

NON-LINEAR ANALYSIS OF SPIKE WAVEFORMS  
FOR CLASSIFICATION OF BENIGN ROLANDIC EPILEPSY OF CHILDHOOD

by

Edward J. Rzempoluck

B.A., Simon Fraser University

THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Psychology

© Edward J. Rzempoluck 1992

SIMON FRASER UNIVERSITY

December, 1992

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

APPROVAL

Name: Edward Joseph Rzempoluck

Degree: Master of Arts

Title of Thesis: Non-linear Analysis of Spike Waveforms  
for Classification of Benign Rolandic  
Epilepsy of Childhood

Examining Committee:

Chair: Dr. Bruce Alexander

---

Dr. Harold Weinberg  
Senior Supervisor

---

Dr. Barry Beyerstein

---

Dr. Douglas Cheyne

---

Dr. Tom Richardson  
External Examiner

Date Approved: 22 January 93



## ABSTRACT

Benign rolandic epilepsy of childhood (BREC) occurs in two forms. Children with the Typical form eventually fail to show symptoms, while those with the Atypical form continue to manifest symptoms into adult life. BREC is associated with characteristic interictal electroencephalographic spike waveforms. In order to develop a classification procedure that would distinguish between the two forms I analyzed 44 sets of averaged spike waveforms. Each set, an average of 6 to 20 spikes organized as 21 channels of 256 points, was available diagnosed Typical or Atypical on the basis of clinical and behavioural measures. From each set, using each of several criteria, I selected a single channel for the correlation dimension calculation. The correlation dimension is an estimate of the complexity of a multivariate dynamical system, based on measurements taken of a single variable. When the channel was selected on the basis of maximum spike amplitude, Atypical spike waveforms had a larger correlation dimension than Typical waveforms, suggesting that the neural generating system is more complex for Atypical than for Typical spikes. Correlation dimension thus appears to be a differential diagnostic measure for BREC.

Key Words: time-series analysis, chaos, epilepsy, childhood.

Let cloud-shapes swarm

Let chaos storm

I wait for form

... Robert Frost

## ACKNOWLEDGEMENTS

The work in this thesis has benefitted substantially from discussions with my senior supervisor, Dr. Hal Weinberg, and from suggestions made by Dr. Doug Cheyne. I want to thank also Diane Crisp for making available source material on BREC. The name QuickBasic is a registered trademark of the Microsoft Corporation.

## TABLE OF CONTENTS

LIST OF TABLES . . . . .		viii
LIST OF FIGURES . . . . .		ix
1 INTRODUCTION . . . . .		1
1.1 BENIGN ROLANDIC EPILEPSY OF CHILDHOOD . . . . .		1
1.2 THE CORRELATION DIMENSION . . . . .		6
1.2.1 APPLICATIONS TO EEG ANALYSIS . . . . .		6
1.2.2 CORRELATION DIMENSION, CHAOS AND DETER-		
MINISM . . . . .		8
1.2.3 CALCULATING CORRELATION DIMENSION . . . . .		12
1.2.4 PARAMETERS . . . . .		17
1.3 ATTRACTOR GEOMETRY AND UNDERLYING DYNAMICS . . . . .		26
1.4 ALGORITHMS . . . . .		28
2 EXPERIMENT 1: Evaluation of the CORDIM Algorithm . . . . .		32
2.1 PSEUDO-RANDOM NOISE . . . . .		33
2.2 RANDOM NOISE . . . . .		37
2.3 THE 2-TORUS . . . . .		42
2.4 THE 3-TORUS . . . . .		46
2.5 SPIKE WAVEFORM SIMULATION USING 2 VARIABLES . . . . .		50
2.6 SPIKE WAVEFORM SIMULATION USING 3 VARIABLES . . . . .		54
3 EXPERIMENT 2: Multi-dimensional Reconstruction . . . . .		59
3.1 APPLICATION TO 2-TORUS DATA . . . . .		59
3.2 APPLICATION TO 3-TORUS DATA . . . . .		64
4 ANALYSIS OF BREC SPIKE WAVEFORMS . . . . .		68
4.1 INTRODUCTION . . . . .		68

4.2	EXPERIMENT 3.1: Single Common Channel . . .	73
4.2.1	METHOD . . . . .	73
4.2.2	RESULTS . . . . .	76
4.2.3	DISCUSSION . . . . .	79
4.3	EXPERIMENT 3.2: Maximum Amplitude Channel .	80
4.3.1	METHOD . . . . .	80
4.3.2	RESULTS . . . . .	81
4.3.3	DISCUSSION . . . . .	88
4.4	EXPERIMENT 3.3 Maximum Peak/RMS Ratio Chan- nel . . . . .	91
4.4.1	INTRODUCTION . . . . .	91
4.4.2	METHOD . . . . .	92
4.4.3	RESULTS . . . . .	92
4.4.4	DISCUSSION . . . . .	94
4.5	EXPERIMENT 3.4: Multichannel Reconstruction	95
4.5.1	INTRODUCTION . . . . .	95
4.5.2	METHOD . . . . .	98
4.5.3	RESULTS . . . . .	99
4.5.4	DISCUSSION . . . . .	101
4.6	EXPERIMENT 3.5 - Multichannel Reconstruction	102
4.6.1	METHOD . . . . .	102
4.6.2	RESULTS . . . . .	102
4.6.3	DISCUSSION . . . . .	104
4.7	GENERAL DISCUSSION . . . . .	105
	APPENDIX I - Algorithms . . . . .	113
	BIBLIOGRAPHY . . . . .	117

LIST OF TABLES

1	- Data Summary, Exp. 3.1 - Channel T4. . . . .	78
2	- Statistics, Exp. 3.1 - Channel T4. . . . .	79
3	- Data Summary, Exp. 3.2 - Maximum Amplitude Channel	81
4	- Statistics, Exp. 3.2 - Maximum Amplitude Channel .	82
5	- Hierarchical Cluster Analysis - Exp. 3.2 . . . . .	86
6a	- K-Means Cluster Analysis; 2 Groups . . . . .	87
6b	- Mean Squares; 2 Groups . . . . .	87
6c	- Classification by Cluster; 2 Groups. . . . .	87
7a	- K-Means Cluster Analysis; 3 Groups . . . . .	87
7b	- Mean Squares; 3 Groups . . . . .	87
7c	- Classification by Cluster; 3 Groups. . . . .	87
8	- Data Summary, Exp. 3.3 - Max. Peak/RMS Ratio Channel	93
9	- Statistics, Exp. 3.3 - Max. Peak/RMS Ratio Channel	94
10	- Data Summary, Exp.3.4 - Multichannel Reconstruction	100
11	- Statistics, Exp. 3.4 - Multichannel Reconstruction	101
12	- Data Summary, Exp.3.5 - Multichannel Reconstruction	103
13	- Statistics, Exp. 3.5 - Multichannel Reconstruction	104

LIST OF FIGURES

Figure 1 - Time-series generated using pseudo-random number generator function (BASIC RND function); length is 1024 data points . . . . . 34

Figure 2 - Phase-space representation generated by plotting pseudo-random time-series against itself using a lag of 1. . . . . 35

Figure 3 - Correlation dimension as a function of embedding dimensions 1 to 5, for the pseudo-random time-series of Figure 1. No saturation is evident. . . 36

Figure 4 - Random time-series, comprised of thermal and 1/f noise components; 1024 data points. . . . . 39

Figure 5 - Phase-space attractor constructed by plotting random time-series against itself with a lag equal to 1. 40

Figure 6 - Correlation dimension as a function of embedding dimensions 1 to 5, for random time-series of Figure 4. No saturation is evident . . . . . 41

Figure 7 - Time-series generated by the equations defining the 2-torus, the summation of 2 sine-waves of incommensurate frequencies. . . . . 43

Figure 8 - Phase-space attractor constructed by plotting the two-torus time-series against itself using a lag equal to 5. . . . . 44

Figure 9 - Correlation dimension vs. embedding dimensions 1 to 5 for 2-torus time-series. Saturation is clearly evident . . . . . 45

Figure 10 - Time-series generated using equations defining the 3-torus, the summation of 3 sine-waves of incommensurate frequencies. . . . . 47

Figure 11 - Phase-space attractor constructed by plotting the three-torus time-series plotted itself using a lag equal to 5. . . . . 48

Figure 12 - Correlation dimension vs. embedding dimensions 1 to 5 for the 3-torus time-series. . . . . 49

Figure 13 - An approximation to the BREC spike waveform generated using 2 equations in 2 variables . . . 51

Figure 14 - Phase-space attractor generated by plotting the two variable spike waveform approximation against itself using a lag equal to 10. . . . .	52
Figure 15 - Correlation dimension vs. embedding dimensions 1 to 5 for the 2-variable spike approximation time-series . . . . .	53
Figure 16 - An approximation to the spike waveform generated using 3 equations in 3 variables. . . . .	56
Figure 17 - Phase-space attractor constructed by plotting the 3 variable spike approximation against itself using a lag equal to 10. . . . .	57
Figure 18 - Correlation dimension vs. embedding dimensions 1 to 5, for the 3-variable spike approximation. . . . .	58
Figure 19 - The 2 torus time-series used for the multichannel reconstruction method. . . . .	61
Figure 20 - Two-torus attractor used with the multichannel reconstruction method . . . . .	62
Figure 21 - Correlation dimension vs. embedding dimensions 1 to 5 for multichannel reconstruction	

method using the 2-torus data . . . . .	63
Figure 22 - The 3 torus time-series used with the multichannel reconstruction method. . . . .	65
Figure 23 - Phase-space attractor for the 3-torus time- series used with the multi-channel reconstruction . . .	66
Figure 24 - Correlation dimension vs. embedding dimen- sions 1 to 5, for the multi-channel reconstruction method using the 3-torus data . . . . .	67
Figure 25 - Example of a Typical spike waveform; 256 data-points sampled at 200 points per second, for a time- span of 1.28 seconds (patient no. 1, channel T4). . . . .	71
Figure 26 - Example of an Atypical spike waveform; 256 data-points sampled at 200 points per second, for a time- span of 1.28 seconds (patient no. 12, channel T3) . . .	72
Figure 27 - Phase-space attractor corresponding to the spike waveform for patient no. 1, channel T4. A lag of 8 was used for this construction . . . . .	77
Figure 28 - Cases are sorted by correlation dimension, and labelled by clinical diagnosis of BREC type . . . . .	84

# 1 INTRODUCTION

## 1.1 BENIGN ROLANDIC EPILEPSY OF CHILDHOOD

Childhood epilepsy may occur in the form of either generalized or focal epilepsy. Generalized epilepsy is characterized by diffuse seizures and an abnormal EEG pattern distributed over the entire surface of the brain. Focal, or partial, epilepsy is marked by localized seizures with EEG abnormalities that are similarly localized to an area of the surface of the brain. The partial form of childhood epilepsy is often found to be associated with focal seizures involving the facial muscles, an EEG pattern localized over the rolandic fissure, and a benign prognosis. This form of epilepsy has been termed benign rolandic epilepsy of childhood (BREC) (eg. Gregory and Wong, 1984).

Evidence has been found for a genetic component to the etiology of BREC. Blom, Heijbel, and Bergfors (1972) studied 40 children, and found that 18% had siblings or parents with seizures. Heijbel, Blom, and Rasmuson (1975) studied 19 children with rolandic spikes, and no other neurological symptoms. They found that 47% had parents or siblings who had seizures in childhood. They conclude that the familial distribution pattern best supports a model of an autosomal dominant gene with age-dependent penetrance.

The prevalence of BREC has been found to be about 7 times greater than the rate for petit-mal seizures, in Uppsala,

Sweden (Blom et al., 1972). For about 15% of all children below age 15 with seizures, the syndrome appeared to be of the BREC type. Heijbel, Blom and Bergfors (1975) report that BREC accounts for 16% of childhood epilepsy, and is 4 times as prevalent as petit-mal seizures in the normal population. BREC has also been found to be more prevalent in males (O'Donohoe, 1985).

BREC is an idiopathic form of epilepsy, not associated with any known organic cause. It is defined by a cluster of symptoms which have been identified by several groups. Blom et al. (1972) suggested that BREC is marked by early onset, between 5 and 10 years of age, a hereditary predisposition, brief usually nocturnal clonic seizures, interictal spike and wave EEG with a centro-temporal focus, seizures which are easily controlled with drug treatment, little evidence of neurological abnormalities, and remission of seizures and abnormal EEG's after puberty. Beaussart (1972) studied 221 cases, and added absence of focal lesions, and interictal EEG with frequent spike-waves localized over the rolandic area, unilateral in 70% of patients but possibly changing sides between recordings.

Dreifuss (1983) describes the rolandic EEG spike-wave discharge associated with BREC as consisting of high amplitude bi-polar spikes followed by a single slow wave. This activity is localized and synchronous in the centro-temporal area. Distribution is unilateral in two-thirds of patients. In the

remaining one-third, spikes are bilateral and asynchronous in amplitude and frequency of discharge. Discharges are frequent and occur in clusters. Discharge characteristics (for example amplitude) are uncorrelated with seizure characteristics.

BREC has been found to occur in two forms (Gregory and Wong, 1984; Wong, Gregory, and Farrell, 1985). Aicardi and Chevrie (1982) studied 32 children with symptoms consistent with a diagnosis of BREC. They found however that a subgroup of 7 children presented additional symptoms, including absences, atonal or myoclonic fits, and abnormal EEG. For these 7 patients, common features were early onset, but not before age 2, and at least 2 seizure types: partial motor seizures, absences similar clinically to petit-mal epilepsy, and brief atonic or myoclonic seizures. Sleep EEG was marked by an asymmetrical (higher on right side) 3 Hz spike and wave pattern during periods of frequent seizure activity. Aicardi and Chevrie (1982) termed this syndrome Atypical BREC. Follow-up studies showed that 5 of these 7 children had remission of symptoms, while 2 showed an improving trend.

The remainder group of 25 children showed no abnormalities other than the anomalous EEG. Follow-up studies showed that all had complete remission of symptoms after age 12. Aicardi's (1988) definition of this BREC syndrome includes absence of cognitive deficits, neurological findings and brain damage, onset after age 2, infrequent seizures of simple partial (facial) type, and good response to therapy.

The existence of the Atypical form of BREC is supported also by a study comparing 14 cases with the benign pattern and 8 with the atypical pattern of development (Beydoun, Garofalo, and Drury, 1989). These authors suggest that these two groups cannot be differentiated on the basis of their EEG findings or initial clinical evaluations.

Wong (1989) referred to these two forms of BREC as Typical and Atypical, defining the Typical form by symptoms including early onset (4 to 12 years), nocturnal partial seizures usually involving the face, absence of neurological findings and brain lesions, and a centrotemporal EEG distribution. The Atypical form is defined by presence of these same symptoms and including the presence of neurological findings and cognitive deficits.

