc.. Additionally the processor will loop four times in order to calculate for and output the results of the differently combined equations. Here the symmetric characteristics about the range of the cross-

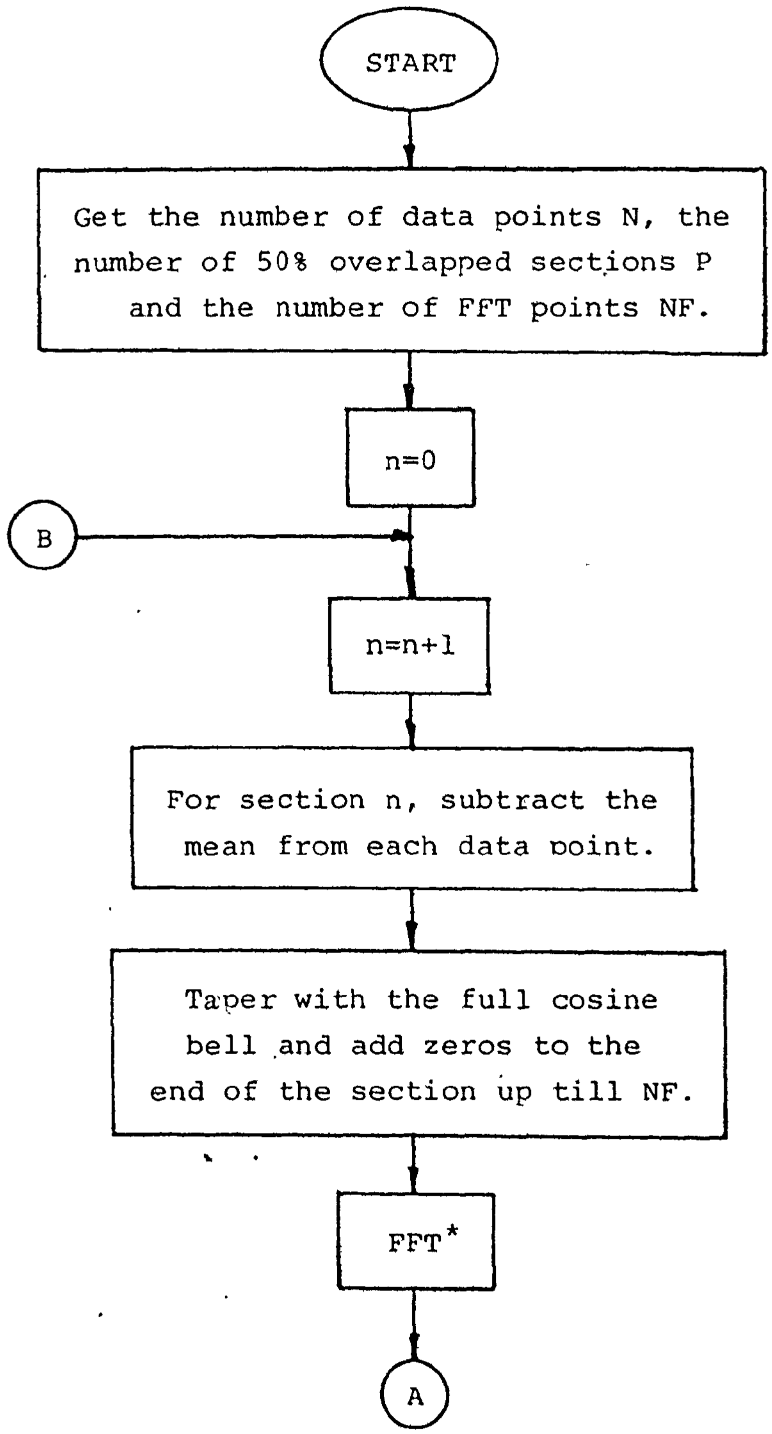


Fig. 3.1 Flow diagram for calculating a bispectrum.

* See {12} for the FFT algorithm.

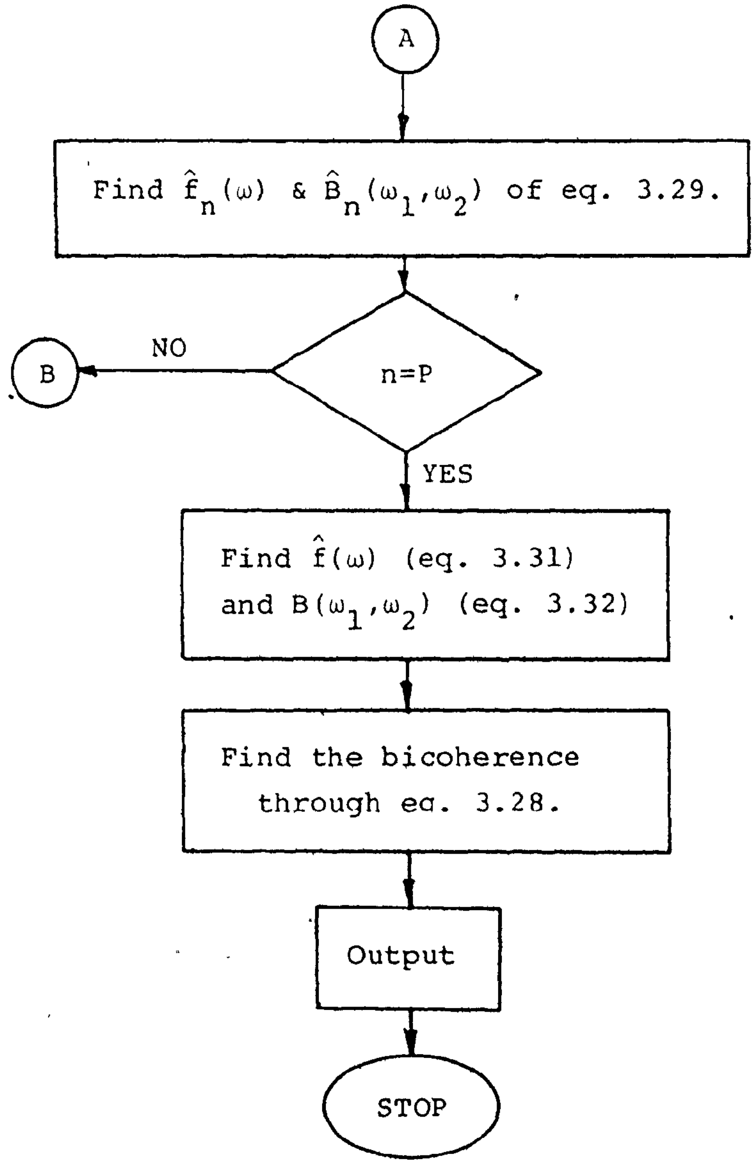


Fig. 3.1 (continued)

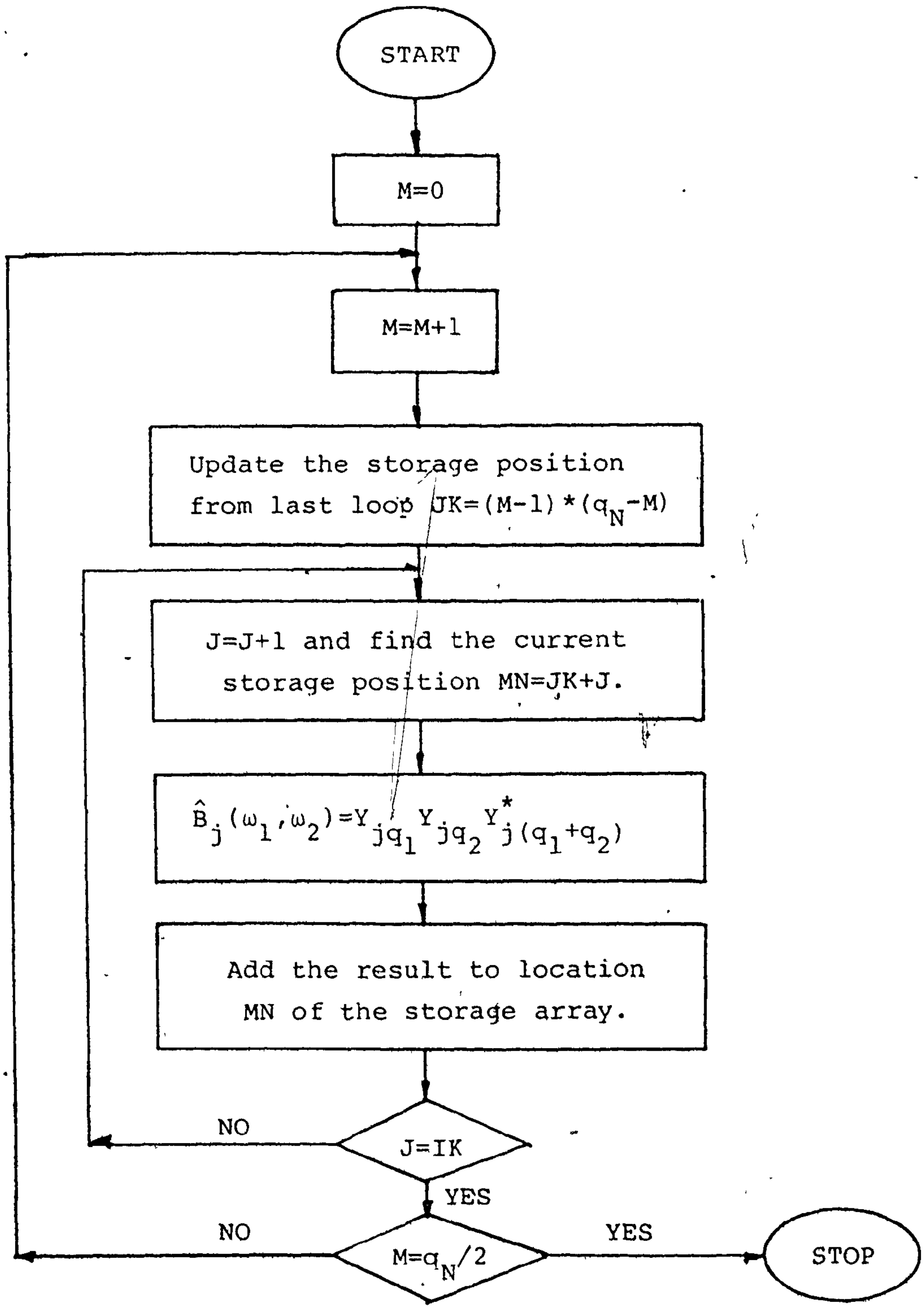


Fig. 3.2 Flow diagram for calculating the biperiodogram of a section of data.

bispectrum are different. For equations (3.19) and (3.20), where the cross-bispectrum is calculated according to $X_1X_1X_2$ and $X_2X_2X_1$ respectively, the range will be the same as that of the bispectrum of a single series, i.e. a triangle with a base length of q_N and a height of $q_N/2$. However, for equations (3.21) and (3.23), where the cross-bispectrum is calculated according to $X_1X_2X_1$ and $X_1X_2X_2$ respectively, the cross-bispectral plane will be a right-angled triangle with both the base length and the height equal to q_N. Hence the calculations and the outputs will be carried out according to these criteria. Fig. 3.3 shows the particular characteristics of this program.

There is one final note on the procedures. The output presented here are actually plots of contour maps of the magnitude of bicoherence and cross-bicoherence values. The magnitudes are divided into ten levels with the top level equal to the maximum bicoherence value in this particular map. Only the upper five levels (presented by symbols '#', '4', '3', '2' and '1' with '#' being the highest contour etc.) are shown for the sake of clarity. Before any real experimental data is tested by these programs, artificial data are used to test the validity of these algorithms and this is being described in the next section.

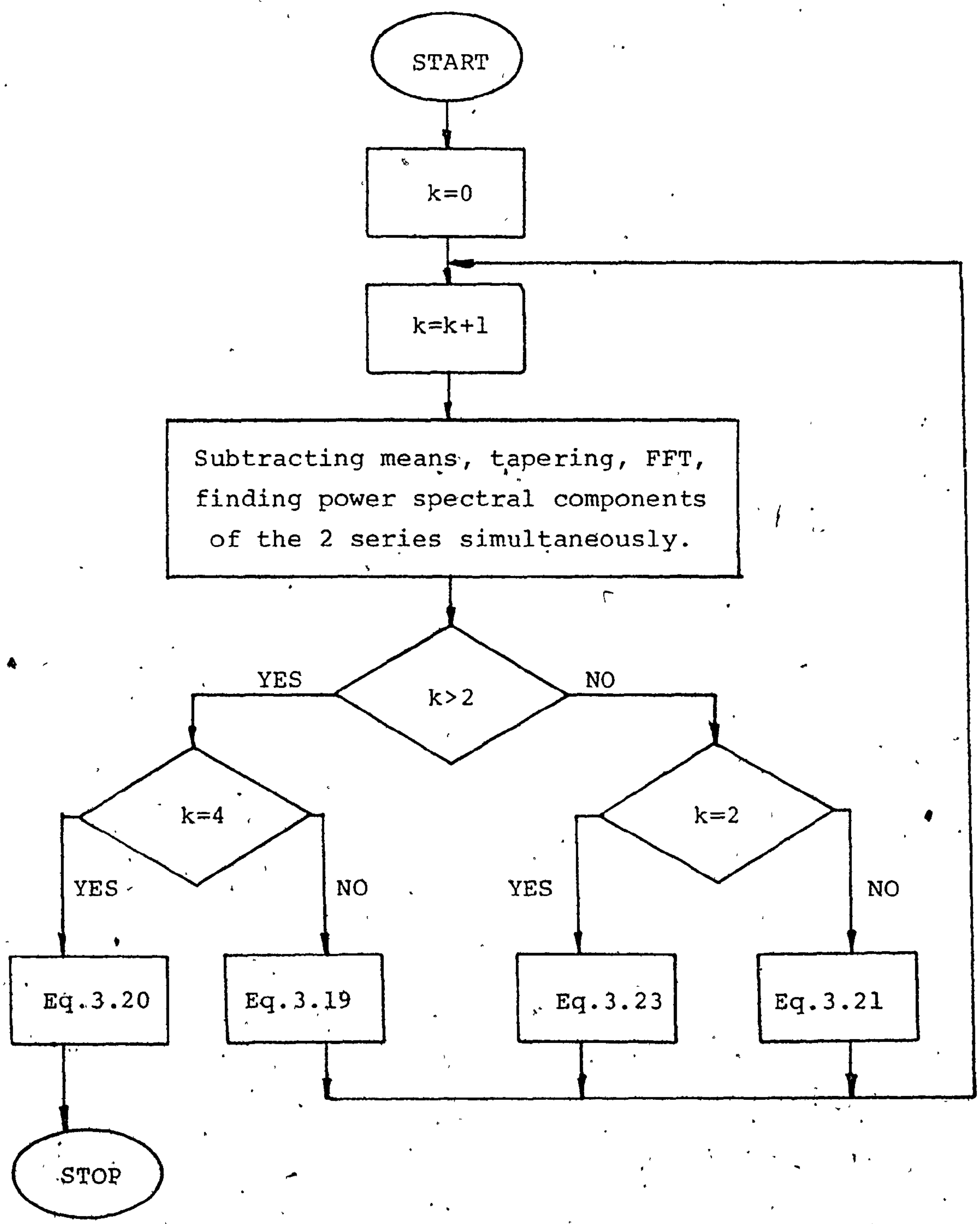


Fig. 3.3 Flow diagram for calculating the cross-biperiodogram.

3.4 BISPECTRUM OF ARTIFICIAL DATA

The generation process of the artificial data presented below is taken from reference {11}. The author feels that there is no need to use another kind of data because the one presented in {11} is very general and it is an excellent process to test the special characteristic of a bispectrum.

Here two stochastic processes are generated. They will have the same power spectrum but entirely different bispectral outputs as will be shown later. This analysis will show that sometimes people may make wrong interpretations by examining the power spectrum, which suppresses the phase information, alone. On the other hand, the bispectrum serves as an enhancement for the interpretation of the results and provides more information that a researcher may need.

The first process x(t) is taken from a periodic function f_1 which is composed of six harmonic components - the fundamental and its five subsequent harmonics. For each harmonic, say j, a time series is being generated by the following equation :

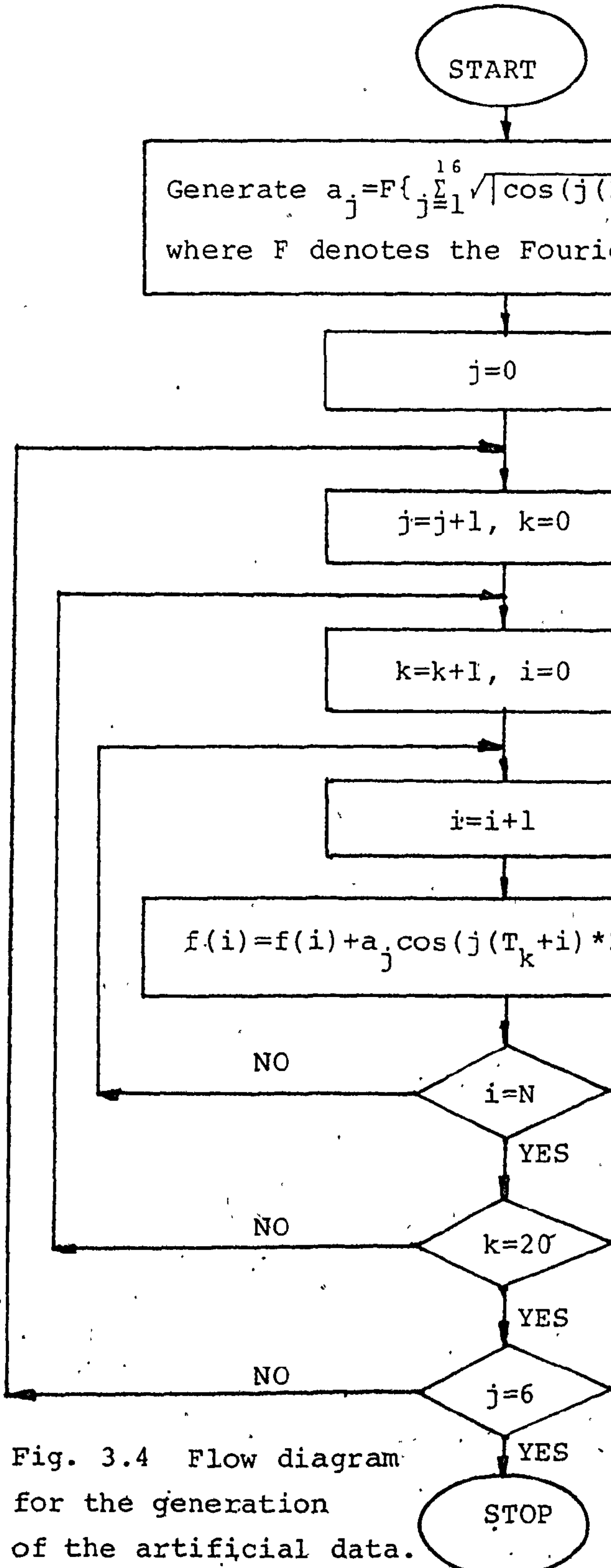
$$f_{1j} = \sum_{t=1}^{N} \sum_{k=1}^{20} a_j \cos\{j\omega(t-T_k)\} \qquad (3.34)$$

where a_j is the Fourier coefficient of $\sqrt{|\cos(\omega)|}$ at 5.4Hz. Hence $\omega = 2\pi(5.4)/f_s$ with f_s being an arbitrary sampling frequency and is taken to be 75 Hz here. Furthermore T_k is an independent random time value and 20 of them are being used. So, for each harmonic, the time series is produced by the summation of twenty distinct series with different time delay values. After all the six series for the six harmonics have been generated, the corresponding component of each series is added together. Additionally, a cosine bell is introduced to every 200 points of the final series for the tapering purpose. The overall procedure in generating the stochastic process with six harmonics is shown in Fig. 3.4.