Wong (1989) found evidence for the existence of the two forms of BREC using single dipole source localization analysis of the EEG spike-wave waveform. For Atypical BREC, stable estimates were obtained using the spike portion of the waveform, while corresponding source locations were relatively widely distributed. For Typical BREC, stable estimates were obtained using the wave portion of the waveform, while corresponding source locations were relatively less widely distributed, localized to the right or left centrotemporal areas.

These findings might be interpreted as suggesting that for the Atypical case, the waveform was relatively phase-coherent over only the short interval of initial spike

portion, and relatively phase-incoherent over the subsequent longer interval of the wave portion. The relative incoherence of the Atypical case might then be associated with a relatively complex network of generator sites, an inference which is also supported by the wider distribution of source locations. In contrast, the relative coherence of the Typical waveform suggests a correspondingly simpler network of generator sites, again a view supported by the relatively tighter clustering of generator sites for the Typical cases. Wong (1989) proposes such a model, with a common generator site for both the Typical and Atypical cases. For the Atypical case however, additional generator sites become involved after the initial spike has occurred. Cheyne (1993, private communication) has suggested that the relative distribution of source locations may be influenced by the fact that only a single dipole was used in the analysis. These results can thus not address the question of whether the generator systems consist of generator sites distributed in location or in time.

Early diagnosis of BREC type would allow early intervention for Atypical BREC, and for Typical BREC might prevent the emotional and behavioural disturbances resulting from the social attitudes which often accompany a diagnosis of epilepsy (Geladze, Toidze, and Lomashvili, 1983; Verity, 1988).

It is the aim of this study to develop a continuous diagnostic variable that can be applied in the classification of BREC. This variable is the correlation dimension. The

contribution to previous work in this area would be that this particular diagnostic variable can be easily calculated using a computer algorithm.

## 1.2 THE CORRELATION DIMENSION

The correlation dimension is a lower bound on the number of variables involved in the dynamical behaviour, or evolution over time, of a multivariate system. It has been shown that under suitable conditions, an estimate of the correlation dimension can be computed from a univariate time-series, when that time-series is an adequate sample taken from the multivariate process (Takens, 1981; Farmer, Ott, and Yorke, 1983). The requirements for a sampling to be considered adequate will be discussed in section 1.2.3. The actual calculation of the correlation dimension has been reduced to a relatively straight-forward algorithm (Grassberger and Procaccia, 1983 a, b). For these reasons, the correlation dimension has been applied as a measure of system complexity, in a wide range of fields of endeavour. Relevant to the present work, correlation dimension analysis of EEG recordings has been used to explore brain system dynamics (eg. Rapp, Bashore, Matinerie, Albano, Zimmerman, and Mees, 1989).

### 1.2.1 APPLICATIONS TO EEG ANALYSIS

Correlation dimension has been applied to EEG recordings made in a no-task condition (Pritchard and Duke, 1992), during

a mental arithmetic task (Nan and Jinghua, 1988; Dvorak and Siska, 1986), alpha rhythm production (Soong and Stuart, 1989; Basar, Basar-Eroglu, Roschke, and Schult, 1990), the sleep cycle (Babloyantz, Salazar, and Nicolis, 1985), an epileptic seizure event (Babloyantz and Destexhe, 1986), stages of anesthesia in a medical operation (Watt and Hameroff, 1987), and during ethanol consumption (Palus, Dvorak, and David, 1992). Animal studies include a demonstration of chaotic dynamics in the EEG recorded from the olfactory bulb of a rabbit (Skarda and Freeman, 1987), and from the limbic cortex of a rat during rest, locomotion, and kindled epileptic seizure (Pijn, Van Neerven, Noest, and Lopes da Silva, 1991).

Rapp et al., (1989) survey of the use of the correlation dimension in the analysis of EEG recordings by different groups, and present as the rationale for using the correlation dimension, that the correlation dimension, compared with other simple statistics, uses more of the information present in a time-series such as the EEG. The correlation dimension, they suggest, is therefore a more robust characterization of the behaviour of such a system.

Basar et al. (1990) showed that alpha EEG has a deterministic, task-related component. Subjects were asked to attend to a missing stimulus in a train of regularly presented stimuli. It was found that alpha EEG produced in the 0.5 seconds prior to the missing stimulus was phase coherent between separate missing stimulus events. These EEG segments

were correlated to the extent that the subjects were able to mentally track the missing stimulus.

Gallez and Babloyantz (1991) compared several methods of analysis of EEG recorded during eyes closed, stage 4 sleep, and Creutzfeld-Jakob coma. These methods included Lyapunov exponents, Kolmogorov entropy, (measuring the rate at which new information is produced, or the mean time for which a signal can be predicted), and attractor dimensionality.

It was found that similar values of dimensionality were computed by the correlation dimension, and by two estimates of dimensionality based on Lyapunov exponents. The authors note that typically there is a great deal of variance in the dimensionality values even using the same estimator with different data samples, and suggest that experimental situations should be arranged to provide a clear distinction between the types of tasks that are used, and thus between the associated neural activities. They emphasize that dimensionality estimates are most effective when used to distinguish between the effects of different types of task requirements, rather than when used as indicators of absolute complexity of neural dynamics.

### 1.2.2 CORRELATION DIMENSION, CHAOS AND DETERMINISM

It might be useful here to state some definitions. First, a dynamical system is any physical system one chooses to identify and demarcate. Examples are a pendulum, an ensemble of neurons, the solar system. Such a dynamical

system is characterized by a pattern of behaviour over time. This behaviour may be relatively simple, as in a pendulum at rest, or swinging freely. One form of behaviour of a dynamical system, termed chaotic, has been found to be particularly complex. One example is a billiard table with several ball on it and in motion. The path of any one ball is determined by well-defined physical principles, and yet is not predictable for more than a short time following some given initial set of positions and velocities of the balls. Moreover, the ongoing behaviour of this system is sensitively dependent on these initial conditions. Small changes to the starting positions or velocities will after a short time cause the behaviour of the system to diverge from the behaviour of the undisturbed system. Such sensitivity to initial conditions is a defining characteristic of chaotic behaviour.

Chatterjee and Yilmaz (1992) offer an accessible review of chaos, its background, relation to statistics, and areas of application. They discuss the relationship between different estimators of dimensionality, noting that at present these estimators are necessary but not sufficient conditions for a chaotic dynamical process to be considered deterministic. The authors suggest that sufficiency conditions for labelling a chaotic process as deterministic have yet to be discovered.

An indicator that has been suggested as a necessary condition for a system to be considered chaotic is the Lyapunov exponent, a quantitative measure of sensitivity to

changes in initial conditions. Moon (1987) refers to Lyapunov exponents as being diagnostic of chaos. Jackson (1990) notes that there is wide agreement that a positive Lyapunov exponent must be present before a motion can be considered to have a chaotic dynamic. Farmer et al. (1983) provide a review of several different measures of attractor dimensionality, including the correlation dimension, and their relationship to Lyapunov exponents.

Together, a fractional value of dimensionality, together with a positive Lyapunov exponent would be compelling evidence for the presence of chaos. In the present work, only a measure of dimensionality is considered. This is insufficient to permit a dynamical system to be classified as chaotic, and indeed no such claim will be made. The correlation dimension is being used here only as an estimate of dynamical system complexity.

Chatterjee and Yilmaz (1992) note also that the discovery of chaos is evidence for the position that the question of whether an aperiodic process is deterministic or probabilistic may be undecidable. This question, they say, may join other such undecidable questions, such as the Heisenberg uncertainty principle and Godel's incompleteness theorem. Godel's theorem established that any system that is sufficiently complex, for example, arithmetic, can generate statements which then cannot be proved correct or incorrect from within that system.

Ford (1987) points to Godel's fundamental theorem as the

reason why chaotic dynamics are in principle unpredictable. Such dynamics, though resulting from possibly simple deterministic rules, nevertheless contain more information than can be encompassed by the logical system that is being used to try and predict the dynamical system's behaviour. In essence, a part is trying to know the whole.

Hobbs (1991) offers an argument that attempts to reconcile what he considers to be the unnecessary debate between chaos and determinism. His thesis is in two parts. First, chaotic processes are pervasive. All that is required for a system to exhibit chaotic behaviour is that the system be energy dissipating, and that it contain a non-linearity. These characteristics are seldom absent from natural systems. Second, the exponential sensitivity to initial conditions that is a characteristic feature of chaotic systems may allow these systems to amplify quantum fluctuations to a level where these can have an influence on macroscopic phenomena.

Thus, Hobbs says, quantum level indeterminism becomes indeterminism at the macroscopic level. Determinism may in fact be, rather than a reality, no more than an illusion resulting from the particular scale at which phenomena are commonly observed. Determinism, in the face of a chaotic universe, in fact requires an infinite level of precision in knowing some initial state. Quantum uncertainty does not allow for such infinite precision. A butterfly's wings distant flapping may indeed have an influence on local

weather, but we could never know what the local weather would have been otherwise. Determinism may be less a principled, and more a human, psychological requirement.

### 1.2.3 CALCULATING CORRELATION DIMENSION

The algorithm used for calculating correlation dimension was proposed by Grassberger and Procaccia (1983a, b). Calculation of correlation dimension begins by using the data points of the original time series to generate a group of points, more precisely a set of vectors, in an  $m$ -dimensional variable space, usually termed a phase-space. The motivation for this is a theoretical result proved by Takens (1981). A single observable variable  $y(t)$  is assumed to be an adequate sampling of a multivariate process defined by the vector  $X(t)$ . Typically, some components of  $X(t)$ , that is, some of the variables of the multivariate process, may be difficult or impossible to observe. The stipulation of adequate sampling requires that the variables of the multivariate process be sufficiently strongly coupled to the measured variable (Frank, Lookman, Nerenberg, Essex, Lemieux, and Blume, 1990). If this is the case, then Takens' (1981) result shows that multiple observations of a single variable  $y(t)$  related to the multivariate process, can be considered as a single observation of the multivariate vector  $X(t)$ . In more detail, the multiple observations of  $y(t)$  are used to form a vector  $u(i)$  as follows:

$$u(i) = \{ y(i), y(i+\tau), \dots, y(i+(m-1)\tau) \}$$

where  $\tau$  is the lag parameter. These vectors  $u(i)$  are then treated as points in an  $m$  dimensional variable space, usually termed a phase space. Geometrically, the points corresponding to the vectors  $u(i)$  define an object, an attractor, in the phase space. The original time-series is said to be embedded in the phase-space.

Takens' (1981) result shows that certain dynamical properties of the original multivariate process  $X(t)$  are preserved in the reconstruction. A particular property of interest that is preserved is the dimensionality of the original multivariate process. This dimensionality is reflected in the dimensionality of the attractor. A requirement is that embedding phase-space dimension  $m \geq 2D + 1$ , where  $D$  is the dimensionality of the attractor.

In the next step of the process, the dimensionality of the attractor is calculated. For this calculation, Grassberger and Procaccia (1983a, b) invoked the notion of the correlations between points on the attractor. These correlations are a function of the distance  $r$  around any one point on the attractor, and are measured by the correlation integral,  $C(r)$ . The correlation integral is the proportion of the points on the attractor that are found within a distance  $r$  of a reference point on the attractor, averaged over all points on the attractor as reference points. The correlation

integral is thus the average proportion of points on the attractor found within a volume of length  $r$  around all points on the attractor.

Intuitively, at small values of  $r$ , that is, in a small volume of phase-space around any point on the attractor, there will be found on average a small fraction of the total points on the attractor. As the distance  $r$  increases, this average fraction will increase. At sufficiently large values of  $r$ , all points on the attractor will be found.

Grassberger and Procaccia (1983a, b) pointed out that the  $C(r)$  and  $r$  are related by the equation

$$C(r) = r^d \quad \dots \text{eqn 1}$$

where  $d$  is the correlation dimension. Thus,  $C(r)$  is a power function of  $r$ , in the limit of values of  $r$  that are small with respect to the size of the attractor. As an example, if  $d = 2$ , then  $C(r)$  increases as the square of  $r$ ; the average proportion of points on the attractor that are found in the vicinity of any reference point on the attractor increases as the square of the distance around the reference point. This behaviour is intuitively consistent with more usual notions of dimension: for a 2-dimensional surface, the area size of the surface increases as the square of length along the surface.

In the case of an attractor, we cannot speak of size as

such, because the attractor is composed of discontinuous points. We therefore rather measure an analogue of size, by computing the correlation integral,  $C(r)$ . The discontinuous nature of the attractor also leads to the notion of fractional values of dimensionality, in contrast with the more usual integer values of dimension assigned to physical objects which are geometrically continuous. Typically, attractors that are reconstructed from time-series will have non-integral values of (correlation) dimension. Such objects are termed strange attractors.

Solving equation 1 for  $d$ , we obtain

$$d = \frac{\ln C(r)}{\ln r} \quad \dots \text{eqn 2}$$

Correlation dimension can thus be calculated by finding the slope of the graph of  $\ln C(r)$  vs.  $\ln r$ . Typically, this graph has an ogive or S shape. In practice therefore, a regression line is calculated that best fits some intermediate portion of the graph. The slope of this regression line is then taken as the correlation dimension.

The correlation integral is defined by the equation

$$C(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1}^N H \{ r - |X(i) - X(j)| \} \quad \dots \text{eqn 3}$$

where  $N$  is the number of data points of the time series that

are used for the calculation;  $r$  is the scale length;  $X(i)$  and  $X(j)$  are vectors defining points on the attractor.  $H$  is the Heaviside function defined by

$$H(x) = 0, x \leq 0 \quad \dots \text{eqn 4}$$

$$H(x) = 1, x > 0$$

The Heaviside function thus simply counts the number of pairs of points  $X(i)$  and  $X(j)$  that separated by a distance less than  $r$ . The distance measure indicated in the equation is the usual Euclidean distance between 2 points. The computation of this distance is time-consuming. Alternative distance measures have been proposed that are faster to compute, and are equivalent for the purposes of calculating the correlation integral (Moon, 1987). The particular method adopted in the CORDIM algorithm involves summing the absolute values of the differences between corresponding components of the two vectors in question. This measure is termed the city-block measure: it is the distance between 2 points that one must travel following the constraints of streets organized around blocks. Using different measures of distance has the effect of changing the absolute range of the value of  $r$  in the graph of  $C(r)$  vs  $r$ . The slope of the graph is however unchanged.

#### 1.2.4 PARAMETERS

The calculation of correlation dimension involves the selection of a number of parameters, whose selection can lead to ambiguities in the final result. These parameters have been studied extensively in the last several years. Some of these factors are discussed by Dvorak and Siska (1986) in the context of EEG analysis. The following is a review of several of these parameters that have been shown to be particularly significant in the correlation dimension calculation.

##### LAG:

One major ambiguity in the calculation is the estimation of the value of the lag, used in forming the vectors which define the phase-space attractor. The original Takens formulation states that the value of lag used is not critical, for a time-series that is both sufficiently long, and noise-free. In the practical case of a limited number of potentially noisy data points, the value of lag that is chosen for reconstructing the attractor in phase space is critical. The value of correlation dimension calculated for an attractor is not independent of the value of lag chosen for the reconstruction. Examining an attractor visually, in the case of unlimited noise-free data, adjacent orbits on the attractor would be discriminable regardless of the value of lag.

In contrast, in the case of limited, noisy data, the extent to which such orbits are distinguishable is a function

of lag. At non-optimal values of lag, noise in the form of quantizing error and low-amplitude signals unrelated to the waveform of interest, will result in some orbits overlapping, and therefore being indistinguishable from, other orbits. The attractor will be compressed along certain axes in the phase-space. At an optimal value of lag, the attractor will be most homogeneously distributed in the phase-space; orbits on the attractor will be maximally mutually discriminable. This observation is the basis for the LAGFIND algorithm described in section 1.4. At this optimum value of lag, the components of the resulting vectors will be maximally independent. A set of maximally independent basis vectors will have been found for the embedding phase-space.

Numerous approaches have been attempted to try to deal with the lag question. It has been suggested that independent components for the embedding phase-space are obtained when lag is chosen to be between 10% and 30% of the periodicity of interest in the time-series (Schaffer, Truty, and Fulmer, 1988).

Broomhead and King (1986) have proposed an alternative to the Grassberger-Procaccia algorithm for the estimation of dimension, using a singular value decomposition method. This method attempts to reduce the effect of noise in the original time-series on the calculation of correlation dimension, as well as circumventing the problem of choosing a value of lag. Their study demonstrated that this method provides an increase

in the length of the scaling region in the plot of  $\ln C(r)$  vs  $\ln r$ , allowing more stable estimates of correlation dimension to be made.

This method begins by using the method of lags to construct a sequence of  $n$ -dimensional vectors from the original time-series. The set of these vectors constitutes a matrix  $X$ , from which an  $n \times n$  covariance matrix  $X^T X$ , may be formed. This covariance matrix may then be decomposed, giving a set of eigenvectors which form an orthonormal basis for an embedding space. By using only the  $m < n$  most significant eigenvectors to define the phase-space, an attractor is created which presumably contains less noise than if the attractor were constructed using all  $n$  vectors to define the phase-space. The dimensionality of the attractor is taken as being equal to the number of significant eigenvalues generated by the decomposition. According to Destexhe et al. (1988) however this method also suffers from the sensitivity to an adequate choice of the lag parameter as does the method of lags itself. It has been shown recently however that the number of significant eigenvalues is unrelated to the dimensionality of the attractor (Gibson, Farmer, Casdagli, and Eubank, 1992), but that the decomposition is nevertheless a useful procedure.

Albano, Muench, and Schwartz (1988) combined the singular value decomposition method of Broomhead and King (1986) with the Grassberger-Procaccia algorithm. As before, singular

value decomposition of the matrix of vectors formed using the method of lags is used to define an appropriate subspace using only the most significant eigenvectors. The resulting attractor is then subjected to analysis using the Grassberger-Procaccia algorithm. This method gives a longer scaling region than does the Grassberger-Procaccia algorithm used alone, allowing more stable estimates to be made of the dimension value. The selection of an appropriate value of lag is still a potential problem with this method.

A commonly used approach to the problem of finding a value of lag which will produce orthogonal coordinates for the embedding phase-space involves calculating the autocorrelation function for the time series, and then determining the first minimum of this function. The autocorrelation function is a measure of the linear dependence of two variables. A similar approach is the calculation of the mutual independence function and its first minimum (eg. Fraser and Swinney, 1986). Mutual information is a measure of the general dependence of two variables. Fraser and Swinney (1986) argue in favour of mutual information over autocorrelation, claiming that mutual information will allow phase-space coordinates to be found that are generally, and not just linearly, independent.

More recent work however has shown that neither autocorrelation nor mutual information is invariably successful in determining an optimal value of lag (Martinerie, Albano, Mees, and Rapp, 1992). Interestingly, Martinerie et al. (1992)

found that the first minimum of the autocorrelation function was the most successful of several measures involving both the autocorrelation and the mutual information functions. The time-series used for their study were the Rossler and Lorenz 3-dimensional systems of differential equations, as well as a 3-torus, a non-linear oscillator driven at 3 incommensurate frequencies (see section 2.4).

#### EMBEDDING DIMENSION:

A second potential source of ambiguity in the calculation of the correlation dimension is the maximum value of embedding dimension,  $m$ . Schaffer et al. (1988) suggest that  $m$  should be greater than  $2n+1$ , where  $n$  is the dimensionality of the attractor. Without a-priori knowledge of the value of this dimensionality,  $m$  can be initially set at an arbitrarily high value (in practice, 10 to 15), and the asymptotic value of correlation dimension that results, an estimate of  $n$ , can then be used to select a more appropriate value for  $m$ . Pilot work with the BREC data files indicated an approximate value for  $n$  of approximately 2. This value is consistent with Babloyantz and Destexhe's (1986) finding of a correlation dimension of 2 at an embedding dimension of 5, for a time-series recording 10 seconds of a petit-mal seizure.

With a limited number of data-points, it has been observed that an overly high value for embedding dimension may result in an unstable estimate of correlation dimension, as a

result of a scaling region that is too short to permit a reliable estimate of the slope to be determined (Mayer-Kress and Layne, 1987). According to Schaffer et al.'s (1988) guidelines, with an estimated correlation dimension value of 2, embedding dimension should be set at 5. However, in order to maximize the length of the scaling region, and because only relative rather than absolute values of correlation dimension are considered, the value of embedding dimension used in the present study was set at 4.

#### NUMBER OF DATA POINTS:

Ruelle (1989) argues for a minimum time-series length on the order of  $a^{D_c/2}$ , where  $a$  is proportional to the length of the scaling region:  $a$  is equal to the ratio of maximum to minimum  $\ln r$  values over which the graph of  $\ln C(r)$  vs.  $\ln r$  is approximately linear.  $D_c$  is an initial estimate of the correlation dimension of the time-series. Ruelle suggests a value of  $a = 10$ . For attractors of high dimensionality and with a requirement for a long scaling region, this criterion quickly leads to a need for an in practice unrealizable number of data points. In the present study,  $D_c$  was estimated on the basis of pilot studies to be on the order of 2. These values give the requirement of a minimum of  $10^1$ , or about 10 data points.

Ruelle's (1989) suggestion has been criticized. Essex and Nerenberg (1991) claim that Ruelle's bound applies to cor-

relation dimension but not to the estimates made by computing the slope of the regression line to the intermediate segments of the  $\ln C(r)$  vs  $\ln r$  curve. Essex and Nerenberg (1990, 1991) suggest using a more modest scaling region requirement (eg.  $a = 2$ ), and a more comprehensive calculation taking into account both the statistical and the geometric problems involved. Their calculation yields a value approximately 20 times larger than Ruelle's requirement, with a scaling region length of  $a = 2$ . This translates into a requirement for 200 data points in the present study. The actual number of data-points used, 256, is of the same order as this requirement. Smith (1992) suggests that if accuracy requirements for the correlation dimension estimate are not too strict then modest sample sizes are not unreasonable. For a root-mean-square error of 1, and an estimated dimension of 5, Smith (1992) suggests a sample size of 30. For an RMS error of 0.1, sample size increases dramatically to about 5000. Thus one effect of a small sample size is a larger RMS error, that is a value of correlation dimension that is less stable with respect to sampling.

DeCoster and Mitchell (1991) investigated the behaviour of the Grassberger-Procaccia algorithm for several data sets generated by sets of equations. They found that in some cases as few as 100 data points were sufficient to allow saturation of correlation dimension to be detected, indicating the presence of deterministic dynamics. In other cases even 5000

data points were not sufficient to generate convergence. This latter result would normally indicate that the time-series in question represent a stochastic rather than a deterministic process. DeCoster and Mitchell (1991) conclude that the time-series length requirement is not well defined, and that analysis should be attempted even when the number of available data points is formally inadequate.

Ramsey, Sayers, and Rothman (1990) however describe in thorough detail a study of the behaviour of correlation dimension when applied to small samples of economic data. They were unable to find signs of saturation of the dimension estimate and concluded that there was no evidence for the presence of a chaotic attractor in their small sample data.

There is an inherent limitation on the maximum number of data points available when using EEG recordings. The state of the brain is probably most reasonably assumed to remain stable for periods of time probably not exceeding several seconds. Time-series recorded over intervals greater than several seconds therefore probably contain segments of differing statistical properties, in which case the time-series would not be statistically stationary (Frank et al., 1990). With the typical sampling rates of 100 to 300 samples per second, the number of data points available is on the order of several hundred. Sampling at greater rates will not alleviate this problem. At increasingly high sampling rates, adjacent data points are increasingly correlated. Essentially, no new

information is captured by sampling at greater than some optimal rate, determined by the highest frequency of interest in the data. Such over-sampling can have an effect on correlation dimension. Grassberger (1986) looked at the effect of interpolation on correlation dimension, finding that correlation dimension decreases with increasing amounts of interpolation.

The implications of these results for the present study is that although sample sizes of several hundred points may be analyzed using the Grassberger-Procaccia algorithm, the results cannot be used to infer the presence of a chaotic attractor, whether or not the slope values are seen to tend to a limiting value.

#### VARIABLE TRANSFORMATION:

When the phase-space reconstruction leads to a highly non-uniform attractor, it may be difficult to calculate correlation dimension. The slope of the graph of  $\ln C(r)$  vs  $\ln r$  can in this situation have more than one intermediate straight-line region, with different values of slope for each region. At the same time, the value of slope may not tend to the sought-after limiting value, the estimate of correlation dimension.

In these circumstances, one approach has been to transform the dependent variable in the time-series. Lefranc, Hennequin, and Glorieux (1992) used the log function for this

transformation. The result was a more uniform attractor geometry, longer scaling regions, and saturation of the slope value. This approach assumes that relevant information is contained at all amplitudes of the time-series, and hence in all parts of the attractor. The transformation makes this information more uniformly available to the correlation dimension calculation.

### 1.3 ATTRACTOR GEOMETRY AND UNDERLYING DYNAMICS

The geometry of an attractor is not sufficient to decide whether or not the associated dynamical system is chaotic. For example, deterministic, non-chaotic, attractors generated by systems of equation have been found, which nevertheless have a fractal dimension value, and which are therefore classed as strange (Grebogi, Ott, Pelikan, and Yorke, 1984; Ding, Grebogi, and Ott, 1989; Romeiras, Bondeson, Ott, Antonsen, and Grebogi, 1987). Typically, the generating systems of equations consist of non-linear oscillators driven at incommensurate frequencies (see sections 2.3 to 2.6).

These results emphasize once more the point made earlier that inferences about the nature of the generating dynamical system cannot be made on the basis of correlation dimension calculations. Typically, in order to specify such a dynamical system as chaotic or non-chaotic requires a determination of the extent to which the system is sensitive to perturbations of initial conditions. This determination is usually made by

calculating the Lyapunov exponent for the time-series. Frank et al. (1990) adopt this approach to the analysis of EEG recordings of epileptic seizure events, finding support for the notion that the underlying dynamics are deterministically chaotic. They conclude that the determination of chaos could not have been made without the calculation of Lyapunov exponents.

In a complementary study, Osborne and Provenzale (1989) found a class of stochastic systems with finite values of correlation dimension. A purely stochastic system will not have a finite value of correlation dimension. Dimensionality will in principle be equal to the number of data-points. The particular class of systems used in this study were of the "coloured noise" variety: the time-series exhibited a power-law spectrum. Such noise is common in physical system, and is typically referred to as  $1/f$  noise: the power spectrum decays with increasing frequency as the inverse of frequency. Intuitively, Osborne and Provenzale's (1989) finding results from the "coupling" of the individual degrees of freedom associated with the individual data points, by the power-law function that defines the power spectrum. In the limit of a power-law function with a large exponent, high frequencies in the time-series are effectively filtered out, and a sinusoidal function of correspondingly low dimensionality remains.

The implication of these results for the present study is that it would be unsafe to draw inferences from the behaviour

of the correlation dimension about whether or not the underlying dynamics are or are not chaotic, and about the extent to which the time-series which were studied represent deterministic dynamics or stochastic behaviour.

Pritchard and Duke (1992) emphasize the point that the Grassberger-Procaccia algorithm is most realistically useful in a relative sense, in comparing systems for evidence of dissimilar complexity, rather than in attempting to determine the absolute complexity of a single system. They advocate the term "dimensional complexity" rather than correlation dimension.

#### 1.4 ALGORITHMS

A number of algorithms were developed specifically for carrying out the analysis for the present work. The source code for these algorithms was written in structured BASIC and compiled using QuickBasic version 4.5, a BASIC language compiler.

**CORDIM:** Calculation of correlation dimension was carried out by means of the CORDIM algorithm which implements the Grassberger-Procaccia algorithm, with some modifications. The first modification is the use of the city-block distance measure between points on the attractor. The second modification involves taking an average over only a subset of points on the attractor when calculating the correlation integral, a

strategy proposed by Moon, 1987. These modifications significantly speed up the running time of the algorithm. The algorithm is presented in section A1.

LAGFIND: The LAGFIND algorithm is an attempt to automate the process of finding the optimum value of the lag parameter, for a univariate time-series. In contrast with other, analytic approaches to this problem, LAGFIND implements a geometric approach.

The LAGFIND algorithm successively embeds the time-series in a two dimensional phase space, with increasing values of lag, beginning with lag = 1. It then calculates at these successive values of lag, two sums of squares deviation of points on the attractor; one from each of two orthogonal axes in the phase space,  $r_1$  and  $r_2$ . Denoting the basis axes of the phase space by  $y_1$  and  $y_2$ , the orthogonal axes used for the sum of squares calculation are defined by the equations:

$$\text{axis } r_1: y_2 = y_1 \quad \dots \text{eqn } 5$$

$$\text{axis } r_2: y_2 = -y_1 + 2y_1' \quad \dots \text{eqn } 6$$

where  $y_1'$  is the mean of the values of points on the attractor along the  $y_1$  axis.

At each value of lag, the two sums of squares that are calculated,  $ss_1$  and  $ss_2$ , are compared for magnitude. At low values of lag,  $ss_1$  will be small and  $ss_2$  will be large,

reflecting the compressed shape of the attractor along axis 1. As the value of lag is increased, the value of ss1 will increase and the value of ss2 will decrease. When the ratio of ss1 to ss2 increases to a criterion value (set at .8 for the present work), the associated value of lag is selected as being the optimum value. At this value of lag, the attractor is approximately equally distributed along axes 1 and 2. The attractor is therefore approximately symmetrically distributed in the phase space.