The second process y(t) can be decomposed into two parts. The first part f_2 is exactly the same as that of x(t) in the first process just mentioned, except that here only the fundamental, 2^{nd}, 4^{th} and 5^{th} harmonics are included. The 20 time delay values are also the same. But for the second part f_3, only the 3^{rd} and the 6^{th} harmonics of the fundamental of x(t) are included. Additionally, the 20 random time delays are different from those in the first part. Hence y(t) can be formulated as

$$y(t) = \sum_{t=1}^{N} \sum_{j=1,2,4,5} \sum_{k=1}^{20} a_j \cos\{j\omega(t-T_k)\}$$

$$+ \sum_{t=1}^{N} \sum_{j=3,6} \sum_{k=21}^{40} a_j \cos\{j\omega(t-T_k)\} \qquad (3.35)$$

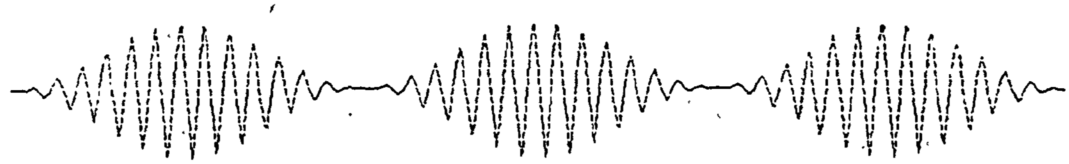


Here j is the harmonic number, T_k is the random delay number and i is the time unit.

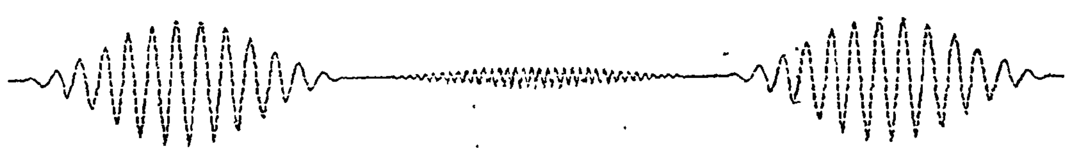
Fig. 3.4 Flow diagram for the generation of the artificial data.

Notice that equation (3.35) represents a series of 2N elements, N for the first part and another N for the second part of the equation, i.e. the two parts are placed side by side to each other in order to form a single series. The value of each element is computed by the second and the third summations of each part of the equation. Hence in each part, the first sigma merely represents a time series and the next two sigmas represent the summation operation. Fig. 3.5 shows the linearly combined time series x(t) and y(t).

Theoretically, x(t) and y(t) will have the same power spectrum and this is proved by Fig. 3.6. On the other hand the bicoherences of x(t) and y(t) are completely different as shown in Fig. 3.7 and 3.8. In Fig. 3.7, which shows the bicoherence of x(t), peaks are observed at (ω,ω), $(\omega,2\omega)$, $(\omega,3\omega)$, $(\omega,4\omega)$, $(\omega,5\omega)$, $(2\omega,2\omega)$, $(2\omega,3\omega)$, $(2\omega,4\omega)$ and $(3\omega,3\omega)$ where ω = 5.4 Hz. They are practically caused by the phase relationships between the corresponding frequency components in f_1. However, in Fig. 3.8, which shows the bicoherence of y(t), peaks can only be seen at (ω,ω), $(\omega,4\omega)$, $(2\omega,2\omega)$, which are caused by the inter-relationships in f_2, and at $(3\omega,3\omega)$, caused by the inter-relationships in f_3. No peaks are observed at $(\omega,2\omega)$, $(\omega,3\omega)$ etc. because the frequency components in f_2 and f_3 are independent to each other and there are no phase relationships existed.



x(t)



y(t)

Fig. 3.2 Artificial Data : x(t) is a series with a fundamental of 5.4 Hz and its 2^{nd}, 3^{rd}, 4^{th}, 5^{th} and 6^{th} harmonics; y(t) is a linearly combination of two series f_2 and f_3. If f_0 is 5.4 Hz, f_2 is made up with the fundamental f_0, its 2^{nd}, 4th and 5th harmonics while f_3 is made up with the 3th and 6th harmonics of f_0 only.

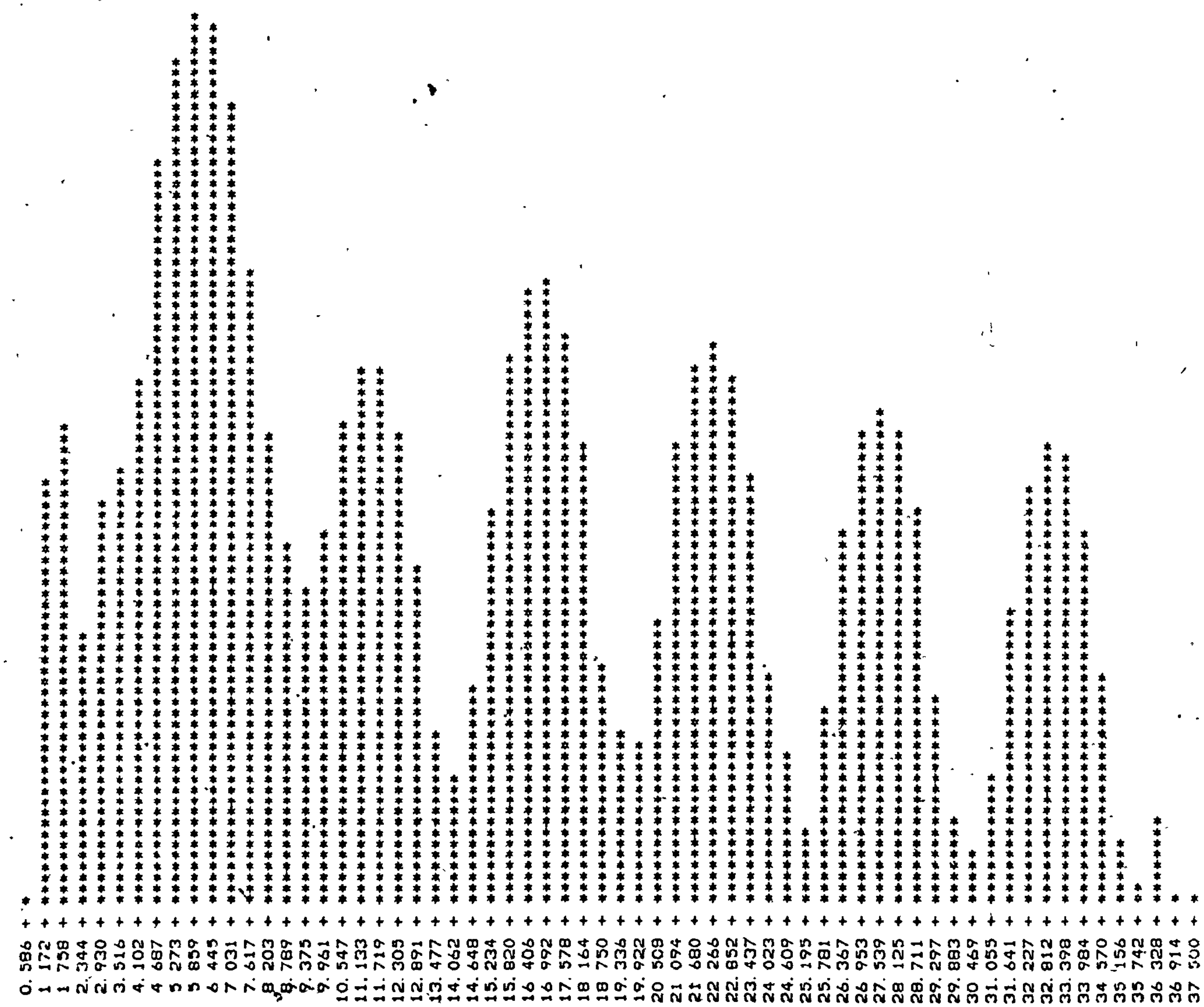


Fig. 3.6a Power spectrum of the artificial series x(t). Here peaks can be seen at all the six harmonics locations.

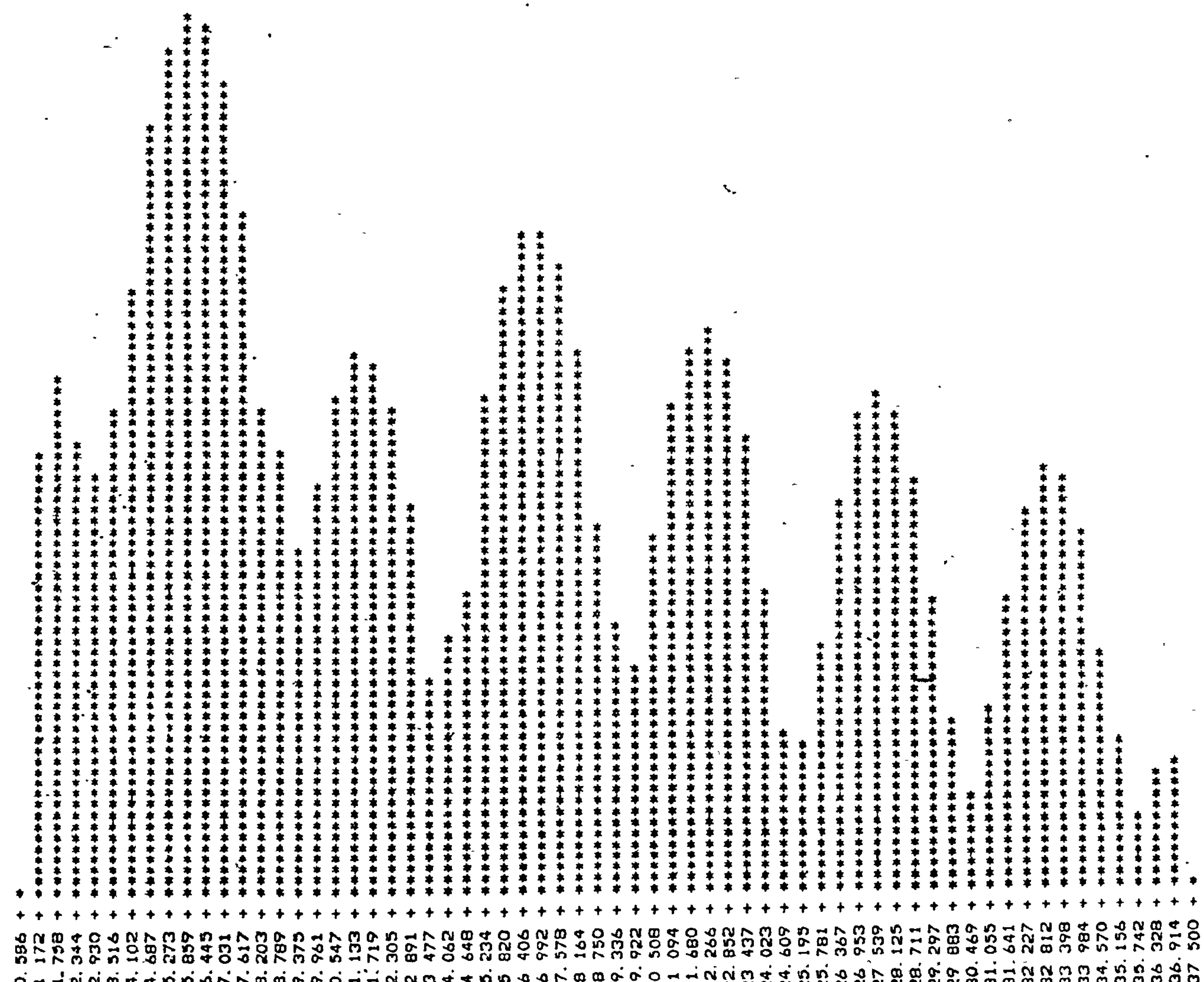


Fig. 3.6b Power spectrum of the artificial series y(t). Here peaks can also be seen at all the six harmonics locations as in Fig.3.6a.

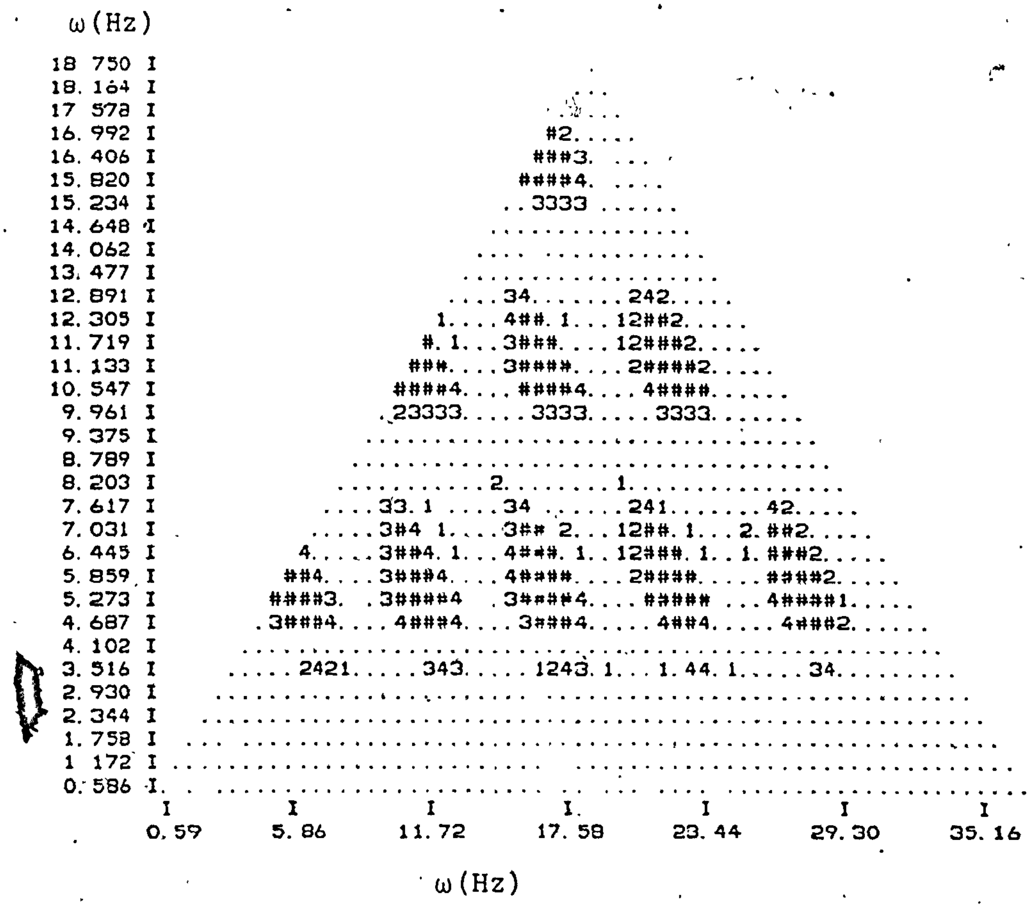


Fig. 3.7 The bispectrum of the artificial series x(t). Here components are phase-locked at (ω,ω,2ω), (ω,2ω,3ω), (ω,3ω,4ω), (ω,4ω,5ω), (ω,5ω,6ω), (2ω,2ω,4ω), (2ω,3ω,5ω), (2ω,4ω,6ω) and (3ω,3ω,6ω) with ω=5.4Hz.

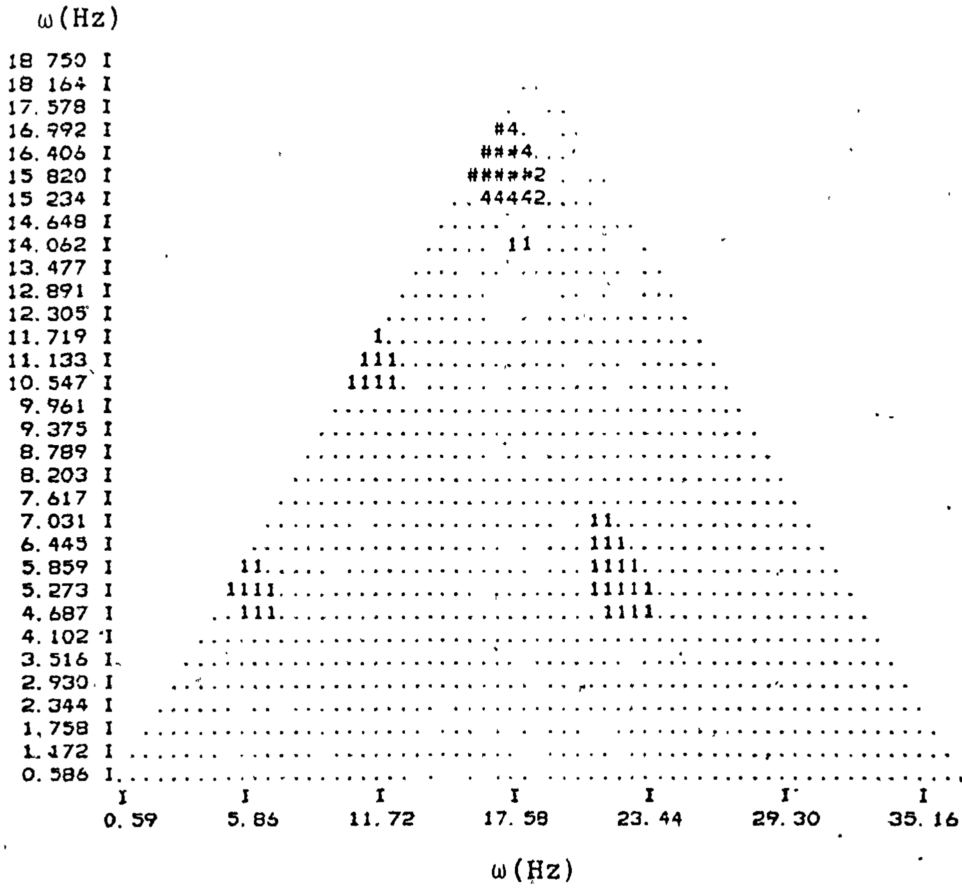


Fig. 3.8 The bispectrum of the artificial series y(t). Here components are phase-locked at (ω,ω,2ω), (ω,4ω,5ω), (2ω,2ω,4ω) caused by f_2 and also at (3ω,3ω,6ω) caused by f_3.