One obvious limitation of this algorithm is that only a single, 2-dimensional phase-space is used for the attractor reconstruction. It is conjectured that a value of lag which results in an optimally distributed attractor in this 2-space will result in an optimally distributed attractor in phase-spaces of higher dimensionality. An improvement on the present algorithm might be to embed the time-series in a phase-space of the nominal dimensionality expected for the associated attractor, rather than in a 2-dimensional space.

Tests of the LAGFIND algorithm on time-series of known properties have indicated that the algorithm correctly finds the optimum value of lag in the following cases:

- uniformly-distributed random noise: computed lag = 1
- sine function: computed lag = one-quarter of the period of the sine wave

The algorithm is displayed in section A2.

REGRESS: The actual computation of correlation dimension involves calculating the slope of a regression line drawn through a series of points along the intermediate portion of the ogive-shaped graph of  $\ln C(r)$  vs.  $\ln r$ . This slope calculation is carried out by the REGRESS algorithm, shown in section A3. REGRESS also computes the error of regression for the computed value of slope. This error of regression is used as an indication of goodness of fit of the regression line to the target points on the curve.

The REGRESS algorithm is also designed to optionally select the region of the curve that will produce the minimum value of error of regression. That is, REGRESS can find the straightest portion of the curve over some preselected number of points, and will then calculate the slope of this optimum curve segment. Although functional, this option was not used in the analyses carried out in this project, because of the limitation in the number of available data points. This limited number of points was conjectured to be the cause of a highly non-uniform attractor. This non-uniformity resulted in there being in some cases several straight-line segments, of different slopes, within the intermediate portions of the curve. When this occurred, it was found that this REGRESS option was not sufficiently sophisticated to consistently select a curve segment which coincided with a segment judged

most appropriate on the basis of visual pattern analysis. Work remains to be done on this portion of the algorithm, although a general impression from perusal of the literature is that most workers estimate slope by means of some heuristically based visual analysis.

## 2 EXPERIMENT 1: Evaluation of the CORDIM Algorithm

The CORDIM algorithm can be tested by analyzing time-series which correspond to generating systems with known characteristics. Tests using three types of time-series were performed. The first set used noise, both random and pseudo-random. The expectation was that correlation dimension would increase and be equal to the dimensionality of the phase-space that the time-series was embedded within. The second set of tests used time-series defined by equations involving 2 and 3 variables. In this case the expectation was that the correlation dimension should reach a limiting value equal to the number of variables in the equations used to generate the time-series. The third set of tests involved time-series which were designed to roughly approximate the spike waveform data, with two different equations: one using 2 variables to generate the simulation, and the other using 3 variables for the simulation.

### 2.1 PSEUDO-RANDOM NOISE

A time-series consisting of 1024 data-points of pseudo-

random noise was generated using the RND function in Quick-Basic. This time-series is shown in Figure 1. This pseudo-random noise generator outputs a uniformly-distributed sequence of data-points. Correlation dimension was calculated for the time-series, using a lag of 1 for the phase-space reconstruction, shown in Figure 2, and an embedding dimension of from 1 to 5. The lag of 1 was chosen because adjacent points in the time-series are independent. The results are shown in Figure 3. There is no clear evidence that the correlation dimension is tending towards a limiting value. This result is consistent with the pseudo-random nature of the mechanism generating the time-series.

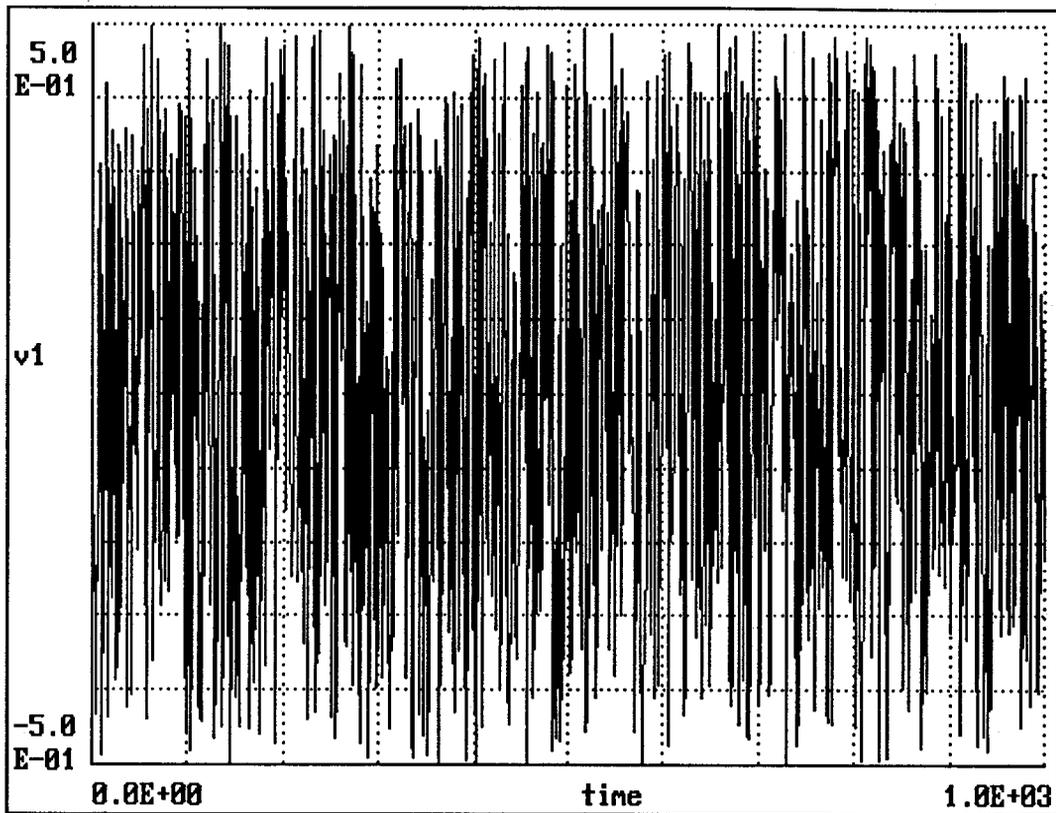


Figure 1 - Time-series generated using pseudo-random number generator function (BASIC RND function); length is 1024 data points.

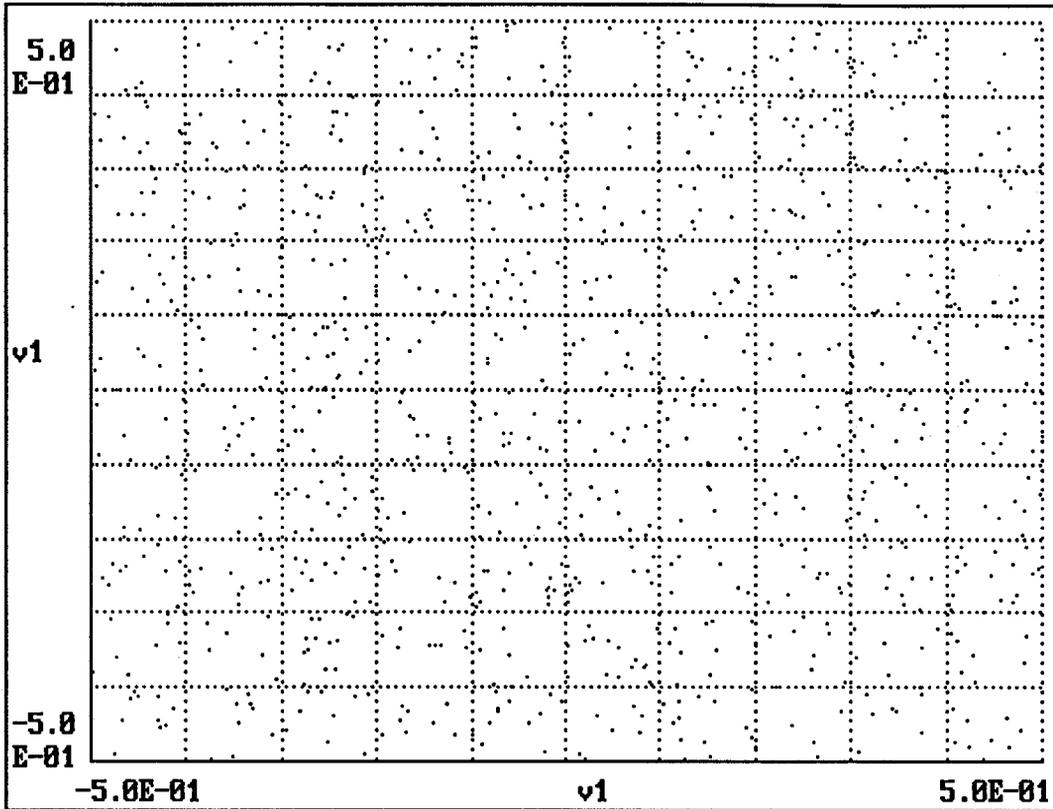


Figure 2 - Phase-space representation generated by plotting pseudo-random time-series against itself using a lag of 1.

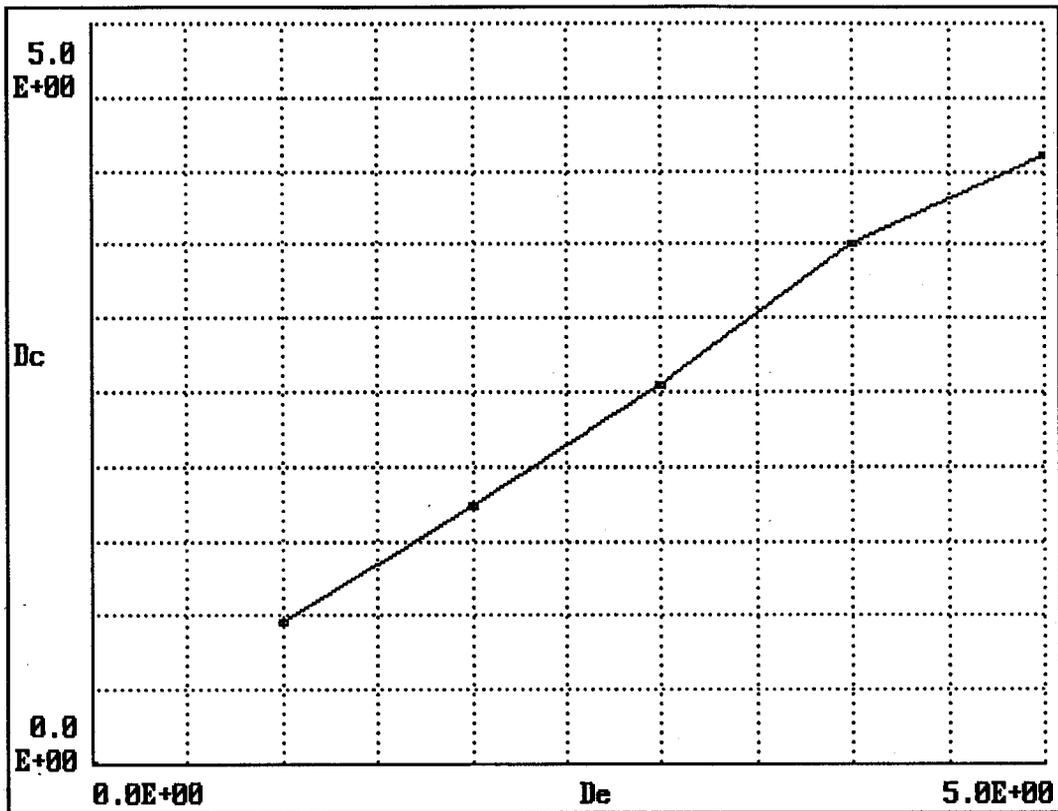


Figure 3 - Correlation dimension as a function of embedding dimensions 1 to 5, for the pseudo-random time-series of Figure 1. No saturation is evident.

## 2.2 RANDOM NOISE

A time-series consisting of 1024 data-points of random noise was generated using an electronic amplifier circuit with a 10 megohm source resistance. The time-series is shown in Figure 4. The output of the amplifier consisted of a combination of thermal noise generated by the source resistance, and amplifier input noise. The thermal noise component results from the discontinuous nature of electron flow through the source resistance, and is uniformly distributed. The amplifier input noise component is generated in the input stages of the amplifier by processes which are circuit dependent. This noise component has a "1/f" power spectrum: Power level in equal frequency bands, for example per 1 Hz, varies in inverse proportion to frequency. Thus large amplitude power components occur at low frequencies, and the amplitude of the power components decreases as frequency increases. This form of noise has an approximately normal distribution: For the 1/f component, relatively more frequent data-points have relatively smaller amplitudes. The output of the amplifier was digitized at 300 samples per second. The resulting time-series was analyzed in the same way as the pseudo-random data, using a lag of 1 for the phase-space reconstruction (Figure 5). The results are shown in Figure 6. Again, and as expected, saturation is not evident.

In principle, with an unlimited number of data-points being available, the calculated values of correlation dimen-

sions should approach the value of embedding dimension. With the limited number of data points used in these analyses, the most that can be expected is that the value of correlation dimension will not tend to a limiting value as embedding dimension is increased. This result is consistent with other studies of correlation dimension. For example, Osborne and Provenzale (1989) studied the properties of random time-series with inverse power-law frequency spectra. Such frequency spectra are described by the function  $1/f^a$ , where  $a$  is termed the scaling exponent. A time-series, such as the one studied in the present experiment, has an associated value of  $a=1$ , giving the  $1/f$  power spectrum. At larger values of  $a$  the power per unit bandwidth (within a 1 Hz band) decreases ever more rapidly with increasing frequency. Osborne and Provenzale (1989) found that for  $a=1$ , corresponding to the situation in the present experiment, there was little evidence of the correlation dimension tending to a limiting value, for embedding dimensions 1 to 5. This is consistent with the result obtained in the present experiment.

Interestingly, Osborne and Provenzale (1989) found that as the value of  $a$  increased, there was a tendency for the correlation dimension to reach a limiting value. At  $a=3$  correlation dimension saturated at a value of 1. The authors suggested that as  $a$  increases, the dynamics of the time-series are increasingly dominated by the effects of the power law function, which involves only a single variable,  $a$ .

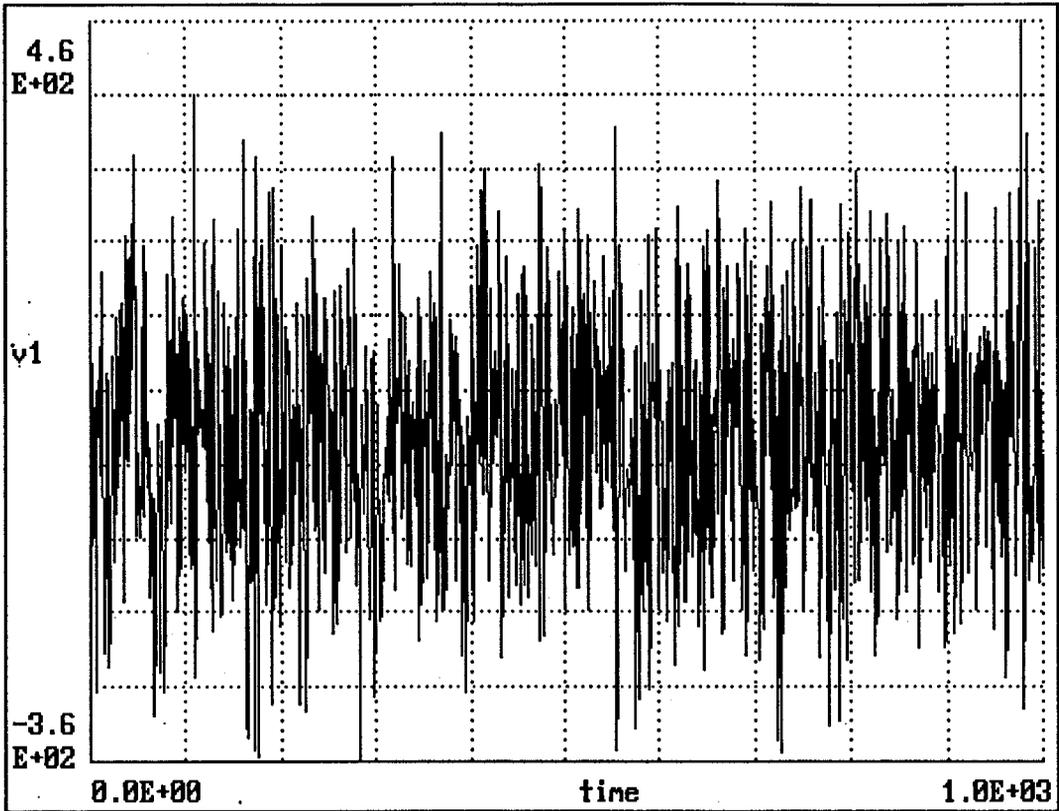


Figure 4 - Random time-series, comprised of thermal and 1/f noise components; 1024 data points.

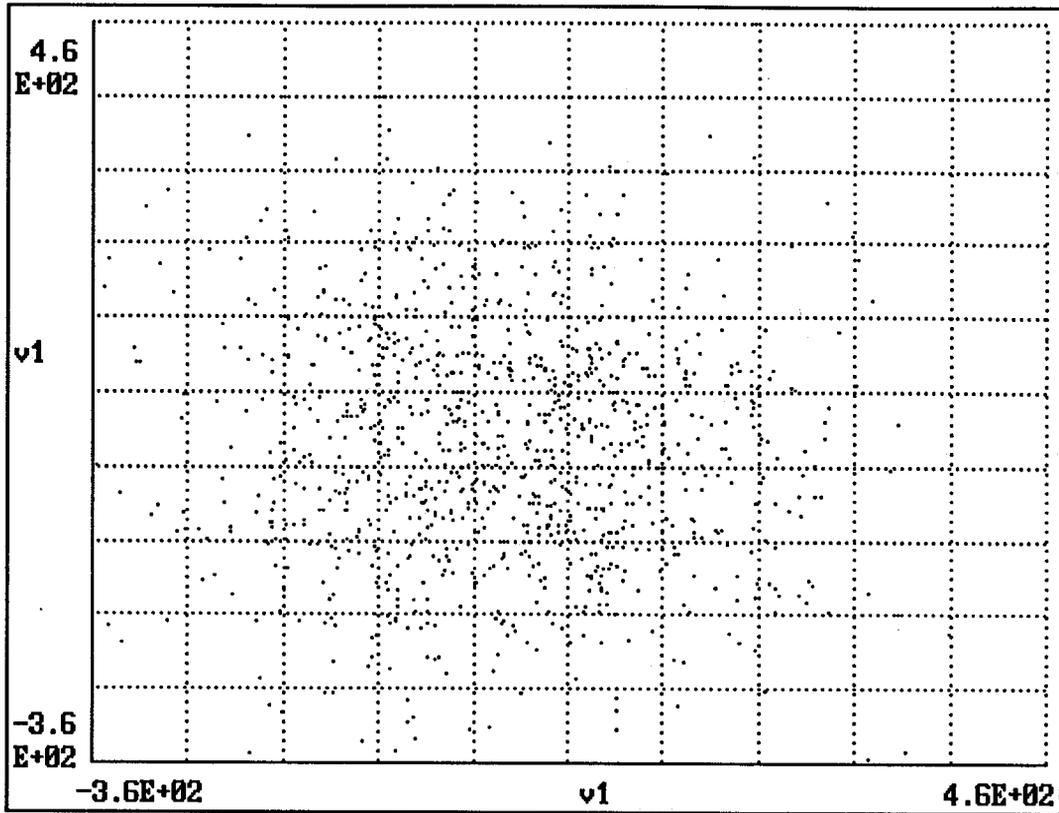


Figure 5 - Phase-space attractor constructed by plotting random time-series against itself with a lag equal to 1.

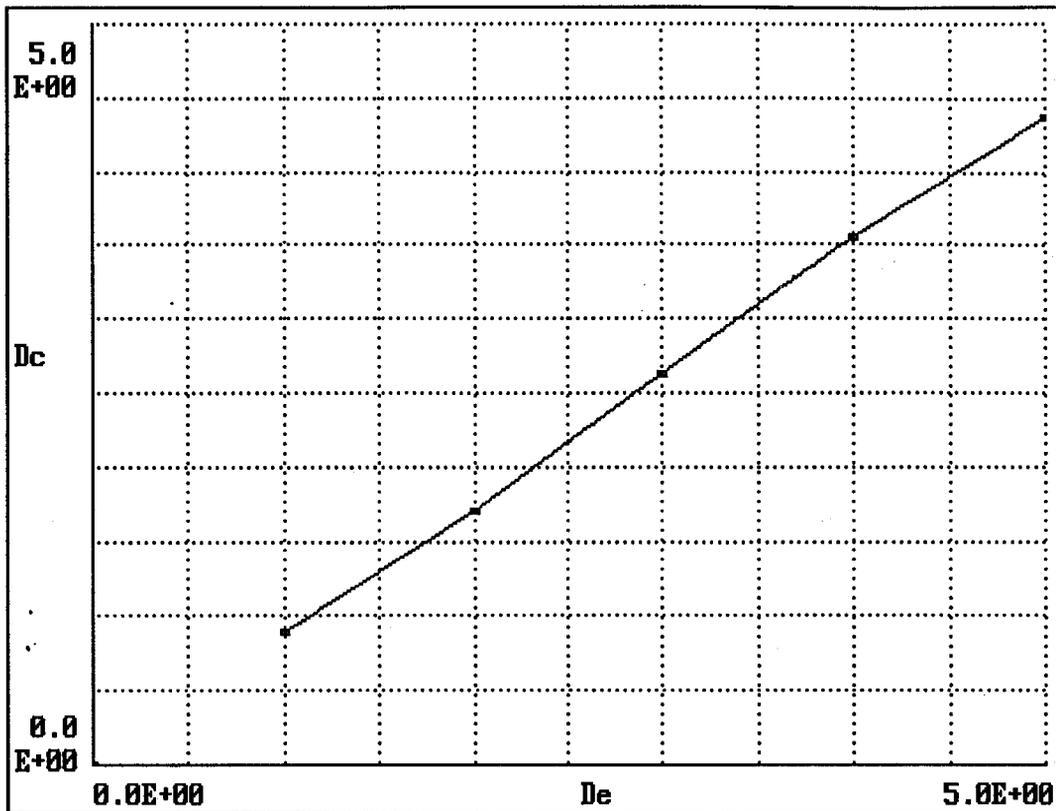


Figure 6 - Correlation dimension as a function of embedding dimensions 1 to 5, for random time-series of Figure 4. No saturation is evident.

### 2.3 THE 2-TORUS

A time-series of 1024 points was generated using the following equation in two variables:

$$y(t) = \sin(2\pi at/n) + \sin(2\pi bt/n) \quad \dots \text{eqn 7}$$

where  $a = 10$ ,  $b = 10\sqrt{2}$ , and  $n$  is the length of the time-series. The resulting time-series is shown in Figure 7. The frequencies have been chosen to be incommensurate; that is, they are not expressible as a ratio of integers. When this time-series is used to construct an attractor in phase-space using the Takens-Packard method of reconstruction, the attractor has the shape of a 2-dimensional torus. Because the frequencies have been chosen to be incommensurate, the trajectory defined by the equation travels along the surface of this torus without ever exactly repeating a previous orbit. The phase-space attractor for this time-series was reconstructed using a value of lag of 18, and is shown in Figure 8. This value of lag was computed using the LAGFIND algorithm, described below. Correlation dimension was then calculated for this attractor, at embedding dimensions 1 to 5. The results are shown in Figure 9. As expected correlation dimension appears to tend to a limiting value of roughly 2, correctly approximating the number of variables used in constructing the original time-series.

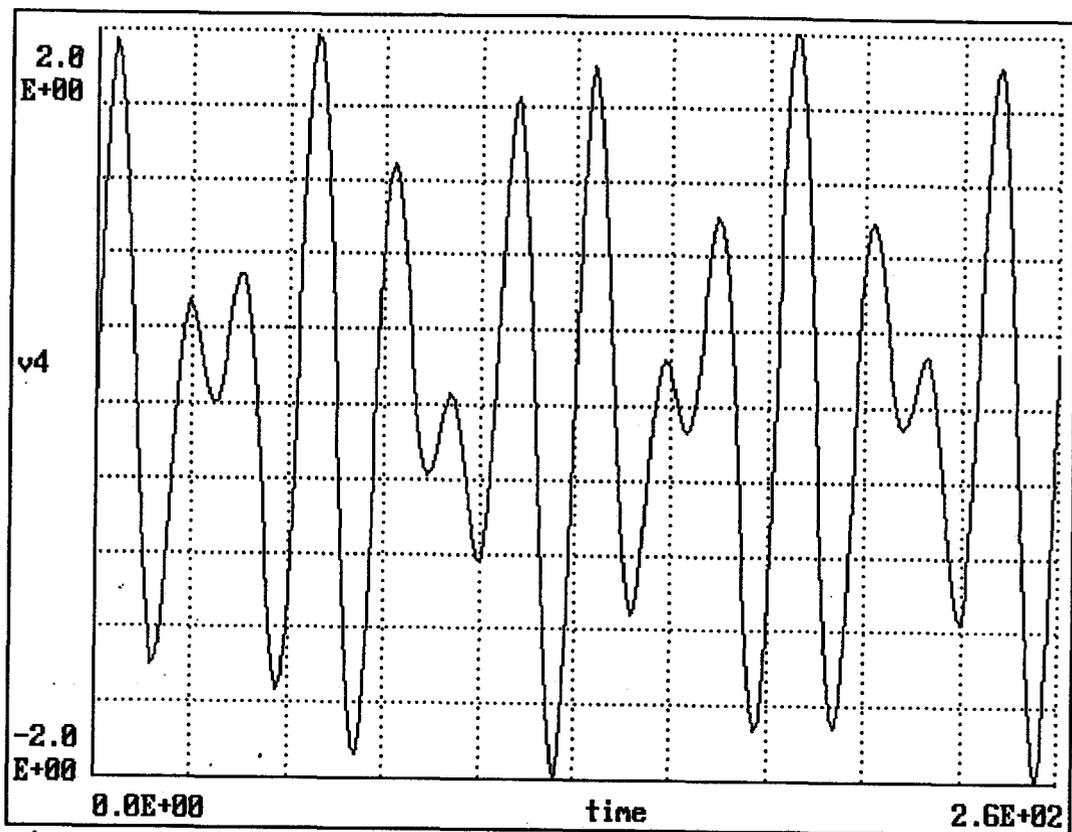


Figure 7 - Time-series generated by the equations defining the 2-torus, the summation of 2 sine-waves of incommensurate frequencies.

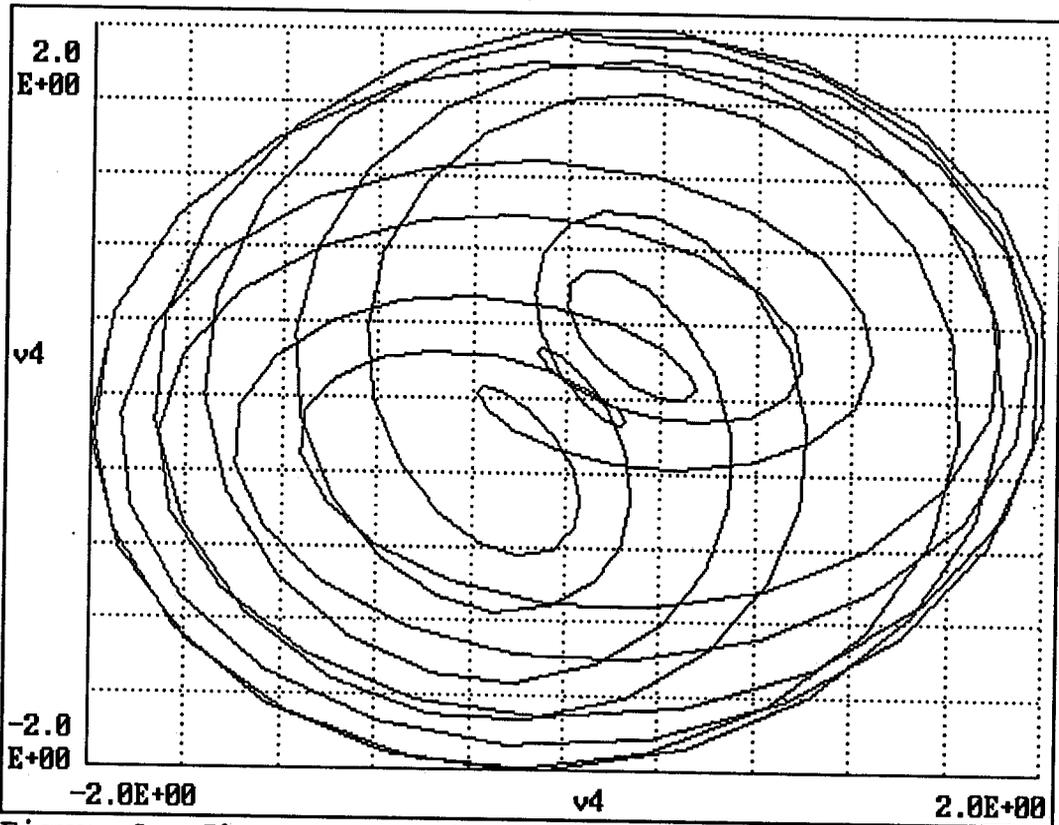


Figure 8 - Phase-space attractor constructed by plotting the two-torus time-series against itself using a lag equal to 5.