Hence one can use the bicoherence to identify distinct independent generators of the signal which may not be revealed by observing the power spectrum alone.

CHAPTER 4

THE ANALYSIS OF THE ELECTRICAL SIGNALS FROM THE DIGESTIVE TRACT

4.1 INTRODUCTION

In this project, electrical signals from various parts of the digestive tract were examined by the methods being proposed. Eight channels of signals with channel 1 being the activity taken from the stomach, channels 2 and 3 from the duodenum, channels 4, 5, and 6 from the distal small intestine and channels 7 and 8 from the colon were used for the analysis. Fig. 4.1 shows a typical portion of the 8 channels of signals under investigation. The stationarity and the normality characteristics of the signal were first investigated in order to examine the nature of the electrical control activity in various parts of the digestive tract. The results also give an idea of what would be expected when the data is analyzed by the bispectral method later. Here the method being used to carry out the analysis, the outcome and the discussion of the results are presented step by step in the subsequent sections of the chapter.

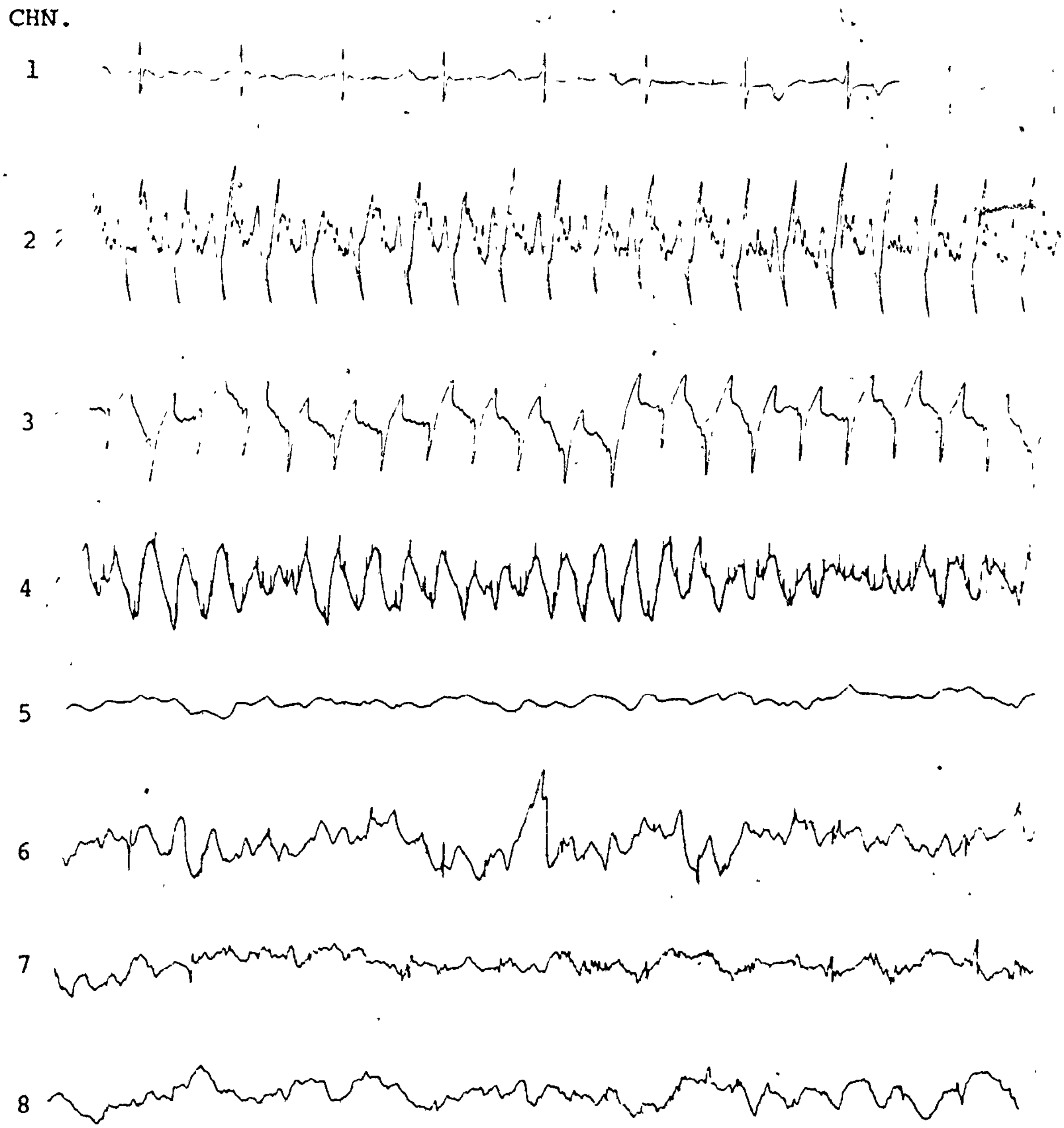


Fig. 4.1 A typical section of the 8 channels of signals used for the analysis here.

4.2 METHODS

At the beginning the signals were recorded by a FM tape recorder which was played back later to feed the signals into the NOVA minicomputer through an 8-channel analog-to-digital converter. Then they were stored in several data files on a disk. These data files were opened by a specifically designed program so that a desired amount of data could be retrieved from the data files. In this project, five minutes of signals were taken for the analysis so that the variation of the signals with time could be observed.

First of all, the stationarity characteristic of each channel of signal was examined by using the RUN test method described in Section 2.3. The algorithm resided in the program file is named 'RUNTST'. It starts off with a series of questions so that some flexibilities can be provided to the user. The questions are as follows :

" Enter the number of data points per section :
Enter the number of sections to be analyzed:
Enter the channel number :
Enter the initial data value :
Enter the sampling frequency :
Enter the significant level value : "

Here the program finds the stationarity of the signal section by section (non-overlapped) and at the end of the analysis, the percentage of the amount of non-stationary sections is calculated. Since each section was divided into 30 segments as mentioned in Section 2.3, the value entered for the number of data points per section must be an integer multiple of 30, e.g. 150. If the sampling frequency is 10Hz, the number of data points per section is 120 and 25 sections are analyzed, a total of 5 minutes of signals are investigated. The channel number in the questionnaire is the channel number of the data file opened. The initial data value is the starting point where the user wants to obtain the data from the file.

In this project, the value set for the number of data points per section was different from channel to channel. This is because the RUN test is designed for the examination of the stationarity of waveforms with irregular pattern. But if the signal has a regular waveform, e.g. a square wave or gastric electrical control activities, a visual inspection can evaluate the stationarity. However, if one insists to use the RUN test to examine those signals, a section with a long enough duration has to be taken, otherwise it will result in an insufficient number of runs and the algorithm will conclude the section to be non-stationary. Here a section of 30 seconds was taken for channel 1 (signal from the stomach), a section of 15 seconds for channels 2 to 6 (signals from the duodenum and the small intestine) and a section of 12 seconds for channels 7 and 8 (signal from the colon) as the basic unit. Since the sampling frequency for all the channels was set at 10 Hz, the values for the number of data points per section were as follows :

channel 1 : 300
channel 2 to 6 : 150
channel 7 and 8 : 120

After all the questions have been answered, the processing unit of the computer enters into a number of loops where in each loop, the stationarity of a section of signal is examined by the RUN test as described in Section 2.3. All the 8 channels of signals are tested by the same procedure and the result is tabulated in the next section when the outcome is being presented.

After the stationarity of the signal has been investigated, the normality characteristic becomes the next step of the analysis. The two algorithms utilized for the normality test are : (1) 'CHIFIT' - the chi-square goodness-of-fit test and (2) 'NORTST' - the Fisher's normality test. For these two tests the same set of questions as that of the RUN test is asked. However, because of the different natures of the normality tests, it is not necessary to assign a different value to each channel for the number of data points per section. Hence a section with the same duration (15 seconds) was set for all the channels. After the test, the result is again tabulated in the same manner as that of the RUN test and it is presented in the next section.

Then the data was passed over to the bispectral method for a further investigation of the behaviour of the electrical signals from the digestive tract. Two separate analyses were performed here : (1) the bispectral analysis for a single channel of signal and (2) the cross-bispectral analysis for two independent channels of signals. This was done because some of the objectives of this project were to find out the inter-relationships between frequency components, the behaviour of the possible generators of the signals and the interaction between two independent channels of signals. Here the program named 'BIS1' computes the bicoherence (the normalized bispectrum) of a single channel of signal. Similar to the case of the stationarity and normality tests, it starts off with a series of questions as follows :

" Enter the sampling frequency :
Enter the channel number :
Enter the initial data point :
Enter the number of data points :
Enter the number of 50 percent overlapped sections :
Enter the number of FFT points (max.1024) : "

There are a few points that are worth mentioning here concerning the above questions. The channel number asked is the channel number in the data file 'MDAT, not those that have been used before. 'MDAT' is created by a program called 'MCS'. It rearranges the original data file and puts the data into a separate data file, i.e. 'MDAT'. Then the algorithm 'BIS1' opens the file and retrieves the required data there. The number of data points is the length of the signal that the user wants to analyze at one time. It was set to be 600 here (which is equivalent to a one minute period for a 10 Hz sampling frequency). Since a smoothing and averaging procedure has to be done during the computation of a bispectrum, a further piece of information is needed. 50 percent overlapped sections of signal are being taken one at a time and the resultant bi-periodogram is averaged. But there is not any definite value for the number of 50 % overlapped sections which are to be divided among the 600 data points. It has to be an odd integer number with a minimum value of 3. From the experience of trial and error, either 7 or 9 50% overlapped sections being divided produces a good output. It was set to be 9 in this project. This number may not be good for some other kinds of signals. Hence a user should try a few analyses with different numbers of 50% overlapped sections before a definite value is assigned. Finally the number of FFT points depend on the resolution of the output diagram as well as the frequency range that is of interest. The resolution can be calculated from the following equation :

$$\text{Resolution(cpm)} = \frac{\text{sampling frequency}}{\text{\# of FFT points}} * 60 \qquad (4.1)$$

The number of points was taken to be 1024 here and hence the resolution is (10/1024)*60 = 0.586 cpm.

For the cross-bispectral analysis, exactly the same set of questions was asked at the beginning with one exception : now 2 channel numbers have to be inputted since this algorithm is designed to analyze 2 independent channels of signals. In this project, cross-bispectra were computed between alternate channels only. Additionally, a consecutive 5 minutes of signals were analyzed for each channel so that we can examine how the behaviour of the signal changes with time. The results are tabulated and discussed in the next two sections. Besides the table of results, a set of typical outputs (contours diagrams) are shown as well in the next section for each channel of signal.

4.3 RESULTS

4.3.1 STATIONARITY AND NORMALITY TESTS

Five minutes of signals from each channel have been analyzed in this project. The percentage of rejection on the null hypothesis, which proposes the signal under investigation to be stationary or normally distributed, has been calculated and listed in Table 4.1.

For the stationarity test, the whole section of signal from the stomach is stationary. This agrees with the visual inspection of the signal. For the signal taken

CHANNEL #	Stationarity (% of non-stationary sections) RUN test	Normality (% of not normally distributed sections) CHIFIT	FISHER
1	0.0	72.0	48.0
2	5.0	16.0	18.0
3	10.0	20.0	28.0
4	15.0	46.0	46.0
5	15.0	48.0	50.0
6	40.0	86.0	82.0
7	90.0	80.0	98.0
8	80.0	92.0	100.0

Table 4.1 Results from the stationarity and normality tests for five minutes of signals from each channel.

from the duodenum, that is, channels 2 and 3, 5.0 and 10.0 %, respectively, of rejection were produced. For the analysis taken from the small intestine (channels 4 and 5), 15% of rejection was recorded for both cases. All of these results show that the electrical signals from the duodenum and the upper part of the small intestine display a regular pattern of oscillations. However when it comes to the lower part of the small intestine, i.e. channel 6, 40 % of the signal was classified as non-stationary and irregular patterns started to appear. Finally, for the signal from the colonic part of the tract, i.e. channels 7 and 8, 90 and 80 %, respectively, of rejections were produced. Hence the electrical signals recorded from the colon are mostly non-stationary.

For the normality test, similar results were obtained from both the chi-square goodness-of-fit test and the Fisher's normality test. So the result described below applies to the outcome of both the tests (with the figures inside brackets being the result from the Fisher's normality test). For the signal taken from the stomach, i.e. channel 1, a 72 % (48 %) of rejection was produced. This large value is probably due to the nature of the signal, i.e. a sharp spike after a relatively long and quiet period. Although the spikes appear periodically, the overall signal cannot be classified as normally or randomly distributed. For the signal from the duodenum, i.e. channels 2 and 3, 16 % (18%) and 20% (28%), respectively, of the signal are not normally distributed. For the upper part of the small intestine, i.e. channels 4 and 5, the percentages of rejection of the null hypothesis are 46% (46%) and 48% (50%), respectively. Finally, for the signal from the lower part of the small intestine (channel 6) and that of the colon (channels 7 and 8), 86% (82%), 80% (98%) and 92% (100%), respectively, of the signal have been rejected from the hypothesis. So they do not satisfy the condition for a Gaussian random time series.

Both the stationarity and the normality tests for the signals display a similar trend from one channel to the next. Starting from the stomach, the signal displays a regular pattern. The signals taken from the duodenum as well as those from the upper part of the small intestine also display the stationary and normally distributed behaviour (with a slight increase in the percentage of rejection of the null hypothesis). But when it goes further down the the digestive tract, highly irregular patterns start to appear. The signals taken from there are classified as non-stationary and not normally distributed. This signifies that highly irregular oscillations occur in this area. More will be discussed about the interpretation of these results in the next section of the chapter.

4.3.2 BISPECTRAL ANALYSIS

Again, 5 minutes of signals from each channel were taken for the bispectral and cross-bispectral analysis. For the bispectral analysis, each minute of signal produces a distinct set of 8 contour map outputs for all the channels. As a result, a total of 40 distinct maps will comprise the the output for the analysis of 5 minute data from 8 channels. Here only the first minute set of output is presented and shown in Fig. 4.2. In each map the point with the maximum height has been chosen as a standard. The upper half of that height is then divided into five equal parts and plotted. The symbol '#' is designated as the top contour, then followed by '4', '3', '2' and '1' respectively. The symbol '.' represents the lower half of the contour map. The frequency triple value that produces a peak can only be chosen by approximation since the top contour usually occupies an area larger than a finite single point. Hence the mid-point of that area is approximated and the frequency triple of that point is assumed to be phase-locked.

The peaks of all the 40 maps have been inspected and the frequency triples that produce them are listed in Table 4.2. In this table, each triple is entered in the form of (f_1, f_2, f_1+f_2). Each of these sets of values signifies that the frequency triple f_1, f_2 and f_1+f_2 are phase-locked in the particular minute that is under

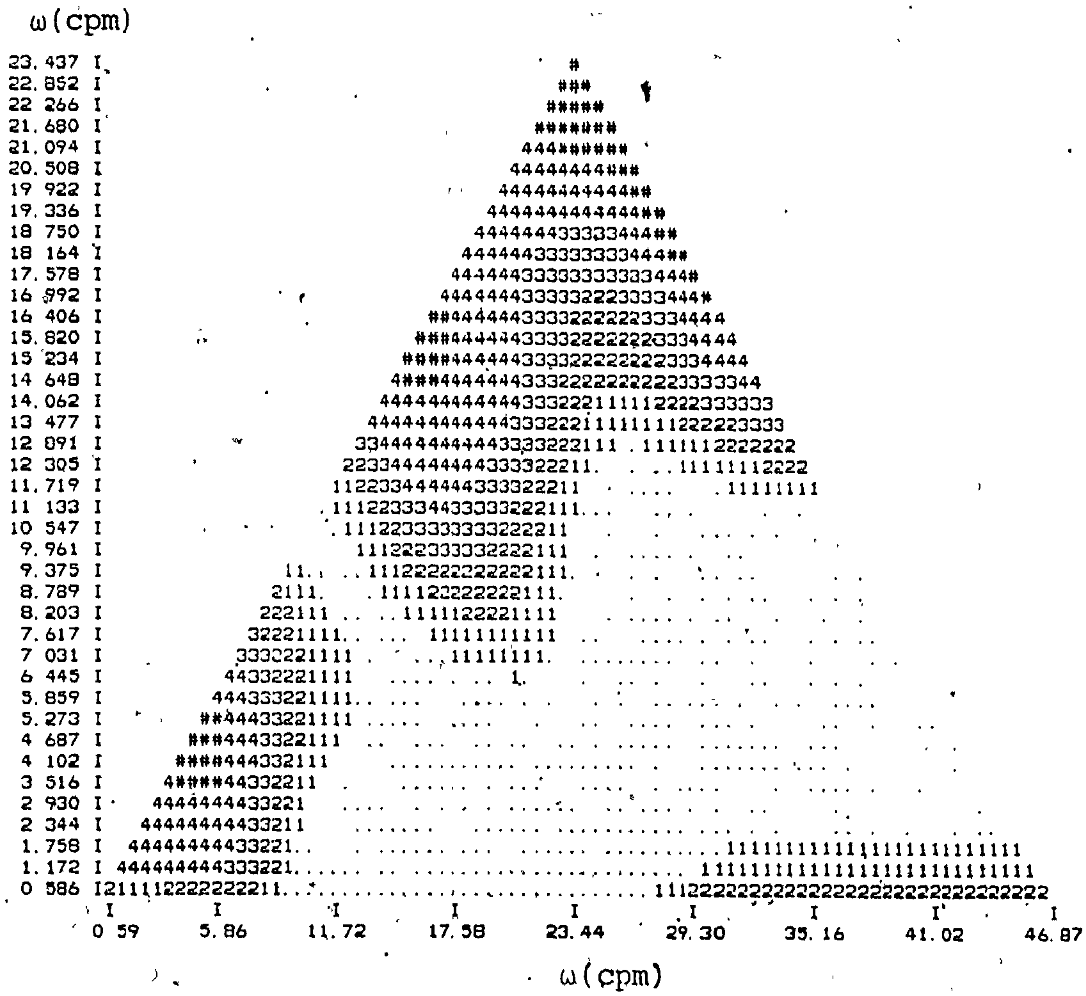


Fig. 4.2a A sample bispectrum of one minute of signal from channel 1. Here components are phase-locked at (15.8,15.8,31.6), (5.0,5.0,10.0) and (22.0,22.0,44.0).