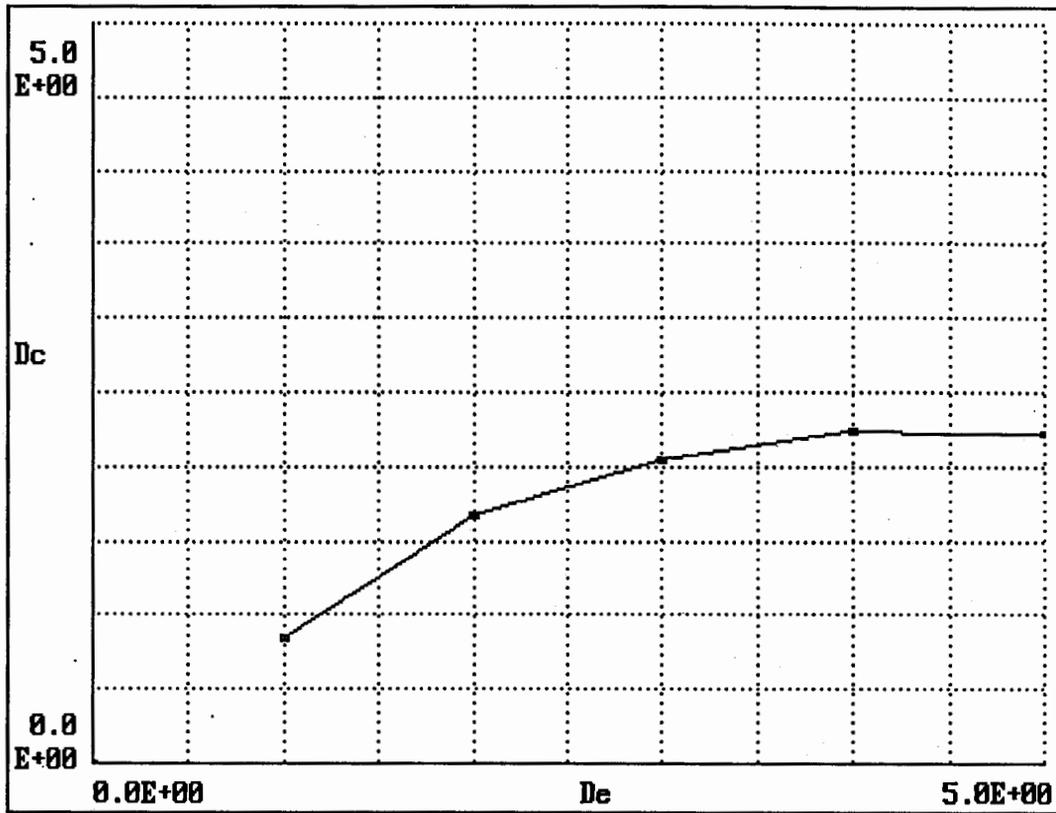


Figure 9 - Correlation dimension vs. embedding dimensions 1 to 5 for 2-torus time-series. Saturation is clearly evident.

## 2.4 THE 3-TORUS

A time-series of 1024 points was generated using the following equation in 3 variables:

$$y(t) = \sin (2\pi at/n) + \sin (2\pi bt/n) + \sin (2\pi ct/n) \dots \text{eqn 8}$$

where  $a = 10$ ,  $b = 10\sqrt{2}$ ,  $c = 10\sqrt{3}$ , and  $n$  is the length of the time-series. The resulting time-series is shown in Figure 10. As for the 2-torus, these 3 frequencies are chosen to be mutually incommensurate. This time-series results in an attractor in phase-space that has the shape of a 3-dimensional torus. A 2-dimensional projection is shown in Figure 11. Correlation dimension was calculated for this time-series using a lag of 18, computed by the LAGFIND algorithm. The results are shown in Figure 12. As expected, the value of correlation dimension tends towards a limiting value of approximately 3, again consistent with the number of variables used in constructing the original time-series.

In accordance with theoretical considerations, the value of correlation dimension for the 3-torus is greater than for the 2-torus. In both cases, the absolute value of correlation dimension is approximately equal to the number of variables in the original equations.

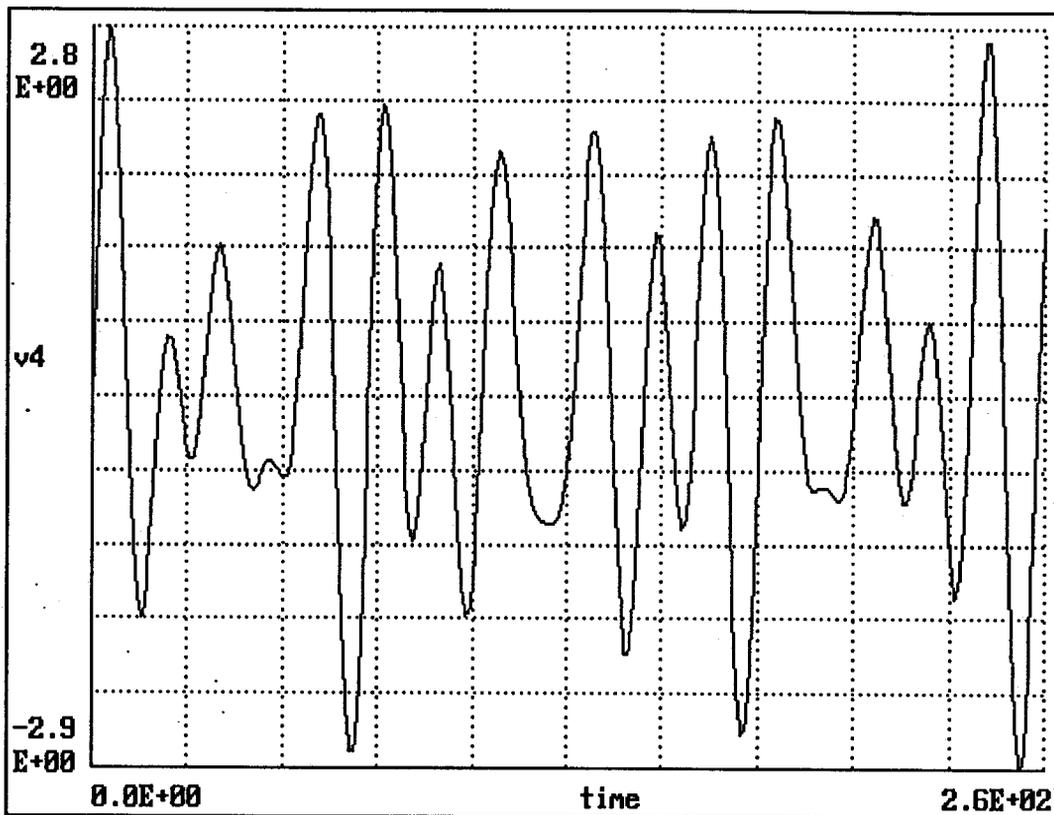


Figure 10 - Time-series generated using equations defining the 3-torus, the summation of 3 sine-waves of incommensurate frequencies.

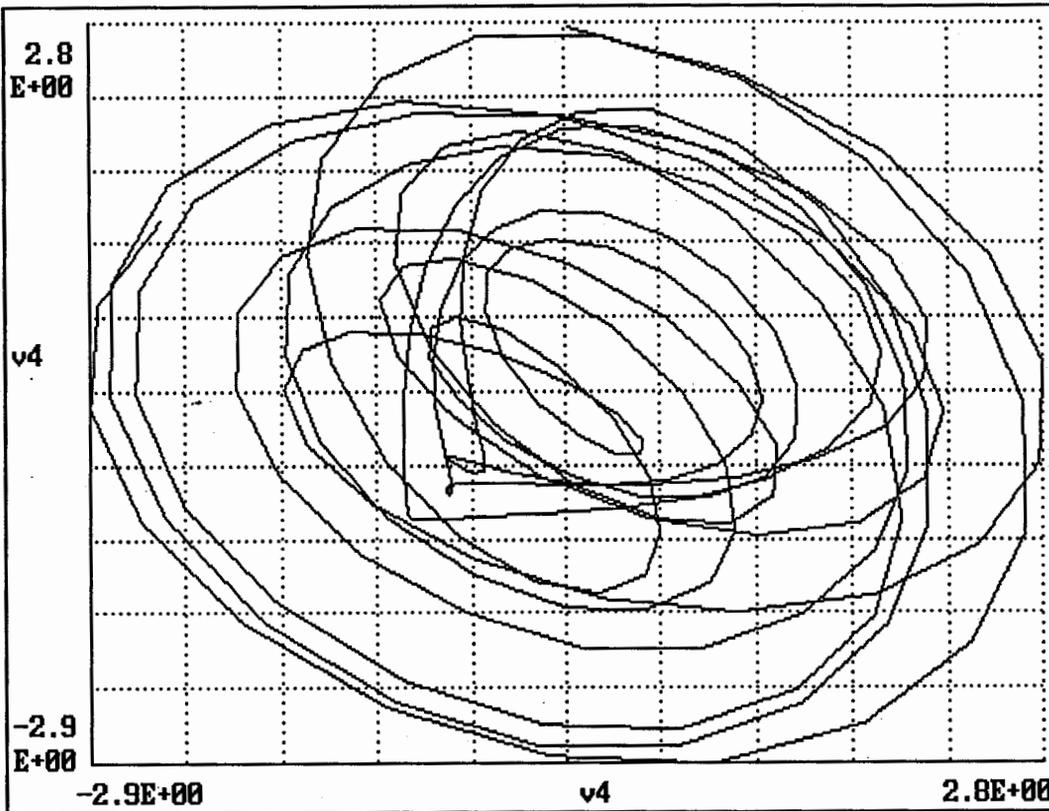


Figure 11 - Phase-space attractor constructed by plotting the three-torus time-series plotted itself using a lag equal to 5.

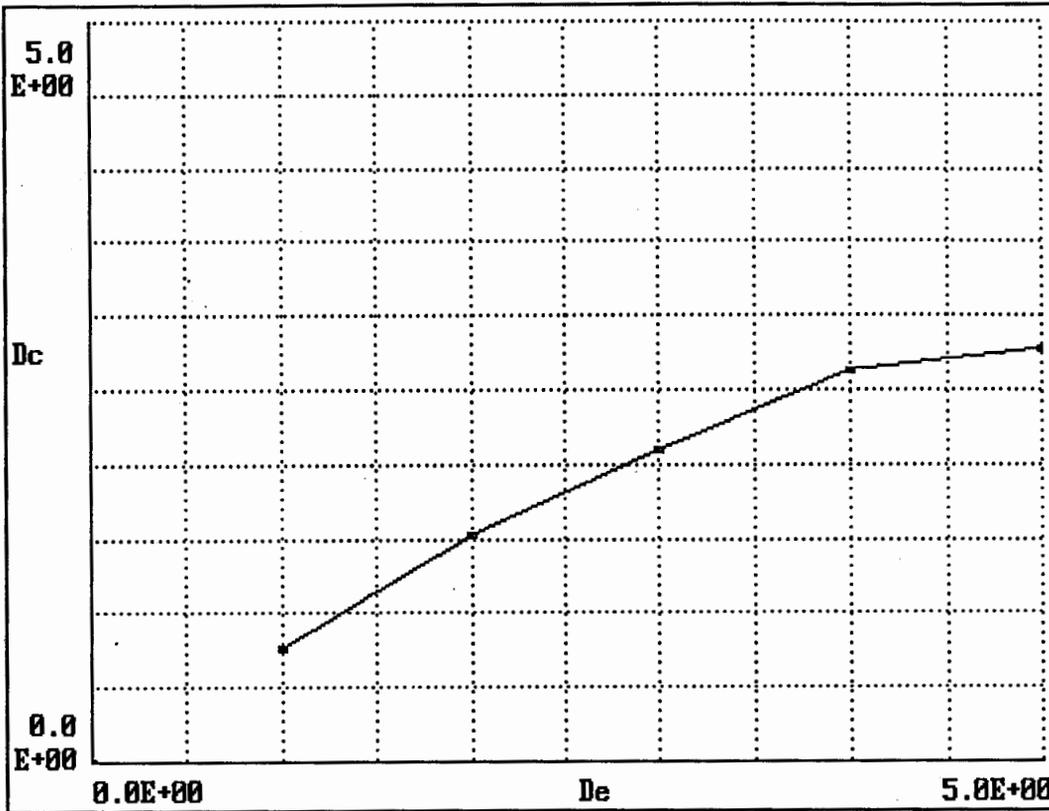


Figure 12 - Correlation dimension vs. embedding dimensions 1 to 5 for the 3-torus time-series.

## 2.5 SPIKE WAVEFORM SIMULATION USING 2 VARIABLES

In order to simulate a spike waveform using only 2 variables, a time-series of 256 points was generated using the following equation:

$$y(t) = (\sin (2\pi at/n)) \cdot (-\cos (2\pi bt/n) + 1) \quad \dots \text{eqn 9}$$

where  $a = 6$ , and  $b = 1$ .

The resulting waveform (Figure 13) is intended to approximate the BREC spike waveform using a trigonometric equation in 2 variables. No attempt was made to optimize the closeness of the approximation. Rather, the equation was designed to be both a simple mathematical model, while generating a time-series visually resembling the BREC spike. The phase-space attractor is shown in Figure 14. This attractor was analyzed for correlation dimension using a lag of 10, computed using LAGFIND, at embedding dimensions of 1 through 5. The correlation integral was averaged over 32 equally-spaced data points on the time-series. The results of this analysis are shown in Figure 15. Correlation dimension reached a value of 1.98 at an embedding dimension of 5, showing some evidence of tending towards a limiting value.