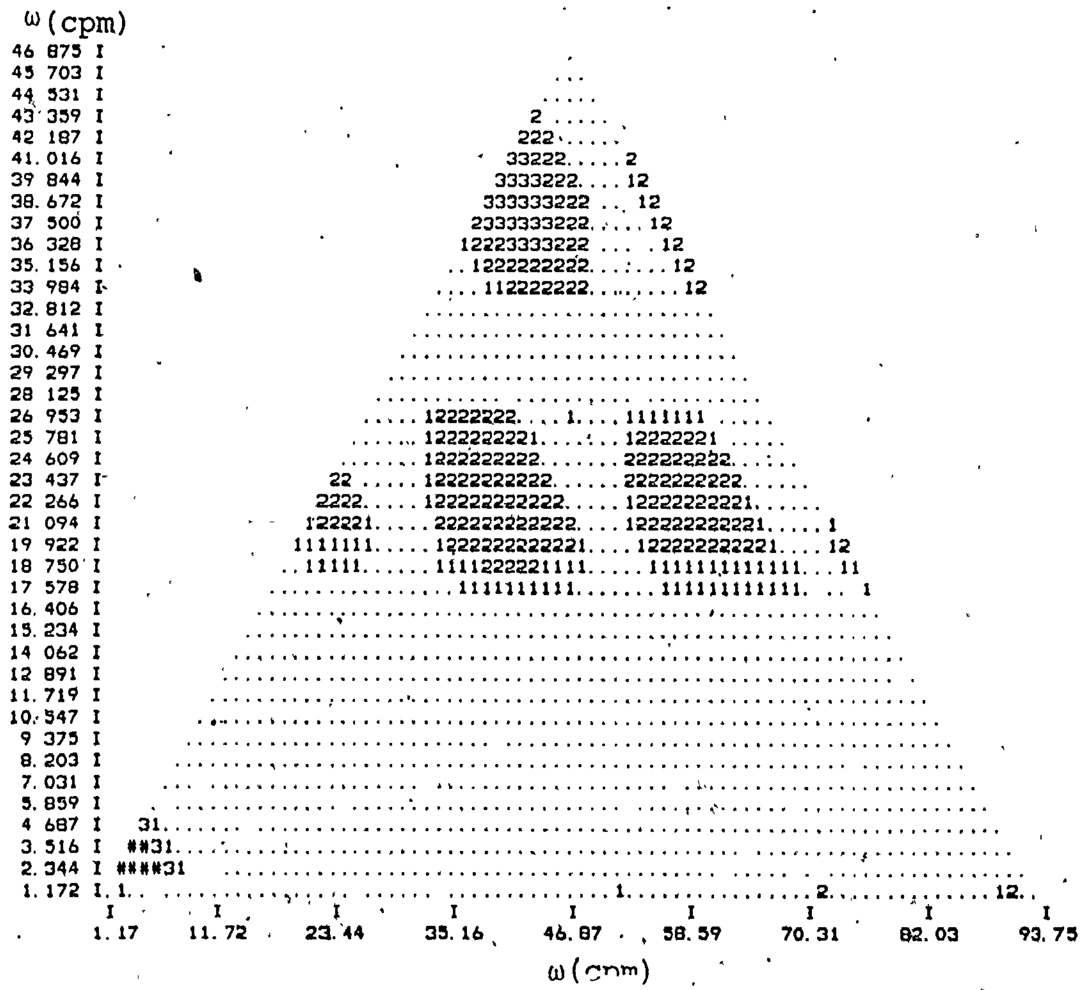


Fig. 4.2b A sample bispectrum of one minute of signal from channel 2. Here components are phase-locked at (20.0,20.0,40.0), (20.0,40.0,60.0), (20.0,60.0,80.0) and (40.0,40.0,80.0).

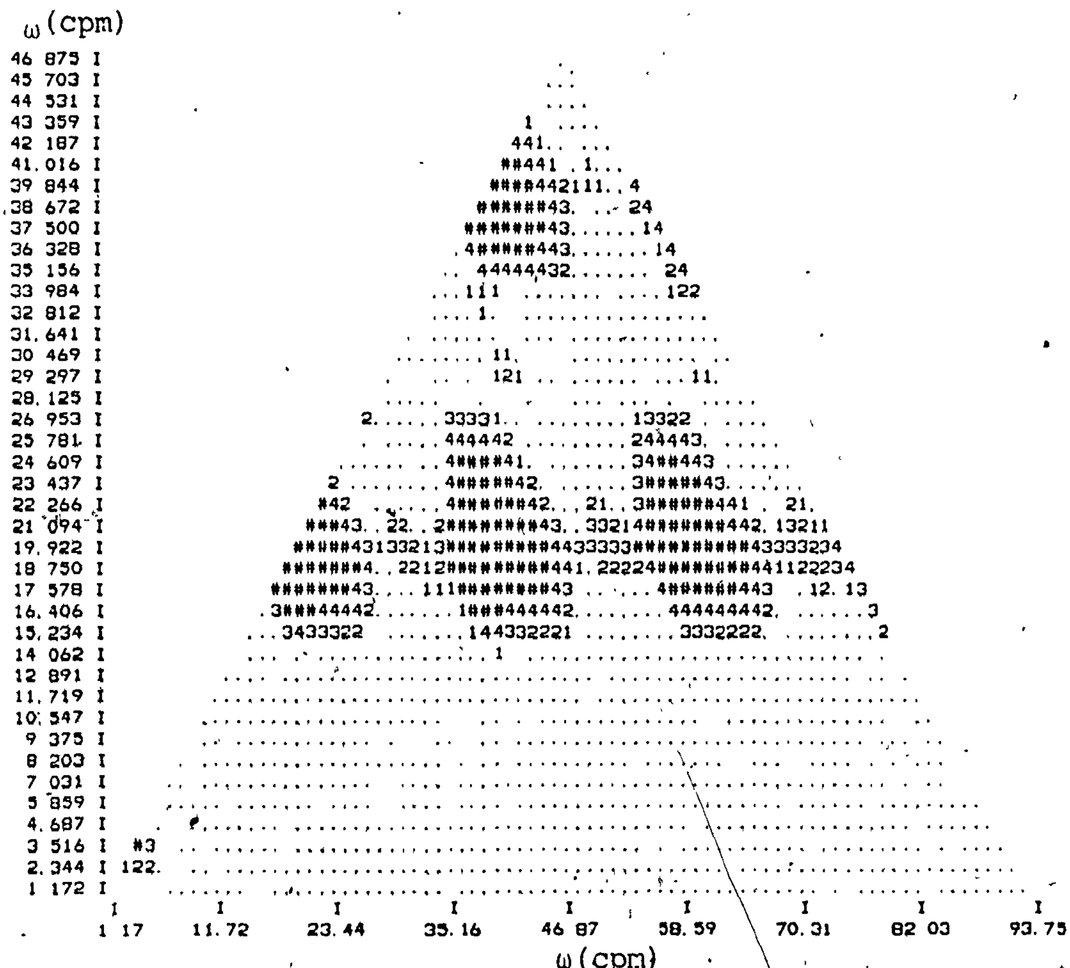


Fig. 4.2c A sample bispectrum of one minute of signal from channel 3. Here components are phase-locked at (19.0,19.0,38.0), (19.0,38.0,57.0), (19.0,57.0,76.0) and (38.0,38.0,76.0).

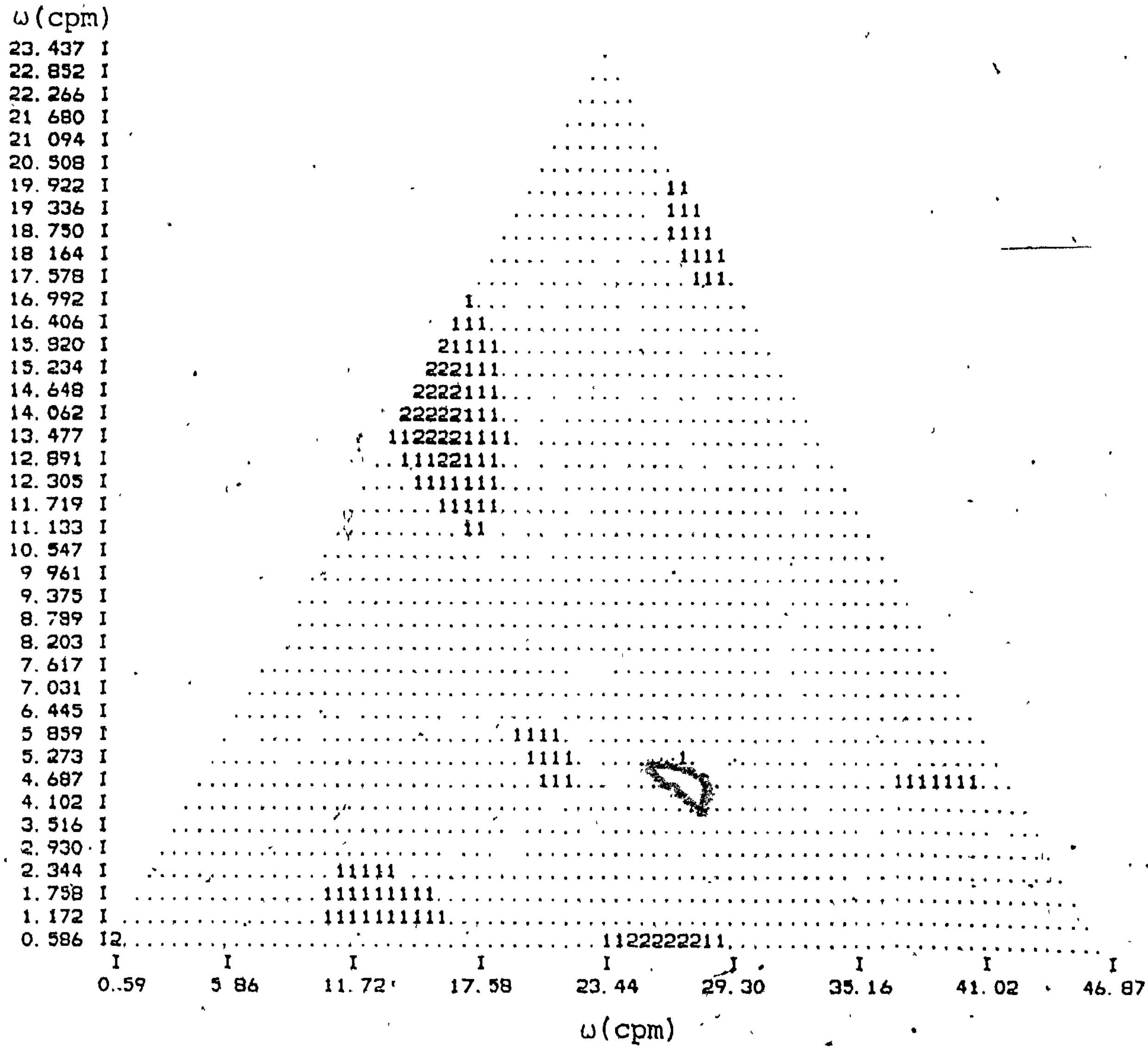


Fig. 4.2d A sample bispectrum of one minute of signal from channel 4. Here components are phase-locked at (15.8,15.8,31.6).

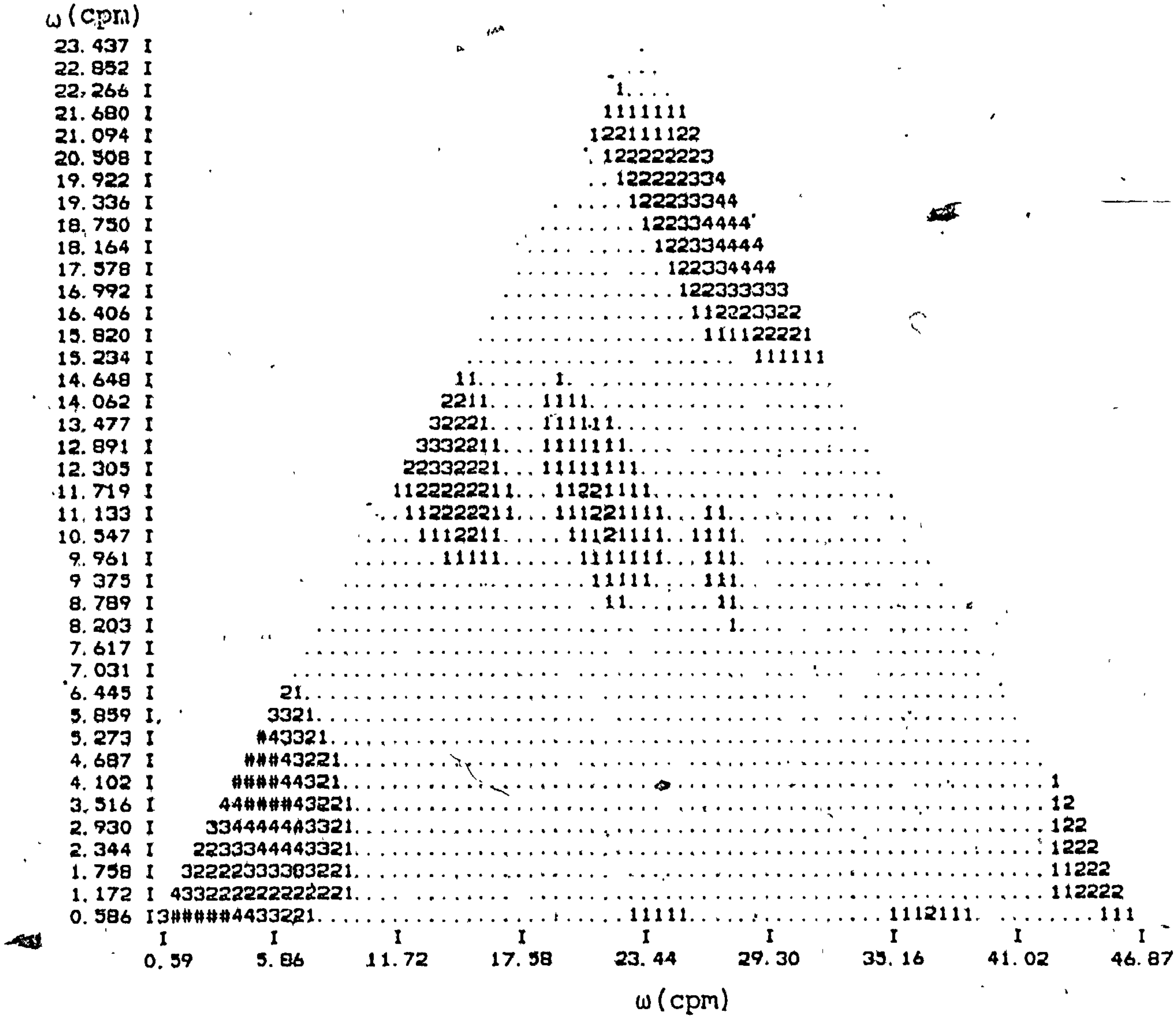


Fig. 4.2e A sample bispectrum of one minute of signal from channel 5. Here components are phase-locked at (4.5,4.5,9.0), (13.48,13.48,26.96), (27.13,18.2,45.33) and (22.5,11.5,34.0).

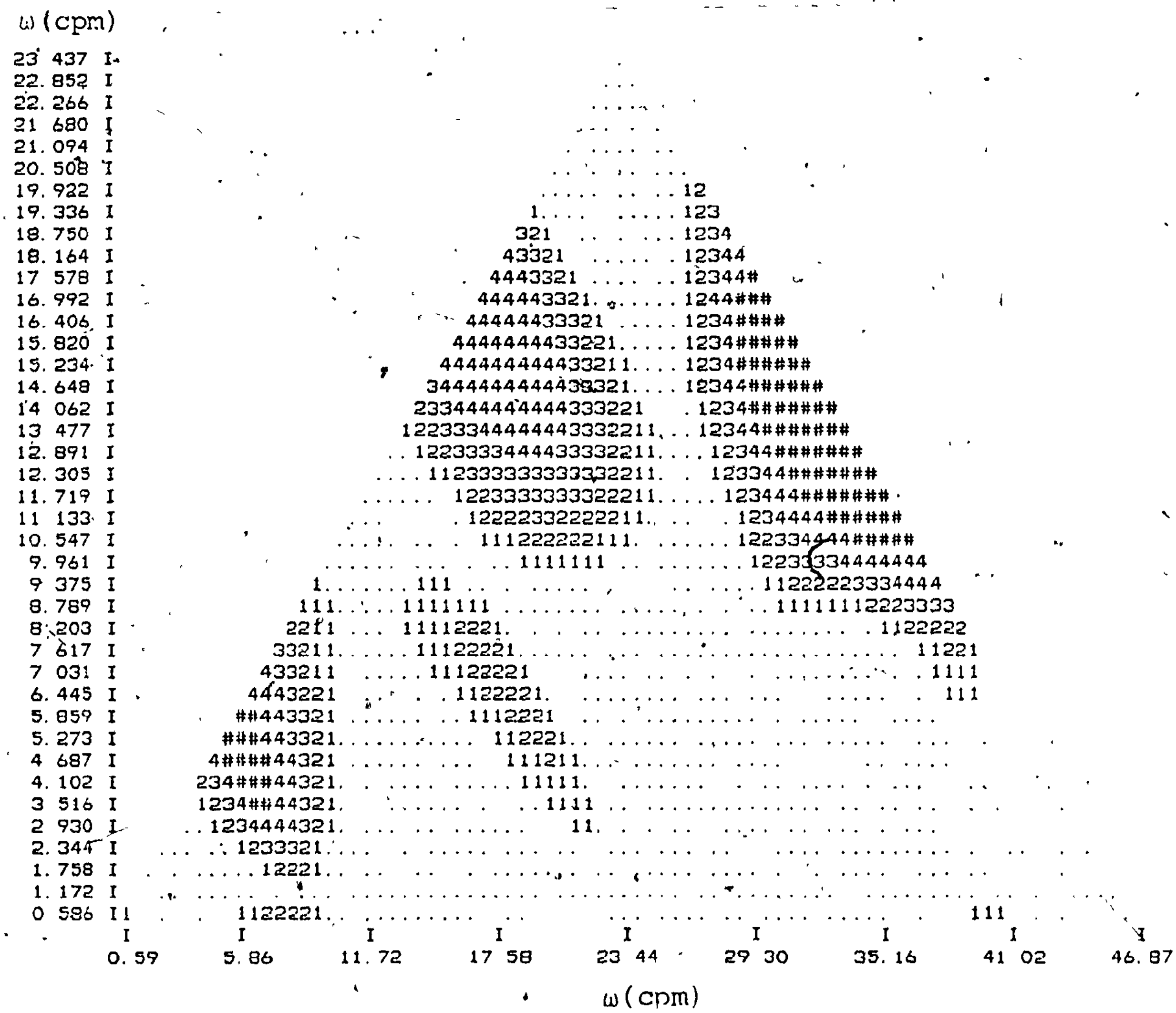


Fig. 4.2f A sample bispectrum of one minute of signal from channel 6. Here components are phase-locked at (5.9,5.9,11.8), (18.16,5.9,24.06), (15.8,15.8,31.6) and (31.05,15.8,46.85).