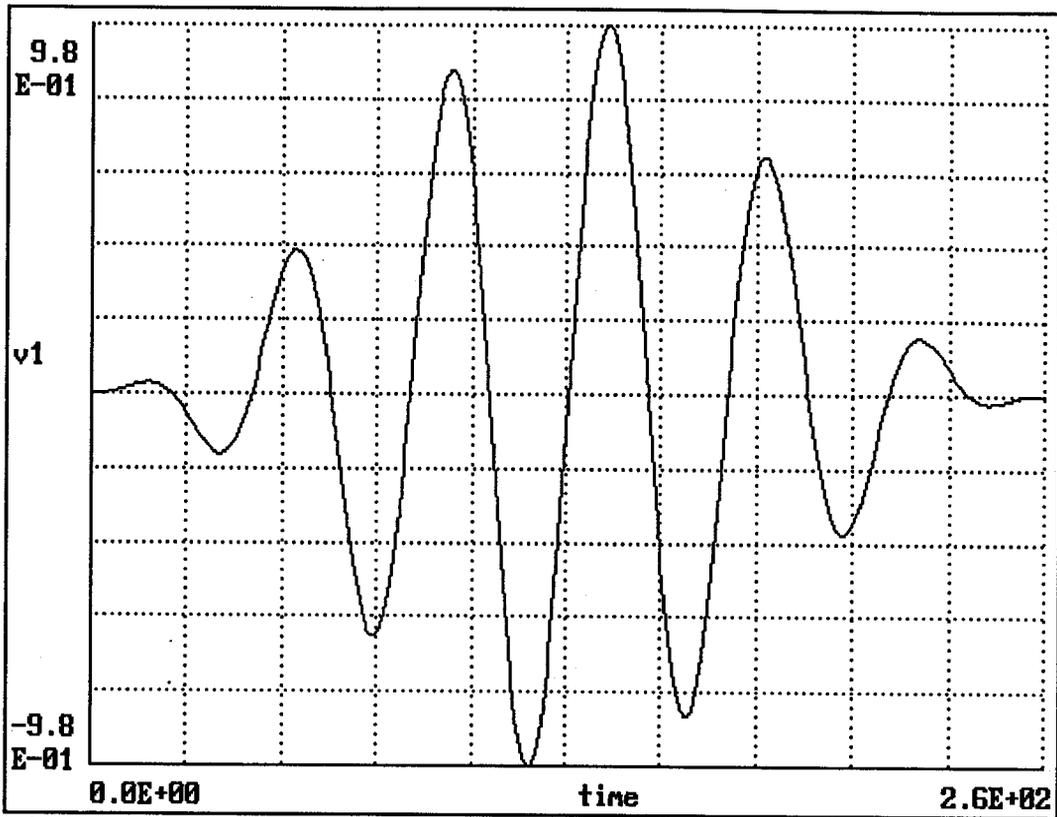


Figure 13 - An approximation to the BREC spike waveform generated using 2 equations in 2 variables.

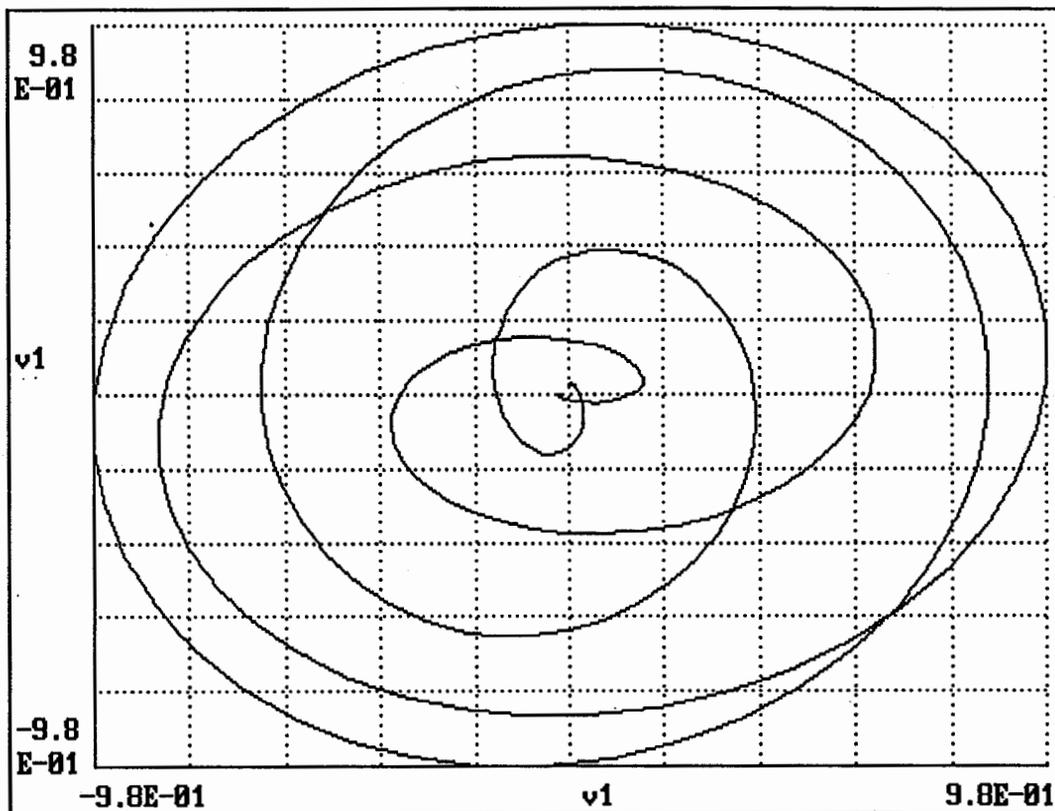


Figure 14 - Phase-space attractor generated by plotting the two variable spike waveform approximation against itself using a lag equal to 10.

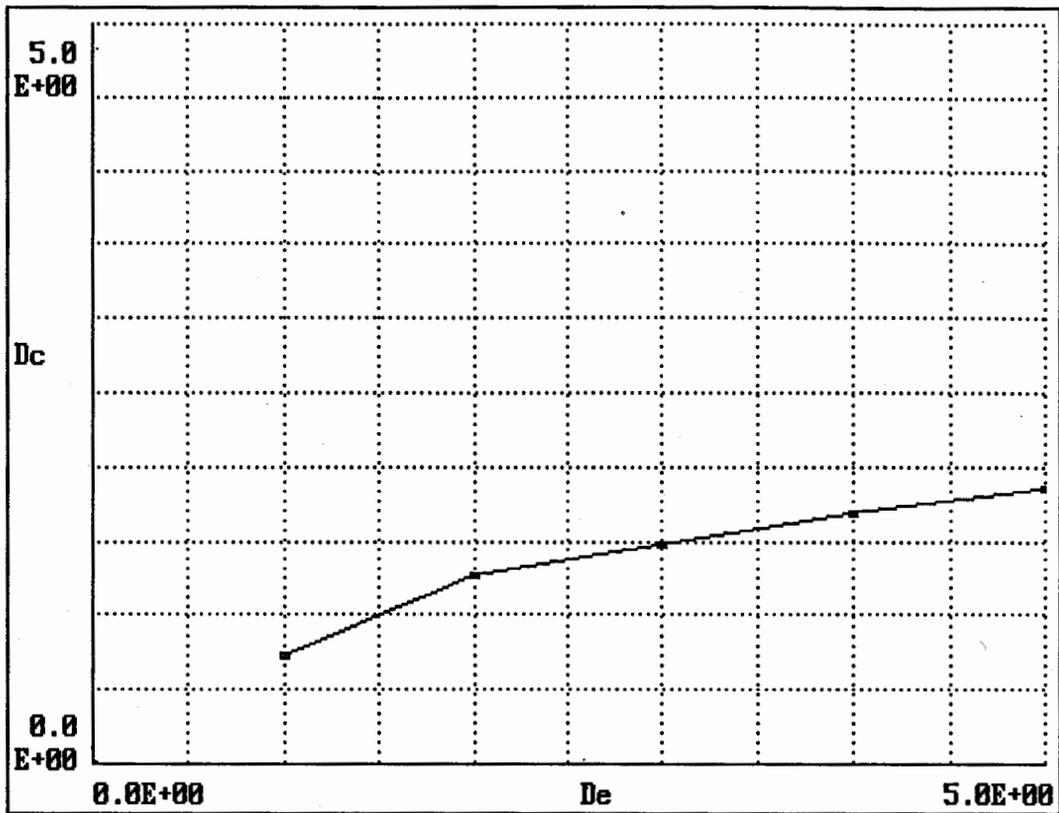


Figure 15 - Correlation dimension vs. embedding dimensions 1 to 5 for the 2-variable spike approximation time-series.

## 2.6 SPIKE WAVEFORM SIMULATION USING 3 VARIABLES

A simulation of the spike waveform was constructed using an equation in 3 variables. A time-series of 256 points was generated using the following equation:

$$y(t) = (\sin(2\pi at/n)) \cdot (\cos(2\pi bt/n)) \cdot (\sin(2\pi ct/n)) \quad \dots \text{eqn 10}$$

where  $a = 6$ ,  $b = 1$ , and  $c = 0.5$ .

The resulting waveform (Figure 16) is intended to approximate the BREC spike waveform using a trigonometric equation in 3 variables. Again, no attempt was made to optimize the closeness of the approximation. The equation was designed to be a simple mathematical model that could generate a time-series visually resembling the BREC spike. The phase-space attractor for this time-series is shown in Figure 17. Correlation dimension was calculated for this attractor using a lag of 10, computed using the LAGFIND algorithm, at embedding dimensions of 1 through 5. As before, the correlation integral was averaged over 32 data-points. The results of the analysis are shown in Figure 18. Correlation dimension attained a value of 2.63 at an embedding dimension of 5, showing some signs of tending to a limiting value.

Consistent with expectations based on theoretical considerations, and on the behaviour of correlation dimension using the 2-torus and 3-torus data, correlation dimension for the 3 variable simulation is greater than for the 2-variable

simulation. In both cases the absolute value of correlation dimension qualifies as being a lower bound on the number of variables in the generating system.

## 2.7 SUMMARY

In all cases, correlation dimension behaved in accordance with theoretical considerations. Two general results may be seen. Firstly, the value of correlation dimension was consistent with the view that this statistic is a lower bound on the number of variables involved in the system generating the time-series. Secondly, the value of correlation dimension correctly reflected the differential complexity of similar pairs of time-series that differed only in terms of the number of defining variables.

It should be expected therefore that correlation dimension will behave similarly when confronted with time-series constituting experimental rather than simulated data. In particular, values of correlation dimension calculated for the BREC spike waveform time-series should reflect the complexity of the underlying neural generating systems.

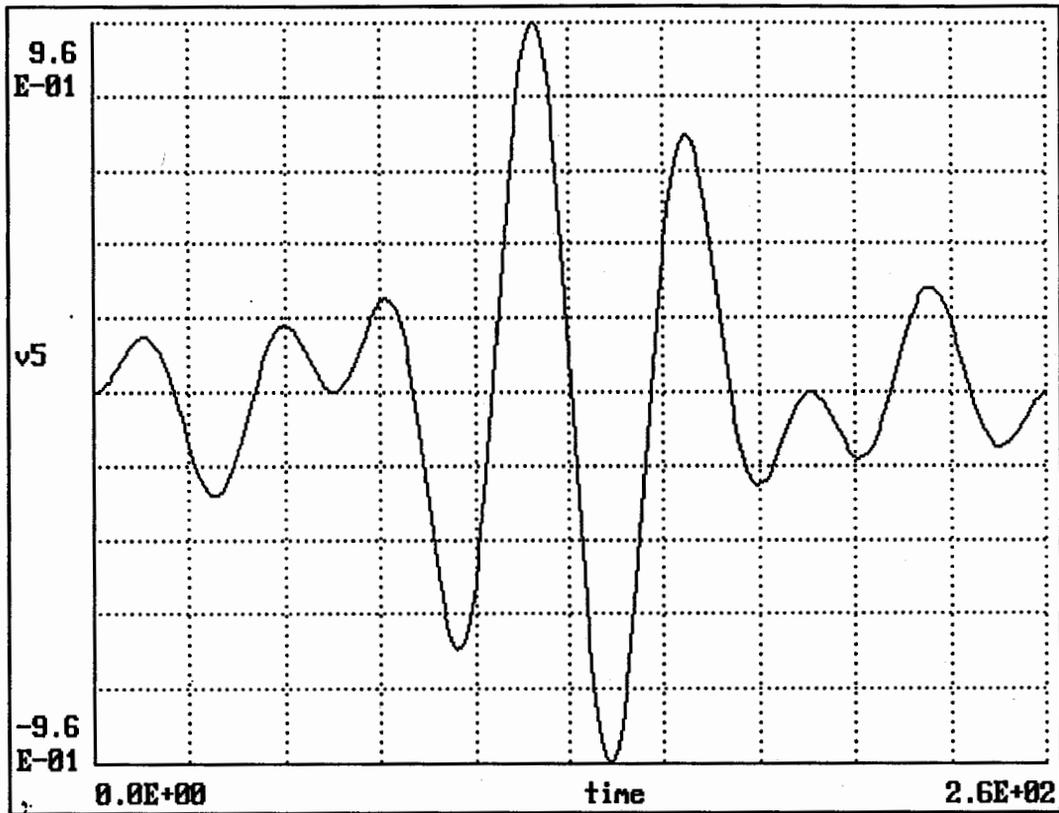


Figure 16 - An approximation to the spike waveform generated using 3 equations in 3 variables.

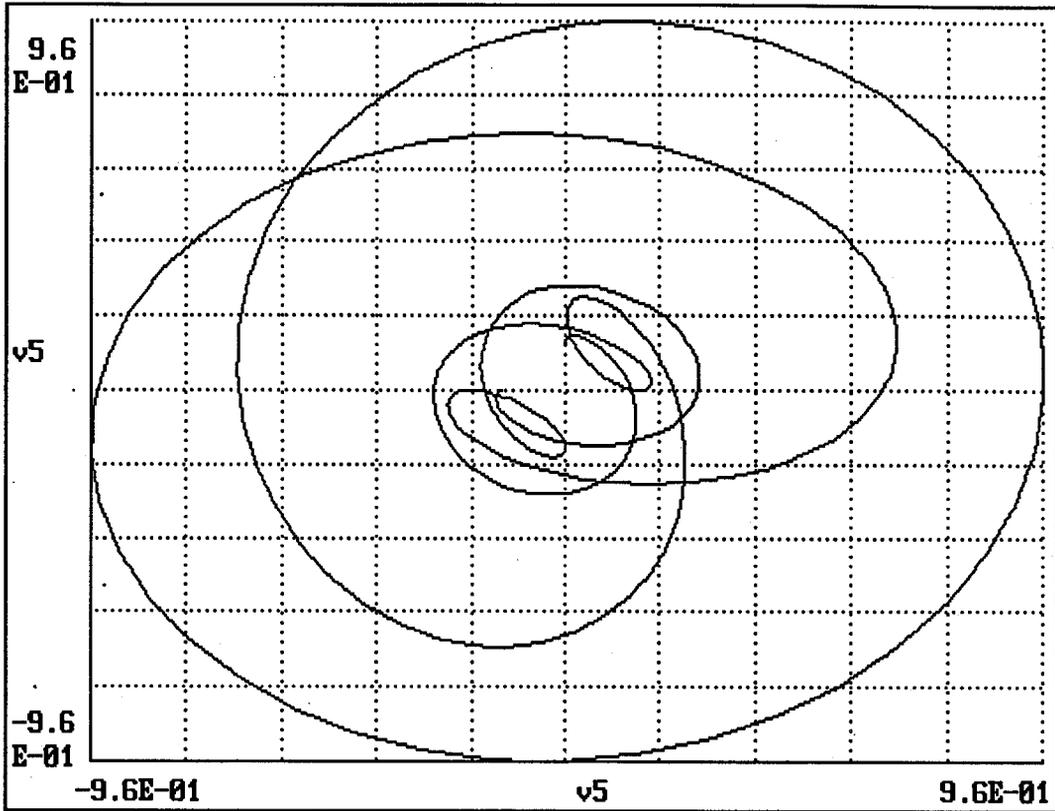


Figure 17 - Phase-space attractor constructed by plotting the 3 variable spike approximation against itself using a lag equal to 10.