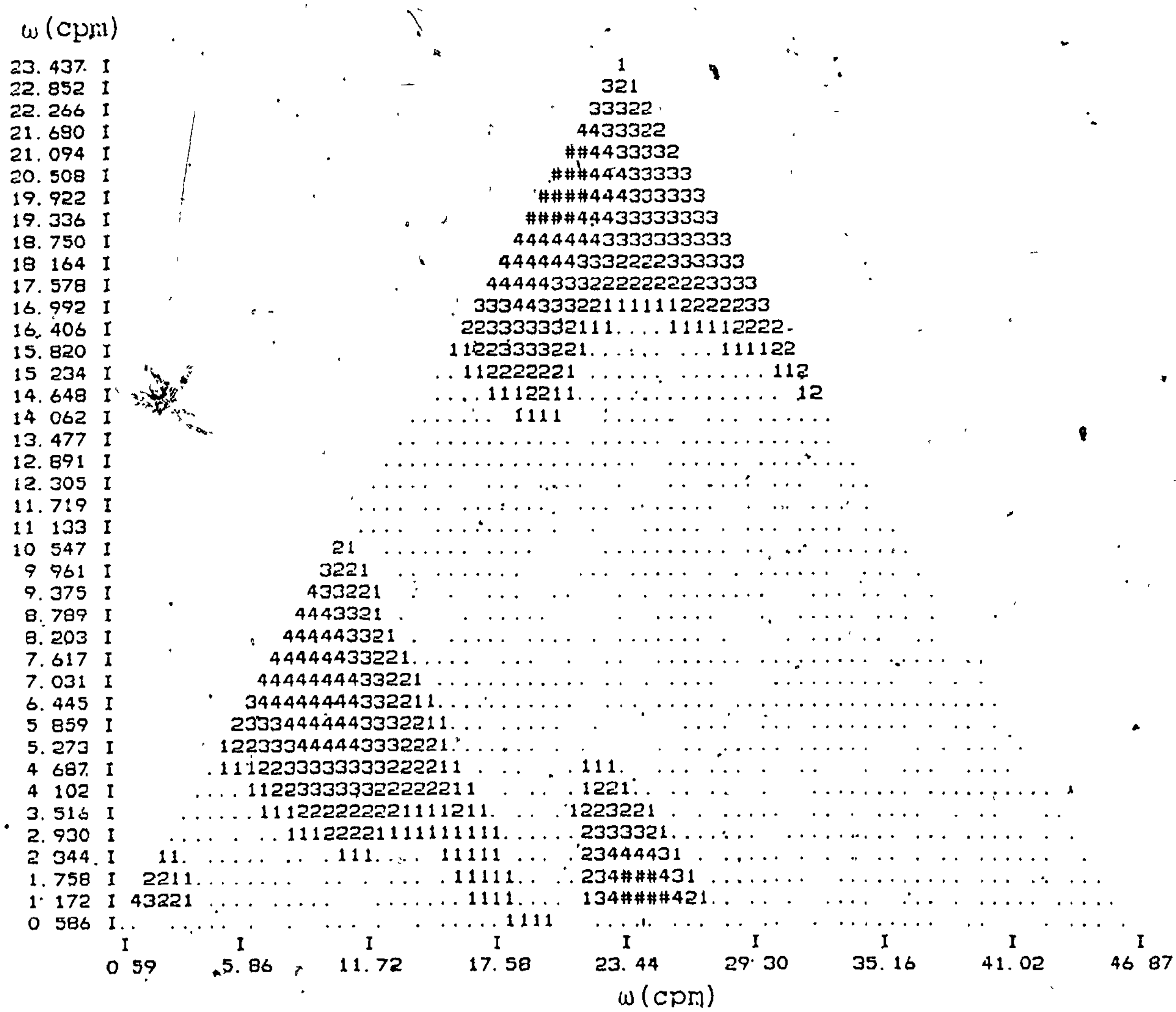


Fig. 4.2g A sample bispectrum of one minute of signal from channel 7. Here components are phase-locked at (8.2,8.2,16.4), (20.0,20.0,40.0) and (24.0, 1.76, 25.76).

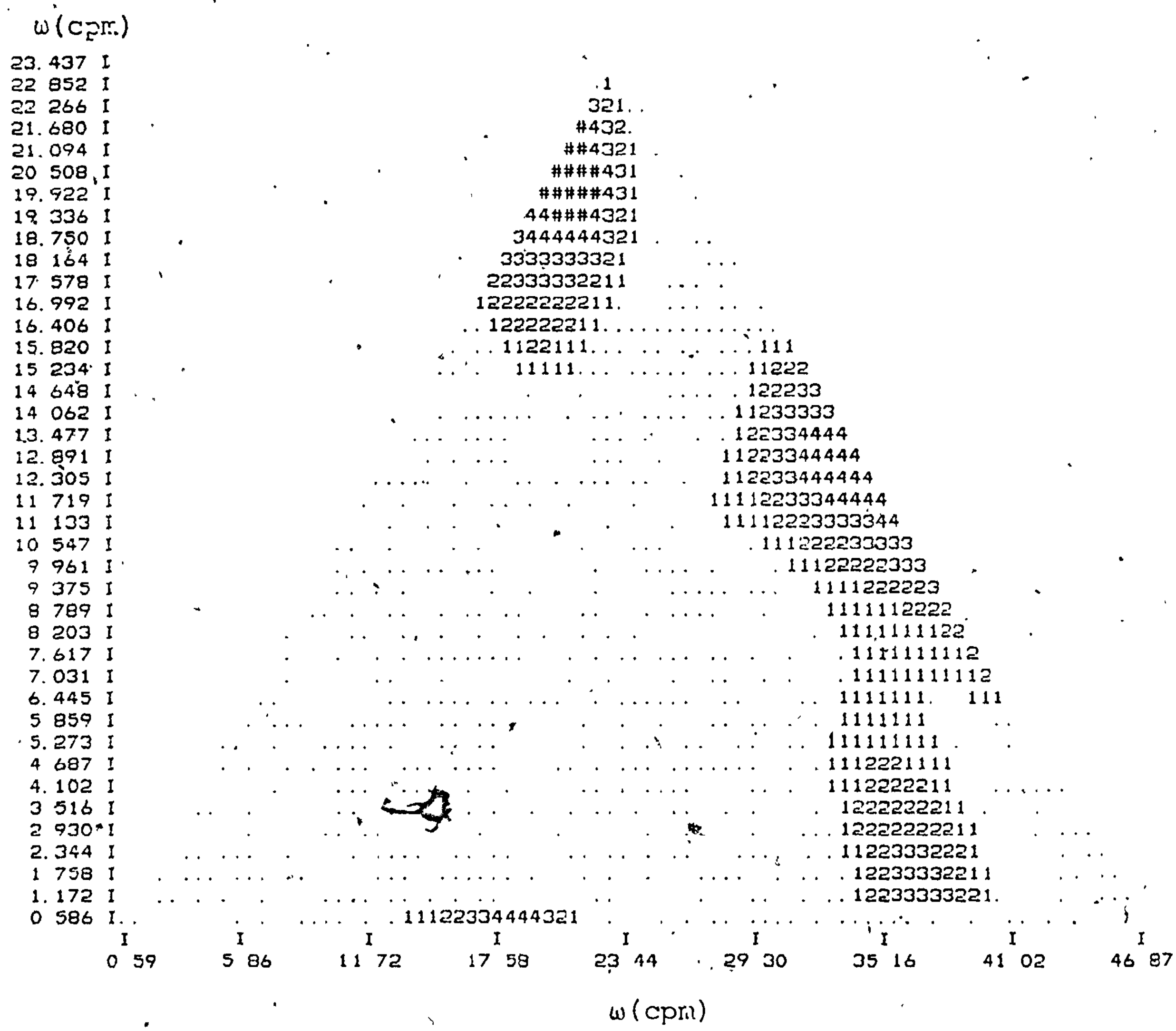


Fig. 4.2h A sample bispectrum of one minute of signal from channel 8. Here components are phase-locked at (20.7,20.7,41.4), (34.6,12.3,46.9) and (36.3,1.76,38.06).

CHANNEL	MINUTE # 1	2	3	4	5
1 (stomach)	(15.8,15.8,31.6) (5.0,5.0,10.0)	(15.8,15.8,31.6) (5.0,5.0,10.0)	(15.8,15.8,31.6) (5.0,5.0,10.0)	(15.8,15.8,31.6) (5.0,5.0,10.0)	(15.8,15.8,31.6) (5.0,5.0,10.0)
2 (duodenum)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)
3 (duodenum)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)	(19.0,19.0,38.0) (19.0,38.0,57.0) (19.0,57.0,76.0) (38.0,38.0,76.0)
4 (small intestine)	(15.8,18.8,31.6)	(4.4,4.4,8.8) (35.2,4.4,39.6) (8.8,8.8,17.6) (35.2,8.8,44.0) (15.8,15.8,31.6)	(4.8,4.8,9.6) (31.0,4.8,35.8) (24.6,14.6,39.2)	(4.5,4.5,9.0) (12.3,12.3,24.6) (27.0,18.0,45.0)	(8.0,8.0,16.0) (12.4,12.4,24.8)

Table 4.2 The results of the bispectral analysis of a single channel of signal for a duration of five minutes. Entries are the frequency triples (in cpm) which produce peaks (i.e. phase-locked) on the bicoherence contour maps.

CHANNEL	MINUTE # 1	2	3	4	5
5 (small intestine)	(4.5,4.5,9.0) (13.5,13.5,27.0) (27.1,18.2,45.3) (22.5,11.5,34.0)	(18.8,18.8,37.6) (23.4,23.4,46.8) (7.6,7.6,15.2) (29.0,7.6,36.6)	(7.6,7.6,15.2) (22.3,7.6,29.9)	(11.7,11.7,23.4)	(22.0,22.0,44.0)
6 (small intestine)	(5.9,5.9,11.8) (11.6,11.6,23.2)	(15.8,15.8,31.6) (31.1,15.8,46.9)	(5.9,5.9,11.8) (18.2,5.9,24.1) (15.8,15.8,31.6) (31.1,15.8,41.9)	(15.8,15.8,31.6) (31.1,31.1,62.2)	(15.8,15.8,31,6) (31.1,15.8,46.9)
7 (colon)	(18.0,18.0,36.0) (30.0,18.0,48.0)	(17.0,9.5,26.4) (31.1,9.6,40.7)	(17.0,17.0,34.0)	(8.2,8.2,16.4) (20.0,20.0,40.0) (24.0,1.76,25.8)	(8.0,8.0,16.0)
8 (colon)	(21.6,9.5,31.1) (21.6,21.6,43.2)	(12.8,12.8,25.6) (21.6,21.6,43.2)	(17.0,10.5,27.5) (27.0,9.5,26.5) (29.9,17.0,46.9)	(20.7,20.7,41.4) (34.6,12.3,46.9) (36.3,1.76,38.1)	(10.0,10.0,20.0) (17.0,26.3,43.3)

Table 4.2 (continued)

examination. For channel 1, the phase-locked triples are almost the same throughout the 5 minute period of analysis. This is not unusual because the signal from the stomach is 100% stationary. Here two oscillators are working together to produce the signal. The first one has a fundamental frequency of 5.0 cpm and a second harmonic with a frequency of 10.0 cpm. In fact the fundamental frequency is the frequency of the regular spikes in the signal. The second oscillator is having a fundamental at 15.8 cpm and its second harmonic at 31.6 cpm. The oscillators are identified here because only the frequency triple (which are multiple integers to each other) from the same generator are phase-locked. For channels 2 and 3, the outcome is exactly the same. The possible reason for this is the electrodes that are used to pick up the electrical signals have been placed close together and also the duodenum in the digestive tract produces a constant and regular waveform throughout. All the harmonics of the fundamental of 19.0 cpm are detected here because the frequency component at 19 cpm and all the others that are at the integer multiples of 19.0 cpm, are phase-locked. This signifies that the electrical signal recorded from the duodenum is composed of a single non-sinusoidal wave component with a fundamental frequency of 19.0 cpm and all its harmonics. Now when it comes to the distal part of the small intestine, multiple sets of oscillators which shift their frequency once in a while become evident. For channel 4 during minute #1, only a fundamental at 15.8 cpm and its second harmonic at 31.6 cpm are present. During minute #2, another oscillator starts to function. However it does not have all its harmonics available. Only the fundamental at 4.4 cpm, the 2^{nd} harmonic at 8.8 cpm, the 4^{th} harmonic at 17.6 cpm, the 8^{th}, 9^{th} and 10^{th} harmonics at 35.2, 39.6 and 44.0 cpm respectively, are detected. The rest of the harmonics, i.e. the 3^{rd}, 5^{th}, 6^{th}, 7^{th} etc. are not present because no phase-locking is found between the fundamental, the 2^{nd} harmonic etc., with them. Besides this new oscillator, the one that was functioning during the last minute (with a fundamental at 15.8 cpm) is still working. During the 3^{rd} minute, the fundamental frequency of the new oscillator that was present during minute #2 shifts slightly to 4.8 cpm. Now only its 2^{nd} (9.6 cpm), 3^{rd} (14.6 cpm), 5^{th} (24.6 cpm) and 8^{th} (39.2 cpm) harmonics are present. Besides these, the fundamental component is also phase-locked with others at 31.0 and 35.8 cpm respectively. When it comes to minute #4, the oscillator shifts its fundamental back to 4.5 cpm and its 2^{nd} (9.0 cpm), 4^{th} (18.0 cpm), 6^{th}(27.0 cpm) and 10^{th} (45.0 cpm) harmonics are present. During this minute, a new generator comes into action with a fundamental frequency of 12.4 cpm and its 2^{nd} harmonic at 24.8 cpm. For the final minute of analysis, the oscillator with the fundamental at 12.4 cpm is still functioning. Additionally, another new generator with components at 8.0 and 16.0 cpm joins this oscillation. Notice that during this 5 minute period no oscillator can persist throughout. They just keep on coming and going and irregular patterns start to appear.

For channel 5, a similar situation is found. During the first minute, an oscillator with a fundamental at 4.5 cpm, its 2^{nd} harmonic at 9.0 cpm, 3^{rd} harmonic at 13.48 cpm, 4^{th} harmonic at 18.2 cpm, 6^{th} harmonic at 27.0 cpm and 10^{th} harmonic at 45.33 cpm are present. Besides this, the frequency triple (22.5, 11.5, 34.0) are also phase-locked. This signifies that another oscillator with a fundamental at 11.5 cpm, the 2^{nd} and 3^{rd} harmonics at 22.5 and 34.0 cpm respectively, is functioning. During minute #2, the former oscillator shifts its fundamental from 4.5 cpm to 18.75 cpm (its 4^{th} harmonic). A couple of other new generators also join the work force. The first one is at 23.4 and 46.8 cpm. The second one is at 7.6 and 15.2 cpm. The fundamental of this later oscillator also phase-locks with the frequency components at 29.0 and 36.6 cpm. During the third minute, this new oscillator with the fundamental at 7.6 cpm persists. Along with the fundamental component, its 2^{nd} (15.2 cpm), 3^{rd} (22.3 cpm) and 4^{th} (29.9 cpm) harmonics are also found. During minute #4; the former oscillator (at 7.6 cpm) disappears and one of the oscillators (at 11.7 cpm) from minute #1 functions again. Only its second harmonic at 23.4 cpm is accompanying it. During the fifth minute, the oscillator shifts its fundamental to a frequency near its second harmonic at 22.0 cpm. Its second harmonic is also shifted to 44.0 cpm.

For channel 6, during minute #1, a single oscillator with a fundamental at 5.9 cpm is present. Its 2^{nd} (at 11.8 cpm) and 4^{th} (at 23.2 cpm) harmonics are also available. During the second minute, another independent oscillator is present. It has a fundamental at 15.8 cpm, the 2^{nd} harmonic at 31.6 cpm and the 3^{rd} harmonic at 46.85 cpm. In fact, this is the same oscillator that was functioning in channel 1 (throughout) as well as in channel 4 during minutes #1 and #2. During the third minute, both the oscillators from the first two minutes are working together. Their frequency sets are (5.9, 11.8, 18.16 and 24.06 cpm) and (15.8, 31.6 and 46.9 cpm) respectively. During minutes #4 and #5, only the second oscillator, i.e. the one with a fundamental at 15.8 cpm, persists. All the results obtained from channels 4, 5 and 6 have a common phenomenon : oscillators only work for a short period of time and they sometimes even shift their fundamental frequencies.

For channels 7 and 8, oscillators appear even more irregularly. Most of them do not have harmonics after the second, e.g., for channel 7 : during minute #1 -(18, 18, 36), minute #3 -(17, 17, 34), minute #4 - (8.2, 8.2, 16.4), (20, 20, 40), minute #5 - (8, 8, 16) and for channel 8 : during minute #1 - (21.6, 21.6, 43.2), minute #2 - (21.6, 21.6, 43.2), (12.8, 12.8, 25.6), minute #4 - (20.7, 20.7, 41.4) and minute #5 - (10.0, 10., 20.0). Some of the oscillators even interact and phase-lock with other frequency components that are not harmonically related to their fundamentals, e.g. in channel 7, during minute #1 (30, 18, 48), minute #2 (17, 9.5, 26.4), (31.05, 9.6, 40.65) and in channel 8 during minute #1 (21.6, 9.5, 31.1), minute #3 (17.0, 10.5, 27.5), (29.9, 17.0, 46.9), and at minute #5 (17.0, 26.3, 43.3). Of course there are also frequency triples phase-locked together but are not either harmonically related or belong to any strong oscillator. These frequency components are just acting on their own but having the same phase angle, e.g. for channel 8 during minute #4 (34.6, 12.3, 46.9) and (36.3, 1.76, 38.06). From these results, one can see that the colonic part of the digestive tract has a very irregular electrical activity pattern and most oscillators are working on a local basis for a short time.

4.3.3 CROSS-BISPECTRAL ANALYSIS

For the cross-bispectral analysis, each of the two successive channels are analyzed at one time. This method gives one a further insight about the relationship between signals recorded from two independent sites. It provides a clue to whether an oscillator on one site has anything to do with the oscillator on the other site by having one or more of its harmonics present there. The result has to be interpreted by making reference to the result of the bispectral analysis of a single channel. because it identifies the sources of that particular signal. However, if any of the elements on the frequency triple, which are phase-locked together, is not from the generators of the signal itself, its presence could signify an independent coupling between it and the components on the other signal or the triple are just having the same phase angle. To identify the exact cause of the phase-locking phenomenon, some other forms of analyses have to be done. Anyway, the cross-bispectrum does identify the phase-locking triple and sometimes the coupling between two or three oscillators in two independent channels of signals.

As mentioned in Chapter 3, four different combinations of cross-bispectrum can be computed. If the frequency triple are of the form $\{X_1(f_1)X_2(f_2)X_3(f_1+f_2)\}$, the combinations are

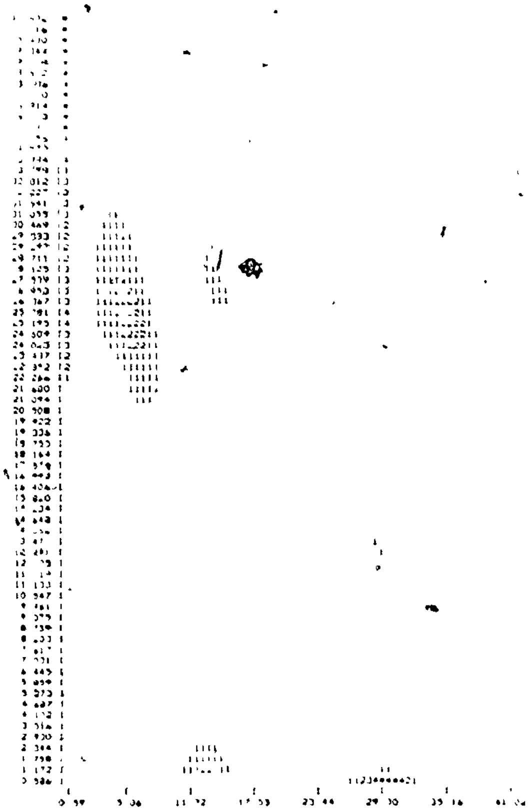
i) $X_1(f_1)X_2(f_2)X_1(f_1+f_2)$

ii) $X_1(f_1)X_1(f_2)X_2(f_1+f_2)$

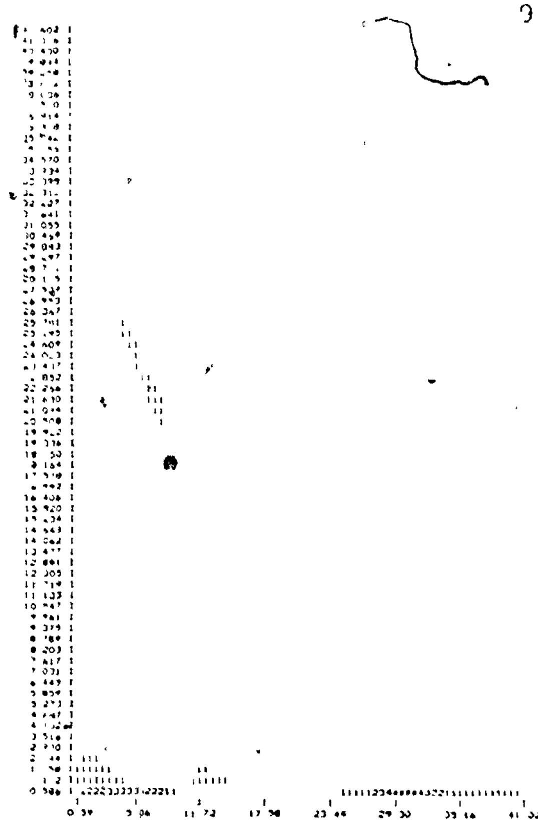
iii) $X_1(f_1)X_2(f_2)X_2(f_1+f_2)$

iv) $X_2(f_1)X_2(f_2)X_1(f_1+f_2)$

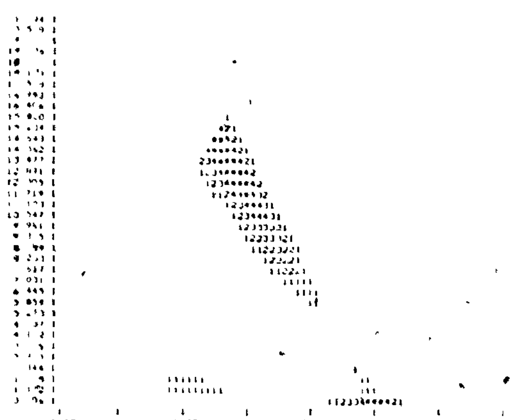
where X_1 is a frequency component from the first signal and X_2 is another component from the second signal. In fact the first two combinations, i.e. $X_1X_2X_1$ and $X_1X_1X_2$ can be analyzed together because they give the same kind of of information : the phase-locking (or sometimes coupling) between two components from the first signal and another component from the second. The same thing can also be done to the other two combinations, i.e. $X_1X_2X_2$ and $X_2X_2X_1$: a coupling between one component from the first channel and two others from the second. In this project all the four combinations have been computed and the output of the first minute of analysis is shown in Fig. 4.3. The complete result is presented on Table 4.3 by listing the peaks of all the maps. Here combinations $X_1X_2X_1$ and $X_1X_1X_2$ are grouped together and the other two combinations are put into another group. The first group, headed by $X_1X_1X_2$, can be interpreted as whether there is any generator in the first channel driving a frequency component in the second channel and the second group, headed by $X_1X_2X_2$, is just the vice versa case of the first. Three distinct possibilities arise from these results : 1), the generator of the signal itself may drive a component on the other channel; 2), an independent generator (not that for the signal) may drive a component on the other channel and 3), the frequency triple may just have the same phase angle by chance. In Table 4.3, all the frequency triples are marked with



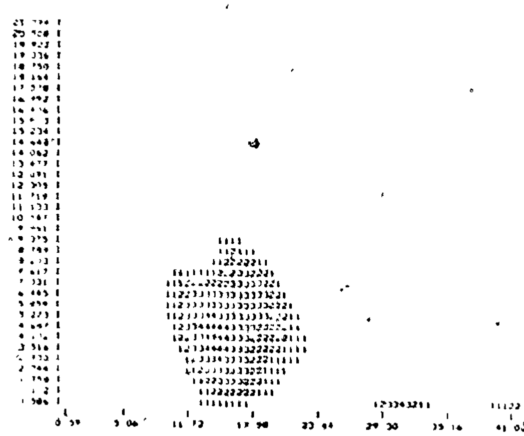
(i) Combination $X_1X_2^*X_1$



(ii) Combination $X_1X_2X_2$

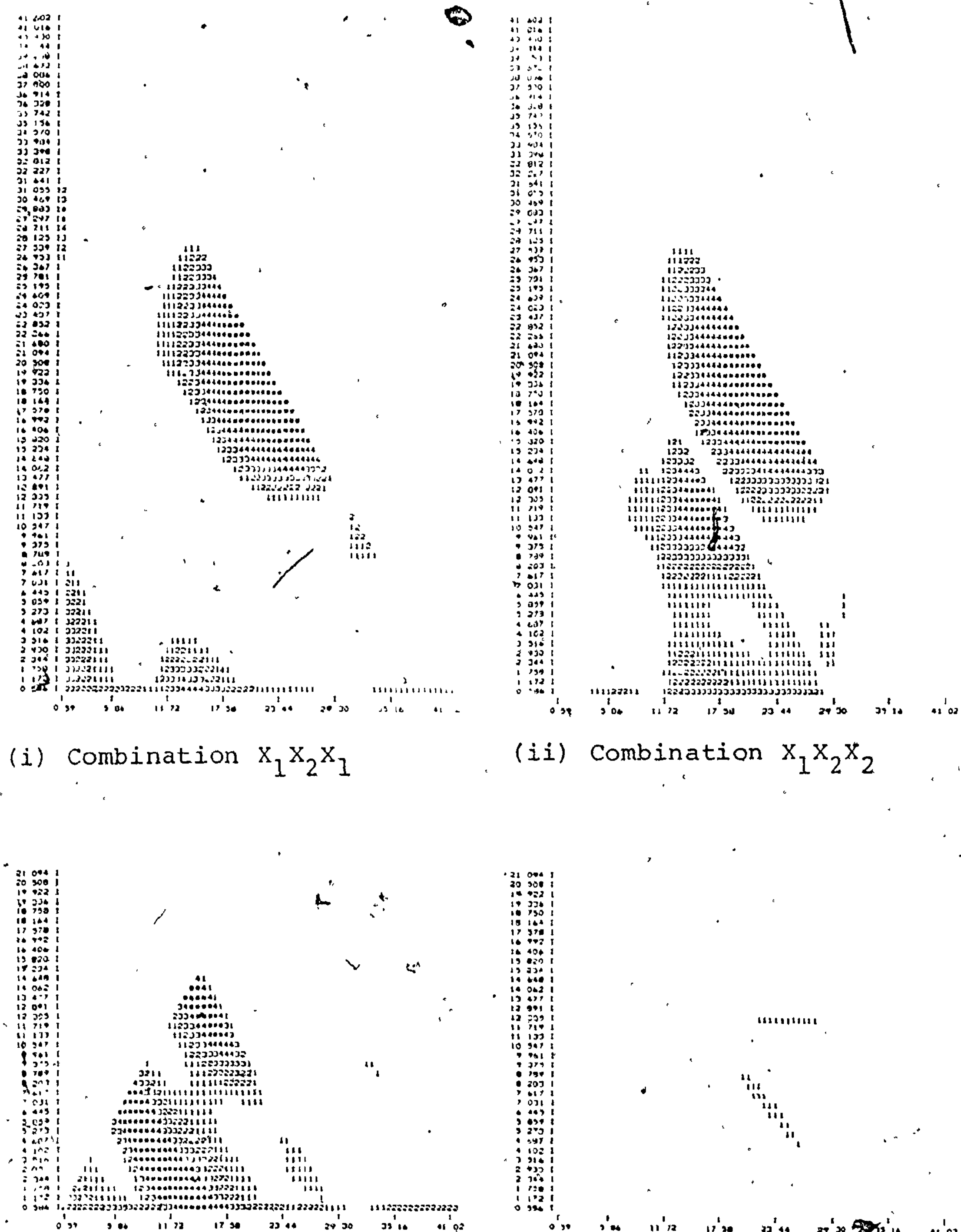


(iii) Combination $X_1X_1X_2$



(iv) Combination $X_2X_2X_1$

Fig. 4.3a A sample cross-bispectrum between channels 1 and 2.



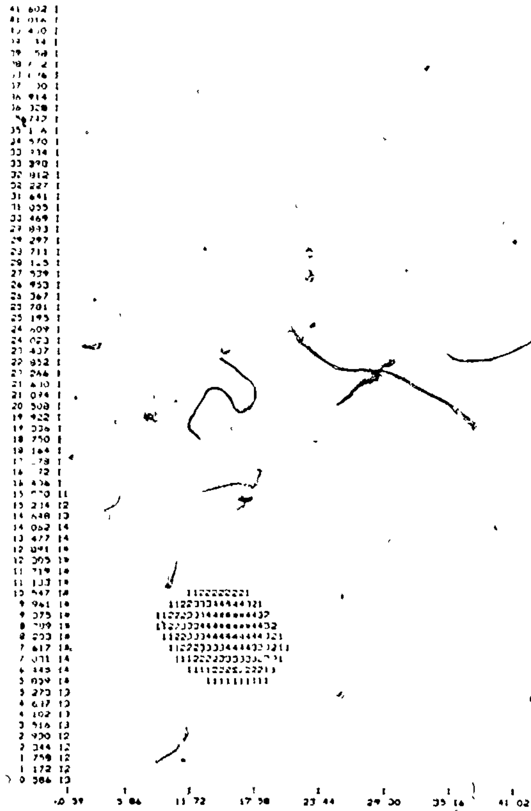
(i) Combination $X_1X_2X_1$

(ii) Combination $X_1X_2X_2$

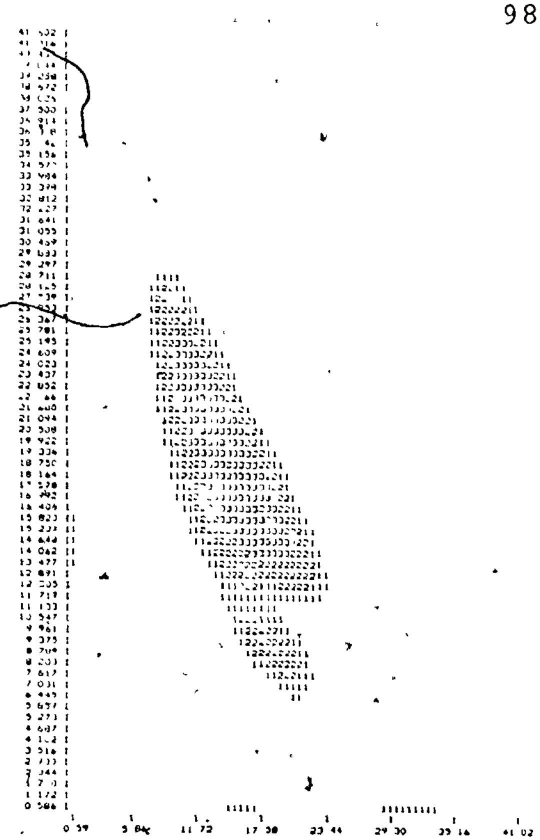
(iii) Combination $X_1X_1X_2$

(iv) Combination $X_2X_2X_1$

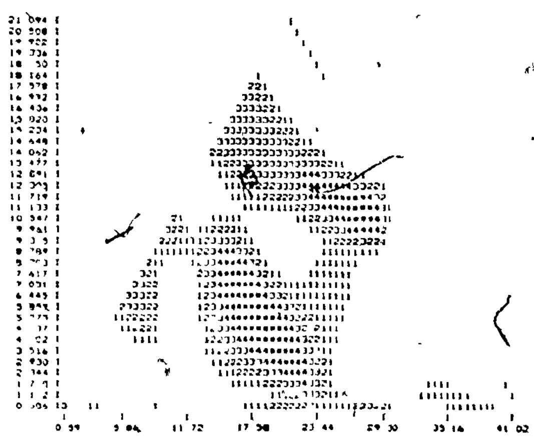
Fig. 4.3b A sample cross-bispectrum between channels 2 and 3.



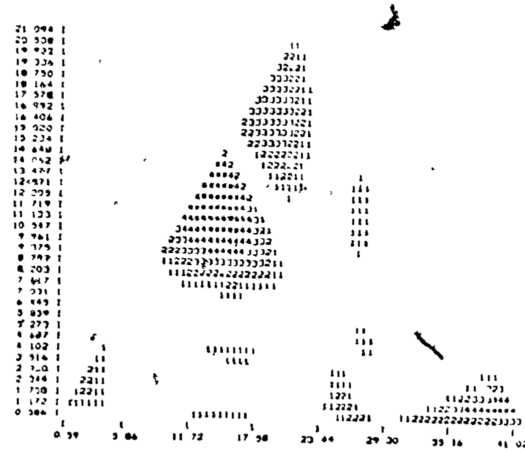
(i) Combination $X_1X_2X_1$



(ii) Combination $X_1X_2X_2$

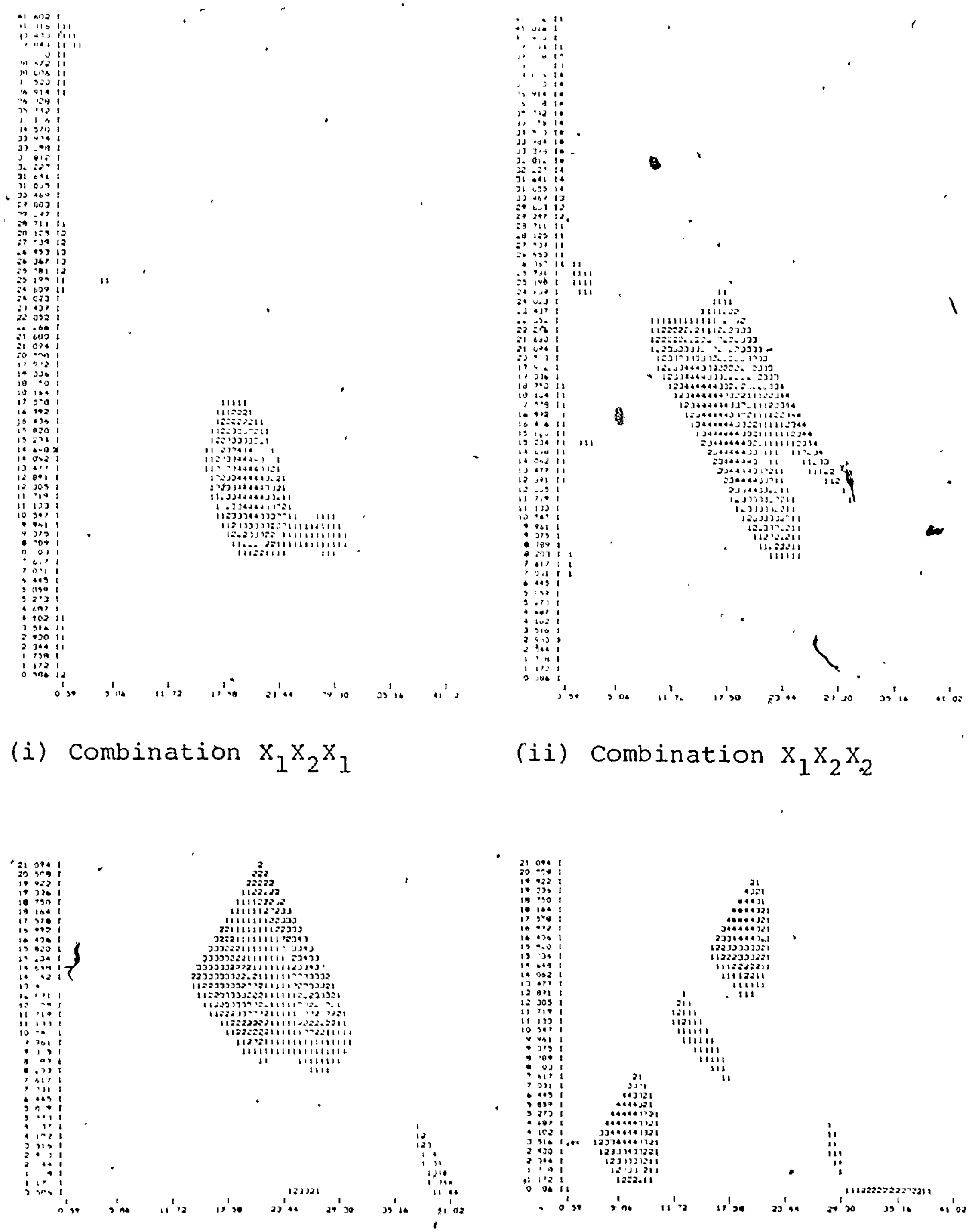


(iii) Combination $X_1X_1X_2$



(iv) Combination $X_2X_2X_1$

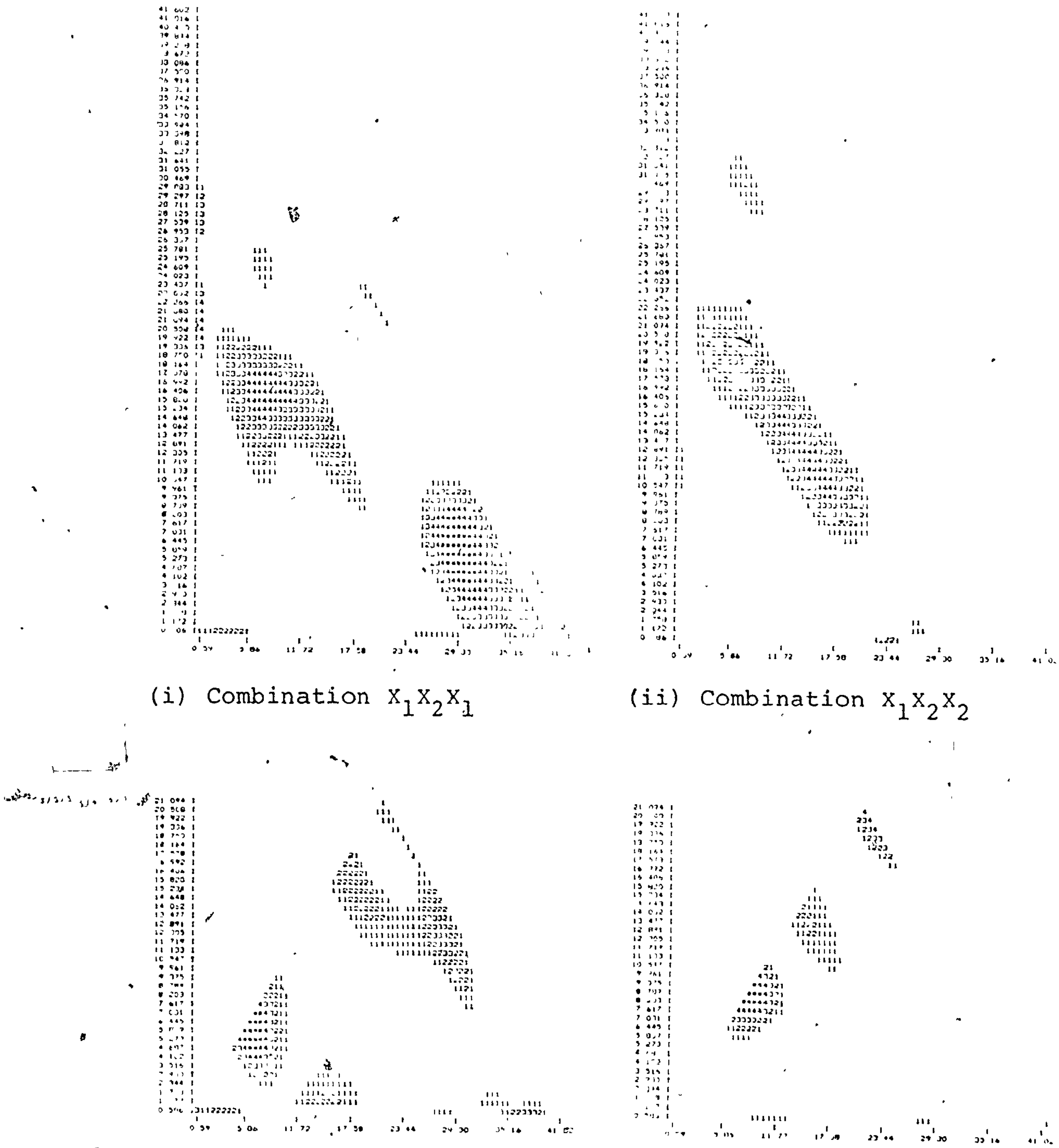
Fig. 4.3c A sample cross-bispectrum between channels 3 and 4.



(i) Combination $X_1X_2X_1$ (ii) Combination $X_1X_2X_2$

(iii) Combination $X_1X_1X_2$ (iv) Combination $X_2X_2X_1$

Fig. 4.3d A sample cross-bispectrum between channels 4 and 5.



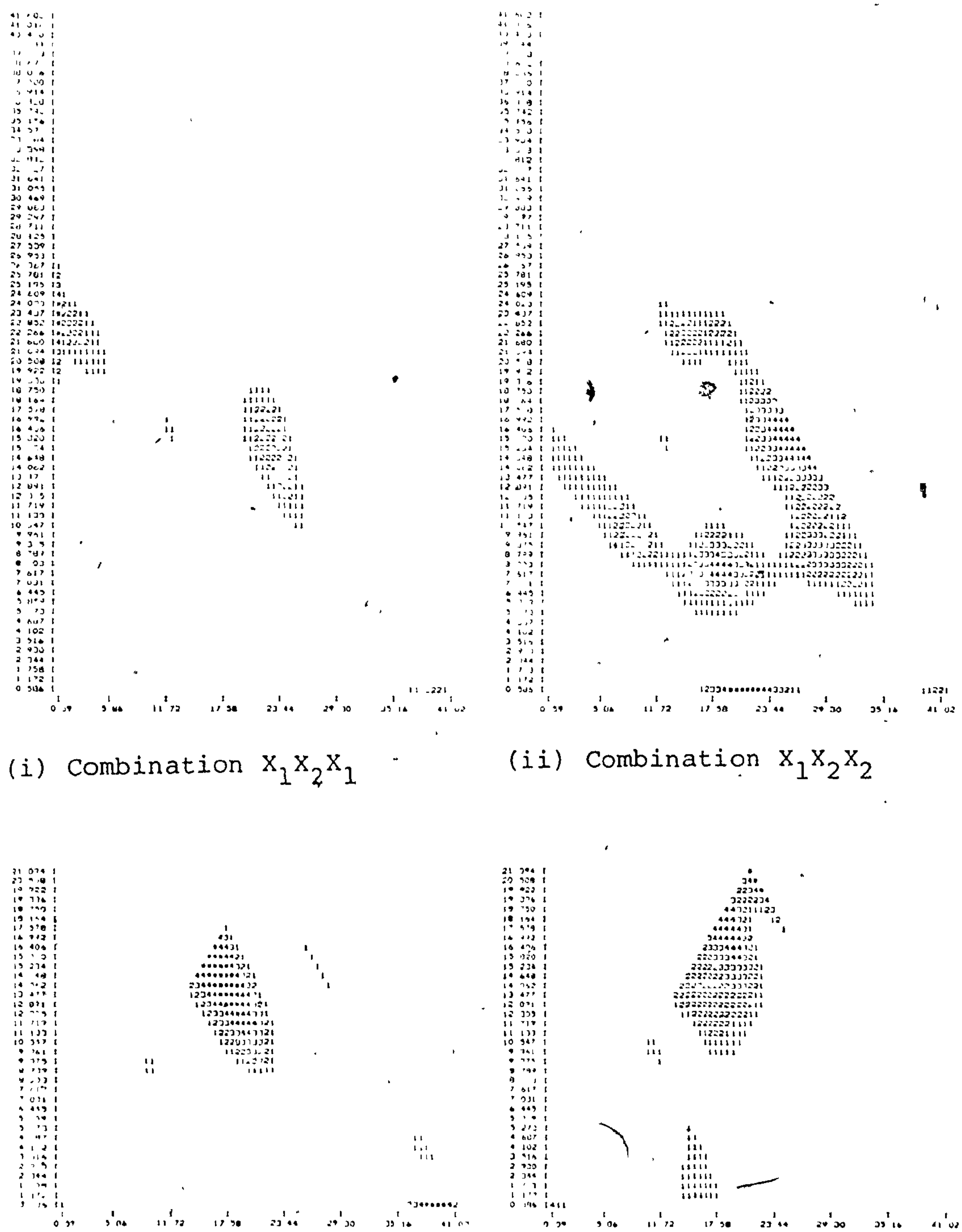
(i) Combination $X_1X_2X_1$

(ii) Combination $X_1X_2X_2$

(iii) Combination $X_1X_1X_2$

(iv) Combination $X_2X_2X_1$

Fig. 4.3e A sample cross-bispectrum between channels 5 and 6.



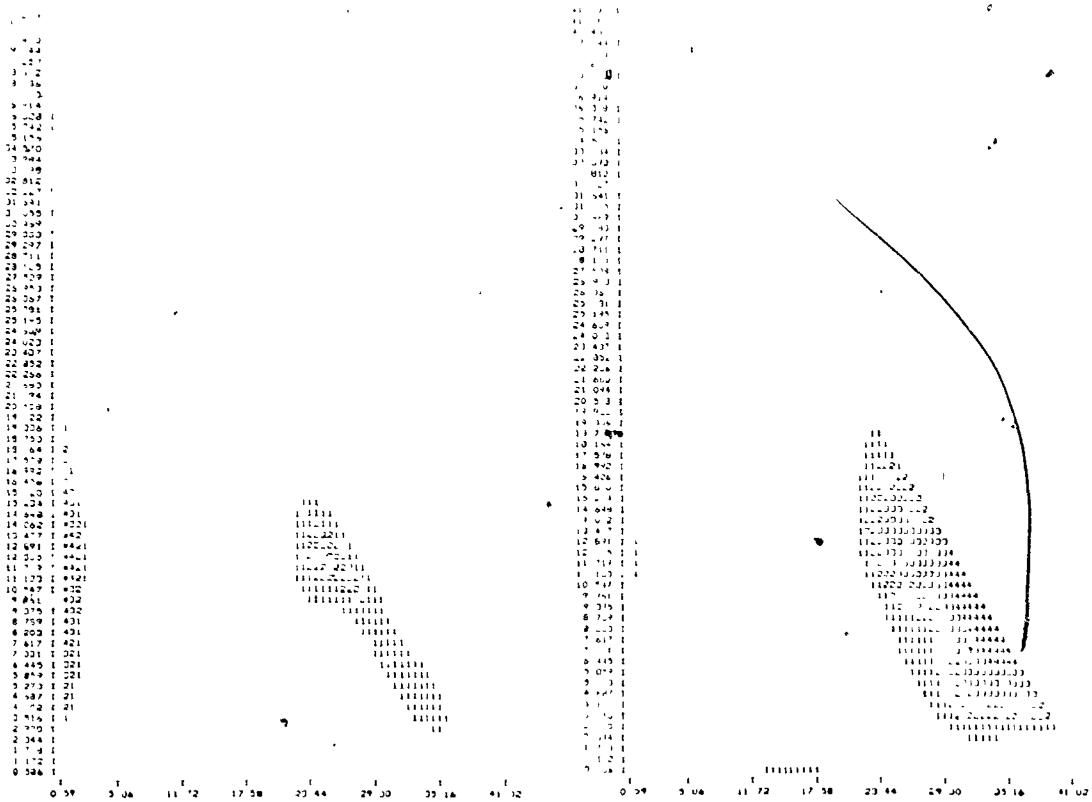
(i) Combination $X_1X_2X_1$

(ii) Combination $X_1X_2X_2$

(iii) Combination $X_1X_1X_2$

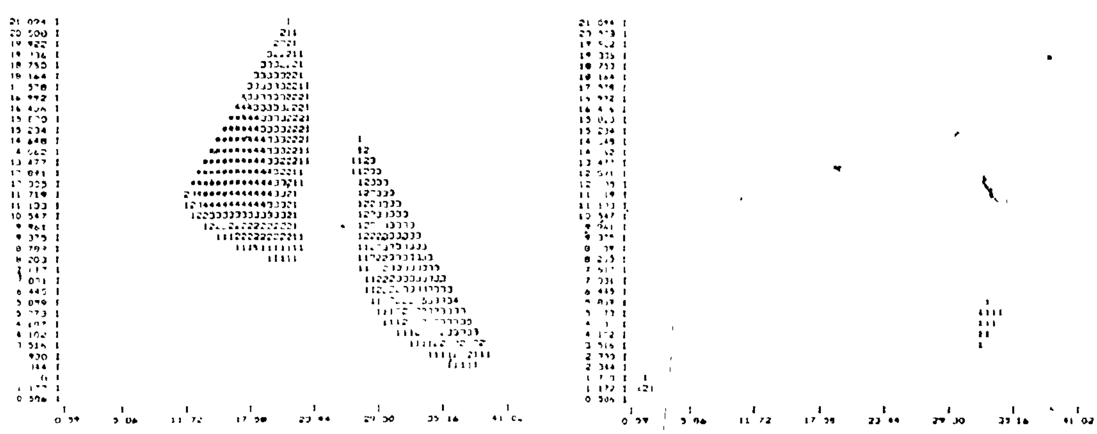
(iv) Combination $X_2X_2X_1$

Fig. 4.3f A sample cross-bispectrum between channels 6 and 7.



(i) Combination $X_1X_2X_1$

(ii) Combination $X_1X_2X_2$



(iii) Combination $X_1X_1X_2$

(iv) Combination $X_2X_2X_1$

Fig. 4.3g A sample cross-bispectrum between channels 7 and 8.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(5.9,31.7,25.8)_1$ $(14.8,14.8,29.6)_1$	$(6.2,24.4,30.6)_3$ $(19.6,14.8,4.8)_3$
2	$(7.03,7.03,14.06)_2$	$(6.2,24.8,31.0)_3$ $(25.6,12.8,12.8)_2$ $(39.3,26.5,12.8)_3$
3	$(14.8,14.8,29.6)_1$ $(14.8,29.6,14.8)_1$ $(7.6,7.6,15.2)_2$ $(11.1,11.1,22.2)_2$	$(15.2,7.6,7.6)_2$ $(29.6,14.8,14.8)_1$
4	$(18.75,42.75,24.0)_3$ $(5.0,29.0,24.0)_1$ $(4.1,4.1,8.2)_2$	$(5.0,24.0,29.0)_3$ $(25.2,20.5,4.7)_3$
5	$(4.5,24.5,20.0)_3$ $(14.0,14.0,28.0)_1$	$(12.3,21.1,33.4)_3$ $(10.5,15.2,25.7)_3$ $(20.0,10.0,10.0)_2$

Fig. 4.3a The results of the cross-bispectrum between channels 1 and 2 for a period of five minutes. Entries are the frequency triples (in cpm) which produce peaks (i.e. phase-locked) on the cross-bicoherence contour maps. Here the subscripts assign each triple to one of the three possible groups described on page 95.

MINUTE #	COMBINATION	
	$X_1X_1X_2$	$X_1X_2X_2$
1	$(19.0,19.0,38.0)_1$ $(13.8,13.8,27.6)_2$	$(19.0,19.0,38.0)_1$ $(15.8,10.7,26.5)_3$
2	$(7.03,7.03,14.06)_2$ $(19.0,19.0,38.0)_1$	$(19.0,19.0,38.0)_1$
3	$(19.0,19.0,38.0)_1$ $(10.0,10.0,20.0)_2$	$(19.0,19.0,38.0)_1$ $(15.8,26.4,42.2)_3$
4	$(19.0,19.0,38.0)_1$ $(5.86,5.86,11.72)_2$ $(11.1,11.1,22.2)_2$	$(19.0,19.0,38.0)_1$ $(15.2,26.5,41.7)_3$
5	$(19.0,19.0,38.0)_1$ $(8.2,8.2,16.4)_2$ $(15.2,15.2,30.4)_2$	$(19.0,19.0,38.0)_1$

Table 4.3b The results of the cross-bispectrum between channels 2 and 3.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(16.4,25.1,8.7)_3$ $(17.5,5.9,23.4)_3$ $(27.0,11.7,38.7)_3$	$(15.0,15.0,30.0)_2$ $(25.8,12.9,12.9)_1$ $(35.2,17.6,17.6)_2$
2	$(27.2,11.9,39.1)_3$ $(6.5,18.8,12.3)_1$ $(6.0,36.5,30.5)_3$ $(13.1,13.1,26.2)_2$	$(6.8,25.0,31.8)_3$ $(10.54,5.27,5.27)_2$ $(35.74,29.3,6.44)_3$
3	$(13.2,13.2,26.4)_2$ $(29.3,3.8,33.1)_3$	$(3.5,30.5,36.0)_3$ $(23.4,11.7,11.7)_2$ $(30.0,27.0,3.0)_3$ $(42.0,39.0,3.0)_3$
4	$(13.0,13.0,26.0)_2$	$(13.0,13.0,26.0)_2$ $(24.5,5.9,30.4)_3$ $(42.2,26.4,15.8)_3$
5	$(28.1,3.8,31.9)_3$ $(14.6,4.0,18.6)_3$	$(24.6,2.93,27.53)_3$ $(30.04,27.5,2.64)_3$ $(29.2,14.6,14.6)_2$

Table 4.3c The results of the cross-bispectrum between channels 3 and 4.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(15.5,15.5,31.0)_1$ $(25.8,15.5,41.3)_3$ $(20.5,43.0,12.5)_3$	$(17.0,15.8,32.8)_3$ $(26.5,15.8,42.3)_3$ $(24.6,12.3,12.3)_2$ $(36.32,18.16,18.16)_2$ $(10.54,5.27,5.27)_2$
2	$(14.06,28.12,14.06)_3$ $(5.2,5.2,10.4)_2$ $(10.5,10.5,21.0)_2$ $(30.0,7.6,37.6)_3$	$(20.5,17.6,38.1)_3$ $(20.5,4.1,24.6)_3$ $(10.0,7.0,17.0)_3$ $(34.6,7.6,42.2)_3$ $(20.0,10.0,10.0)_2$
3	$(26.4,30.5,4.1)_3$	$(19.3,15.8,35.1)_3$ $(30.511.7,42.2)_1$
4	$(14.0,14.0,28.0)_2$ $(19.0,19.0,38.0)_2$	$(4.2,25.2,29.4)_3$ $(24.0,10.0,34.0)_3$ $(27.27,22.0,5.27)_3$ $(32.97,27.7,5.27)_3$
5	$(12.0,32.0,20.0)_3$	$(21.0,10.5,10.5)_1$

Table 4.3d The results of the cross-bispectrum between channels 4 and 5.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(8.2,24.6,16.4)_3$ $(6.45,6.45,12.9)_2$ $(17.0,17.0,34.0)_2$ $(27.5,12.3,39.8)_3$ $(29.3,35.2,5.9)_3$	$(13.5,13.5,27.0)_2$ $(28.0,14.0,14.0)_2$ $(42.0,21.0,21.0)_2$ $(17.58,8.79,8.79)_2$
2	$(14.0,36.3,22.3)_3$ $(9.0,9.0,18.0)_2$ $(21.0,21.0,42.0)_2$ $(15.5,24.0,8.5)_3$	$(12.3,17.6,29.9)_3$ $(14.06,7.03,7.03)_2$ $(34.0,17.0,17.0)_2$ $(41.5,30.0,11.5)_3$
3	$(18.8,37.6,18.8)_1$ $(21.0,21.0,42.0)_1$	$(26.8,13.4,13.4)_2$
4	$(18.8,37.6,18.8)_1$ $(7.61,7.61,15.2)_1$	$(28.6,14.3,14.3)_2$
5	$(9.4,35.8,26.4)_3$	$(35.8,26.4,9.4)_3$

Table 4.3e The results of the cross-bispectrum between channels 5 and 6.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(4.1,15.8,11.7)_3$ $(12.0,24.0,12.0)_3$ $(20.0,32.9,12.9)_3$ $(14.65,14.65,29.3)_2$ $(11.1,11.1,22.2)_1$ $(24.6,14.65,39.25)_3$	$(14.65,14.65,29.3)_2$ $(21.1,14.65,35.75)_3$ $(4.0,12.0,16.0)_3$ $(24.0,12.0,12.0)_2$ $(12.8,6.4,6.4)_2$
2	$(16.4,40.4,24.0)_3$ $(21.0,21.0,42.0)_2$	$(13.7,22.3,36.0)_3$ $(36.2,18.1,18.1)_1$
3	$(22.3,37.53,15.23)_3$ $(15.8,15.8,31.6)_1$	$(18.2,8.2,26.4)_1$ $(27.0,15.8,42.8)_3$ $(29.3,8.8,38.1)_3$ $(36.3,18.15,18.15)_1$ $(42.0,21.0,21.0)_1$
4	$(13.5,19.5,6.0)_3$ $(18.5,29.5,11.0)_3$ $(29.0,10.5,39.5)_3$ $(6.45,6.45,12.9)_2$	$(20.0,13.5,33.5)_3$ $(32.8,9.4,42.2)_3$ $(30.46,15.23,15.23)_2$
5	$(14.5,14.5,29.0)_2$	$(21.6,8.8,30.4)_3$ $(17.7,10.5,28.2)_3$

Table 4.3f The results of the cross-bispectrum between channels 6 and 7.

MINUTE #	COMBINATION $X_1X_1X_2$	$X_1X_2X_2$
1	$(24.6,36.9,12.3)_3$ $(14.0,14.0,28.0)_2$ $(36.2,5.9,42.1)_3$	$(31.6,10.5,42.1)_3$
2	$(7.03,35.15,28.12)_3$ $(14.06,33.81,18.75)_3$	$(4.0,12.5,16.5)_3$ $(16.8,8.4,8.4)_2$
3	$(6.45,26.95,20.5)_3$ $(7.0,33.4,26.4)_3$ $(6.45,42.75,36.3)_3$ $(16.5,23.5,7.0)_3$ $(19.5,19.5,39.0)_2$	$(28.0,14.0,14.0)_2$
4	$(19.3,19.3,38.6)_2$ $(13.5,5.8,19.3)_3$	$(22.9,19.3,42.2)_3$ $(14.6,4.69,29.29)_3$ $(32.8,16.4,16.4)_2$ $(42.0,21.0,21.0)_1$
5	$(14.5,20.8,6.3)_3$ $(12.8,12.8,25.6)_2$	$(35.0,17.5,17.5)$

Table 4.3g The results of the cross-bispectrum between channels 7 and 8.

subscripts which assign them to one of the three possible groups mentioned above.

For the cross-bispectrum between channel 1 and 2, triples of groups 1)and 2) appear quite frequently for the $X_1X_1X_2$ combination. Frequency components from the sources of channel 1 (with fundamentals at 5.9 cpm and ~15.0 cpm) keep on driving components on channel 2, e.g. during minute #1, the fundamental of the first source (5.9 cpm) together with the second harmonic of the second source (31.7 cpm) are phase-locked with a frequency component (25.8 cpm) from the second channel. At the same time, the second harmonic (29.6 cpm) of that second source (14.8 cpm) is also observed at the second channel. This phenomenon happens throughout the 5 minute duration. Additionally, some independent oscillators are coupled with their harmonics on the 2nd channel, e.g. during minute #2 (7.03, 7.03, 14.06), minute #3 (7.6, 7.6, 15.2), (11.1, 11.1, 22.2) and (4.1, 4.1, 8.2). There are also some occurences of frequency triples which belong to group 3), e.g. during minute #4 (18.75, 42.75, 24.0) and minute #5 (4.5, 24.5, 20.0). They should be expected from this kind of analysis because random components may appear in a physical environment. For the $X_1X_2X_2$ combination, most of the frequency triples are classified to group 3). This signifies that channel 2 do not have any great influence on channel 1 and no driving force from channel 2 is coupled with components in the signal of channel 1.

For the analysis between channel 2 and 3, the outcome is somewhat as expected because they both have a a common generator (with a fundamental at 19.0 cpm) as observed from the bispectral analysis of a single channel of signal. From Table 4.3, only triples belonging to groups 1) and 2) are observed for the $X_1X_1X_2$ combination. For the $X_1X_2X_2$ combination, triples belonging to group 1) also appear throughout the entire 5 minute period. Extra triples of group 2) in the $X_1X_1X_2$ combination mean that more influences have been made to channel 3 from channel 2 than the vice versa case.

For the cross-bispectrum between channels 3 and 4, results appear to be much more irregular. Most of the frequency triples belong to group 3), i.e. components are phase-locked in a random basis. They might be coupled together and driving each other. But this only happens in a very short time and on a local basis. The same result is found for the analysis of the relationship between channels 4 and 5, i.e. many triples belong to group 3). But there are a few more triples classified to either group 1) or 2). This is because channel 3 is taken from the duodenum while channels 4 and 5 are from the distal small intestine. This could mean that the duodenum and the distal small intestine are working independently of each other most of the time and no big influence is made on each other. Notice that in Table 4.3d, during minute #1 of the $X_1X_1X_2$ combination and also during minute #2 of the $X_1X_2X_2$ combination, there are a couple of frequency triples (marked 3) which demonstrate a slightly different phenomenon. They are (25.8, 15.5, 41.3) and (34.6, 7.6, 42.2) respectively. An element in each of the triples (e.g. 15.5 or 7.6 cpm) belongs to the fundamental of the generator of the signal itself. So these sources may sometimes interact with independent components and exert a driving force on the other channel. When it comes to channels 5 and 6, one can observe that many triples belong to either group 1) or 2). It also happens in both the combinations. This means that oscillators on both sites are trying to drive the activity at the other site indicated by the presence of their second harmonics at these sites.

Finally, for the cross-bispectrum between channels 6 and 7 as well as between 7 and 8, nothing really unusual occurs. Frequency triples belong to all the groups 1), 2) and 3). But the group 3) elements are in a majority situation. They are mixed with some group 1) and 2) elements so that some oscillators are trying to send their harmonics to the other site and hence make some influence on the activity there. But from the fact that there are a large number of group 3) elements, the activities on the colonic part of the digestive tract are usually acting all by themselves on an independent, local and short duration basis.

4.4 DISCUSSION

The purpose of the stationarity and the normality tests is to find out how a signal behaves when time passes by. Obviously they do not predict the exact characteristic of the signal under investigation. But they do provide a clue about how the signal is going to behave. Here, after the analysis of the signals, it is observed that the stationarity and normality start decreasing from the gastric and the duodenal part down to the colonic part of the digestive tract. This is because the main function of the stomach and the duodenum is to rapidly propel the contents down the tract. Hence a steady and organized electrical signal has to be present all the time. But when it comes to the distal small intestine, major mixing activities come into action. At the same time, food is not going to be transported so rapidly. A larger percentage of rejection of the null hypothesis thus results. In the colon, mainly mixing activities and only local movement occurs. This needs a lot of local independent oscillators with different frequencies. Hence the function of the signal there changes from time to time and the tests conclude that it is non-stationary and not normally distributed. All these reasonings are also proved by the bispectral analysis of both a single and double channels.

The purpose of bispectral analysis is to identify frequency components that are phase-locked. The components may or may not be harmonically related. However, if they are related harmonically, for sure they are generated from the same source. Hence the analysis can sometimes find out the characteristics of the generators of the signal under examination. From the bispectral analysis of a single channel of signal ,most of the generators located in various parts of the digestive tract are identified. In the stomach and the duodenum, the same sources are present throughout the entire period of investigation, e.g. two generators (5.0 and 15.0 cpm) in the stomach and a single source (19.0 cpm + harmonics) in the duodenum are observed. This confirms the theory that the main purpose of the stomach and the duodenum is to rapidly propel the contents and hence a stationary signal has to be present there all the time. When it comes to the distal small intestine, independent oscillators at different frequencies begin to function once in a while. In the jejunum, some of these oscillators still prevail for a relatively long period of time and hence propel the contents with a steady pace. But the oscillators only function for a very short time when it comes to the lower part of the small intestine. The presence of these oscillators prove the result of a non-perfect stationarity characteristic as proposed from the stationarity test being done on this part of the tract. Finally, in the colon, oscillators exist only for a very short time and on a local basis. These local oscillators generate random contractions in small regions of the wall and mix the contents without propelling them.

Finally, from the cross-bispectral method, a further identification of the characteristic of the digestive tract is obtained. The purpose of the analysis is to find out the existence of influences on the function of a local activity from another site. In the stomach and the duodenum, many activities are influenced by oscillators from the proximal site. For example, in the analysis between channels 1 and 2, group 1) and 2)'s frequency triples only appear on the $X_1X_1X_2$ combination, but not on the other. However this phenomenon is changed in the distal small intestine. Here oscillators are trying to exert driving forces in both directions. For instance, in the cross-bispectrum between channels 5 and 6, many oscillators that belong to either group 1) or 2) appear on both the $X_1X_1X_2$ and $X_1X_2X_2$ combinations. Food is still moving around but this time bidirectionally because these driving forces in both directions introduce contractions that are not in sequence. Finally, in the colon, most frequency triples belong to group 3). Little influence is felt at the nearby sites. These unsteady, locally based and short-lived oscillators prevent rings of contractions which move the contents in one direction or the other. They introduce random contractions which mix the food without transporting them. In concluding all these analyses, the functions of the digestive tract are as follows : the gastric and duodenal activities rapidly propel the food. In the distal small intestine, mixing activity starts coming into action. The contents are also moved bidirectionally as oscillators try to influence both the proximal and the distal sites. In the colon, little movement occurs but there are lots of mixing motions involved as local, short-lived and weakly coupled oscillators become the main components there.

CHAPTER 5

CONCLUSIONS

In the first part of this project, algorithms on stationarity and normality tests are designed. They give one an idea of how the time series under investigation may behave. But the length of the section of the series plays an important role on the use of these algorithms. Hence some experiments before the actual analysis are always recommended. But these tests are very straight forward and hence they should be used more often as an initial investigation.

Since the conventional power spectral analysis assume the signal under examination to be stationary, an alternate method, namely the bispectral analysis, is proposed for the analysis of signals which do not satisfy the assumption. It serves as an enhancement of the traditional analysis and provides more valuable information on the signal. It mainly computes the relationship between frequency components by looking at the phase-locking information between them. Sometimes when the components are harmonically related to each other, the generator of the signal can easily be identified by the method. Hence the function of the signal can be studied. But there are some draw backs in using this analysis. First of all, because of its complicated theory and little research has been done in this area so far, many of its statistical properties are still not known. Additionally, it consumes a large amount of computation time. But if the time is not an important factor and with the advancement of the computer technology, this should be a good method in search for the characteristics of both the explored and unexplored signals.

REFERENCES

{1} J. S. Bendat and A. G. Piersol, Random Data : Analysis and Measurement Procedures. Wiley-Interscience, 1971.

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

{3} G. Bodenstein and H. M. Praetorius, "Feature Extraction from the Electroencephalogram by Adaptive Segmentation", Proceedings of the IEEE, Vol.65, No.5, May 1977.

{4} D. R. Brillinger, "An Introduction to Polyspectra", Annuals of Mathematical Statistics, Vol.36, 1965, pp.1351-1374.

{5} M. D. Godfrey, "An Exploratory Study of the Bi-Spectrum of Economic Time Series", Applied Statistics, Vol.14, 1965, pp.48-69.

{6} G. Dumermuth and Th. Gasser, "The Computation of EEG Bispectra", Computer Programs in Biomedicine, Vol.8, 1978, pp.235-242.

{7} G. Dumermuth, P. J. Huber, B. Kleiner and Th. Gasser, "Analysis of the Interrelations Between Frequency Bands of the EEG by Means of the Bispectrum", Electroencephalography and Clinical Neurophysiology, Vol.31, 1971, pp.137-148.

{8} R. A. Fisher, Statistical Methods for Research Workers, Hafner Publishing Company, 1963.

{9} R. A. Haubrich, "Earth Noise, 5 to 500 Millicycles per Second", Journal of Geophysical Research, Vol.70, No.6, March 15, 1965, pp.1415-1427.

{10} K. Hasselmann, W. Munk, G. MacDonald, "Bispectra of Ocean Waves", Time Series Analysis,(Ed.)M.Rosenblatt., John Wiley and Sons, Inc., 1963, pp.125-139.

{11} P. Huber, B. Kleiner, T. Gasser and G. Dumermuth, "Statistical Methods for Investigating Phase Relations in Stationary Stochastic Process", IEEE Trans. on Audio and Electroacoustics, Vol.AU-19, No.1, March 1971, pp.78-86.

{12} J. F. Hung, "Program Listings for the Digital Processing of Non-Stationary Signals", Internal Report No.GMRG-4, McMaster Univ., Hamilton, Ont., Canada. June 1981.

{13} Jenkins, Spectral Analysis and Its Applications, Holden-Day, 1968.

{14} N. Kawabata, "A Nonstationary Analysis of the EEG", IEEE Trans. on Biomedical Engineering, Vol.BME-20, No.6, Nov. 1973, pp.444-452.

{15} B. Lindgren, Statistical Theory, MacMillan Publishing Co., Inc., 1968.

{16} Parzen, "Time Series Analysis for Models of Signal Plus White Noise - Cross-Bispectrum", Spectral Analysis of Time Series, (ED.) B. Harris, Wiley.

{17} L. J. Pinson and D. G. Childers, "Frequency-Wavenumber Spectrum Analysis of EEG Multielectrode Array Data", IEEE Trans. on Biomedical Engineering, Vol.BME-21, No.3, May 1974, pp.192-205.

{18} H. M. Praetorius, G. Bodenstein and O. D. Creutzfeldt, "Adaptive Segmentation of EEG Records : A New Approach to Automatic EEG Analysis", Electro-encephalography and Clinical Neurophysiology, Vol.42, 1977, pp.84-94.

{19} M. B. Priestley and T. S. Rao, "A Test for Non-stationarity of Time-Series", J.R.Statist.Soc.B, 31, pp.140-149.

{20} M. B. Priestley and H. Tong, "On the Analysis of of Bivariate Non-Stationary Process", J.R.Statist. Soc.B, 35, pp.153-166.

{21} S. K. Sarna, B. L. Bardakjian, W. E. Waterfall and J. F. Lind, "Human Colonic Electrical Control Activity (ECA)", Gastroenterology, Vol.78, 1980, pp.1526-1536.

{22} H. Sugimoto, N. Ishii, A. Iwata and N. Suzumura, "On the Stationarity and Normality of the EEG Data During Sleep Stages", Computer Programs in Biomedicine, Vol.8, 1978, pp.224-234.

{23} H. Sugimoto, N. Ishii, A. Iwata and N. Suzumura, "Stationarity and Normality test for Biomedical Data", Computer Programs in Biomedicine, Vol.7, 1977, pp.293-304.

{24} F. S. Swed and C. Eisenhart, "Tables for Testing Randomness of Grouping in a Sequence of Alternatives", Annuals of Mathematical Statistics, Vol.14, 1943, pp.66-87.

{25} D. Wang and A. H. Vagnucci, "TSAN : A Package for Time Series Analysis", Computer Programs in Biomedicine, Vol.11, 1980, pp.132-144.

{26} P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra : A Method Based on Time Averaging Over Short, Modified Periodograms", IEEE Trans. on Audio and Electroacoustics, Vol.AU-15, No.2, June 1967. pp.70-73.