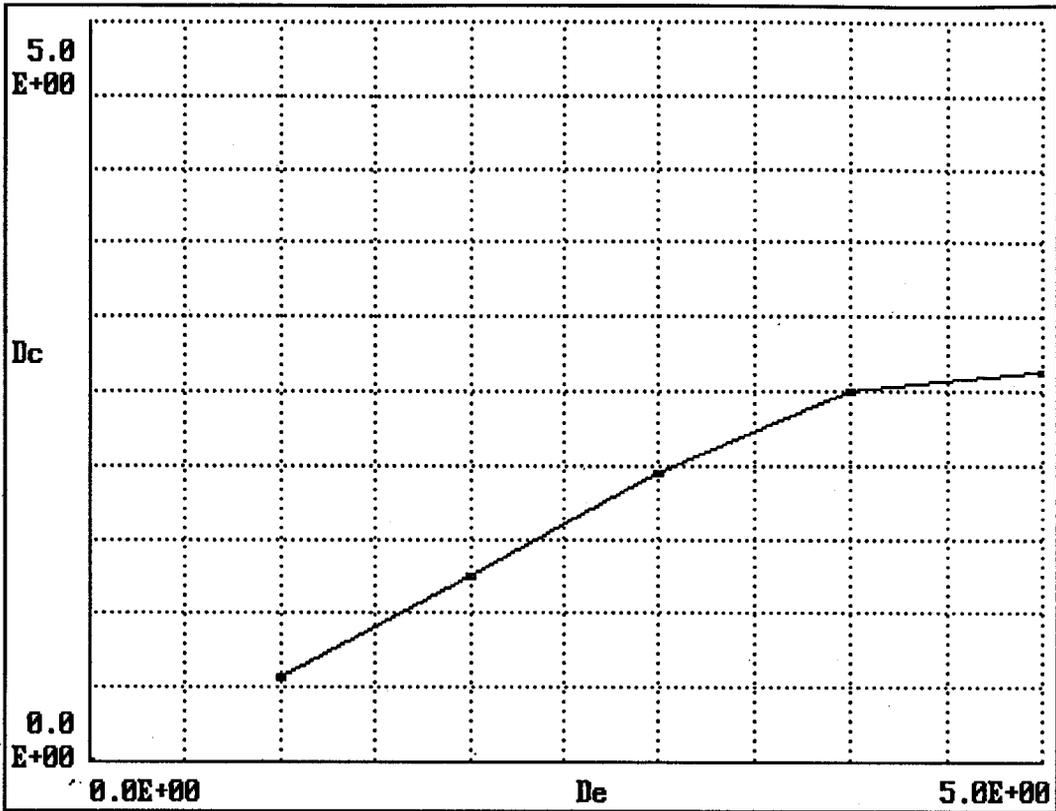


Figure 18 - Correlation dimension vs. embedding dimensions 1 to 5, for the 3-variable spike approximation.

### 3 EXPERIMENT 2: Multi-dimensional Reconstruction

It has been suggested that the Grassberger-Procaccia algorithm for calculating correlation dimension is prone to an upward bias for attractors of low dimension embedding in high dimensional spaces, and that this bias may be reduced by a modification to the algorithm (Dvorak, 1990). This modification involves making use of the multiple channels available in, for example, a typical EEG recording. At any one time-point, data points from a set of channels become the components of the vector that is plotted in phase space to form the attractor.

This approach has the advantage of making use of the spatially as well as temporally extended information available in a multi-channel EEG recording. A second advantage is the circumventing the problem of choosing an optimal lag parameter. An obvious disadvantage is that spatial resolution is sacrificed.

#### 3.1 APPLICATION TO 2-TORUS DATA

For the 2-torus case, vectors for the reconstruction were formed by taking successive sets of points from the 2 time-series formed by the following equations:

$$y_1(t) = \sin(2\pi at/n) \quad \dots \text{eqn 11a}$$

$$y_2(t) = \sin(2\pi bt/n) \quad \dots \text{eqn 11b}$$

where  $a = 10$ , and  $b = 10\sqrt{2}$ . The time-series are displayed in Figure 19. In this way we form the following series of vectors:

$$u(1) = \{ y_1(1), y_2(1) \}, \dots, u(i) = \{ y_1(i), y_2(i) \}$$

A total of 256 vectors were formed in this manner, and the resulting attractor (Figure 20) was analyzed for correlation dimension using the Grassberger-Procaccia algorithm. The results are shown in Figure 21. The correlation dimension showed clear evidence of saturation, reaching a value of 1.8 at an embedding dimension of 3. This result is consistent with the number of variables involved in the generating system of equations, and used in the construction of the phase-space attractor.

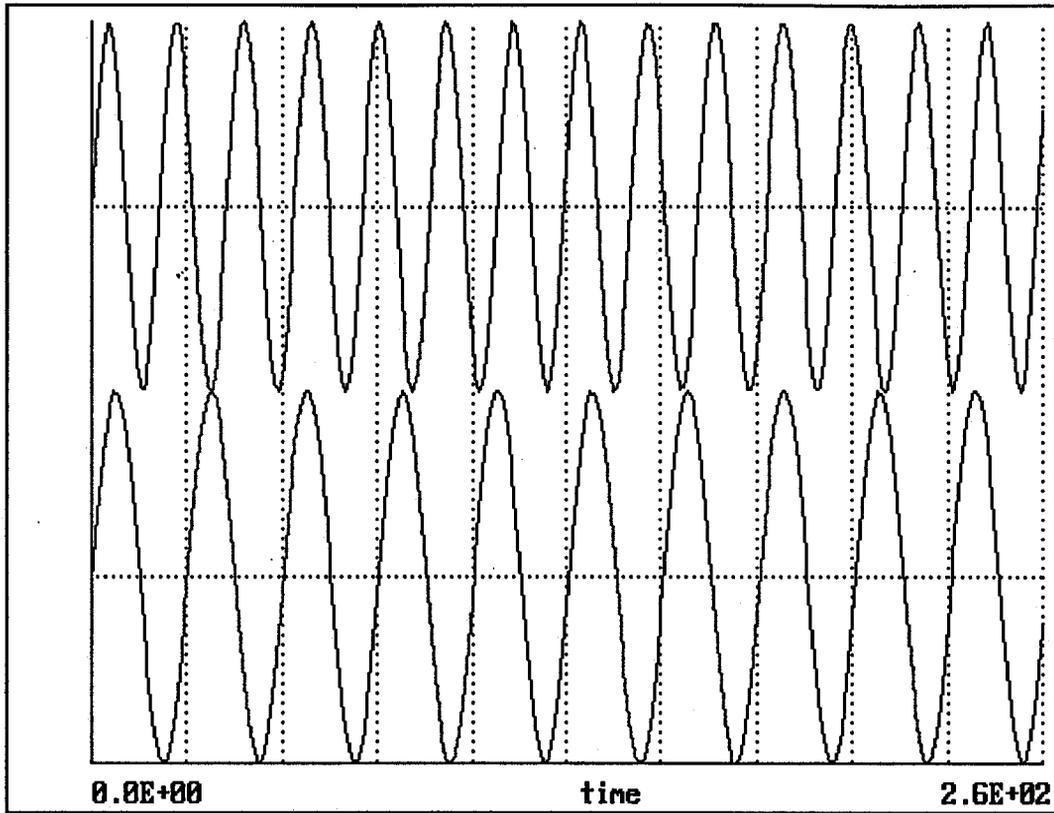


Figure 19 - The 2 torus time-series used for the multi-channel reconstruction method.

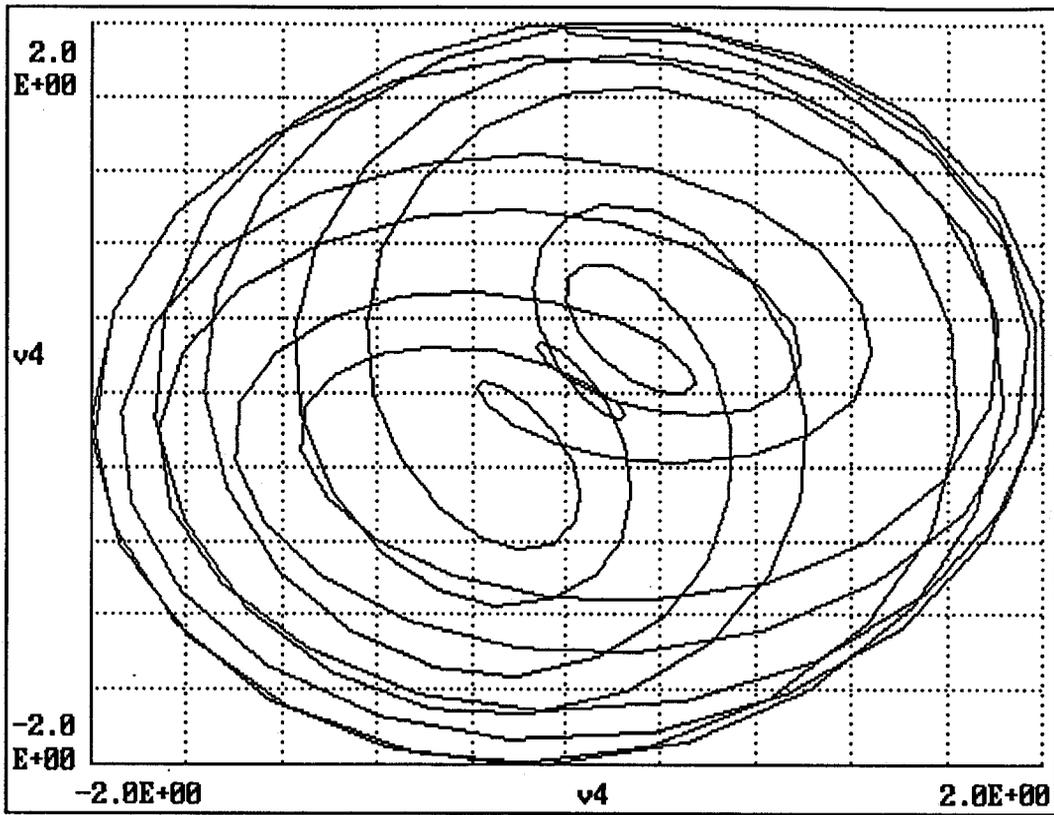


Figure 20 - Two-torus attractor used with the multi-channel reconstruction method.

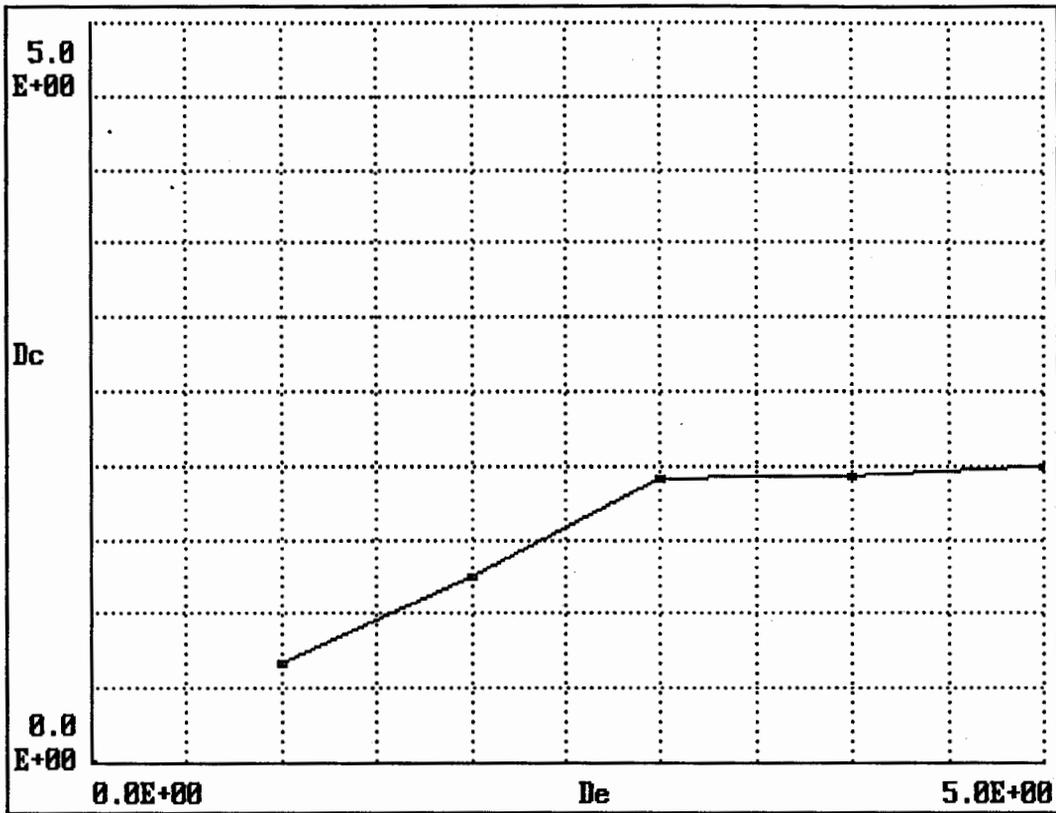


Figure 21 - Correlation dimension vs. embedding dimensions 1 to 5 for multichannel reconstruction method using the 2-torus data.

### 3. APPLICATION TO 3-TORUS DATA

For the 3-torus case, vectors for the reconstruction were formed by taking successive sets of points from the 3 time-series formed by the following equations:

$$y_1(t) = \sin(2\pi at/n) \quad \dots \text{eqn 12a}$$

$$y_2(t) = \sin(2\pi bt/n) \quad \dots \text{eqn 12b}$$

$$y_3(t) = \sin(2\pi ct/n) \quad \dots \text{eqn 12c}$$

where  $a = 10$ ,  $b = 10\sqrt{2}$ , and  $c = 10\sqrt{3}$ . The time-series are shown in Figure 22. Thus, we obtain the series of vectors:

$$u(1) = \{ y_1(1), y_2(1), y_3(1) \}$$

...

$$u(i) = \{ y_1(i), y_2(i), y_3(i) \}$$

A total of 256 vectors were formed in this manner, and the resulting attractor (Figure 23) was analyzed as above for the 2-torus case. The results are shown in Figure 24. The correlation dimension showed evidence of saturating, reaching a value of 2.7 at an embedding dimension of 5. This result is consistent with the number of variables in the generating system of equations.

The results indicate that the multi-channel reconstruction method is able to accurately estimate the dimensionality of the related generating systems of equations.

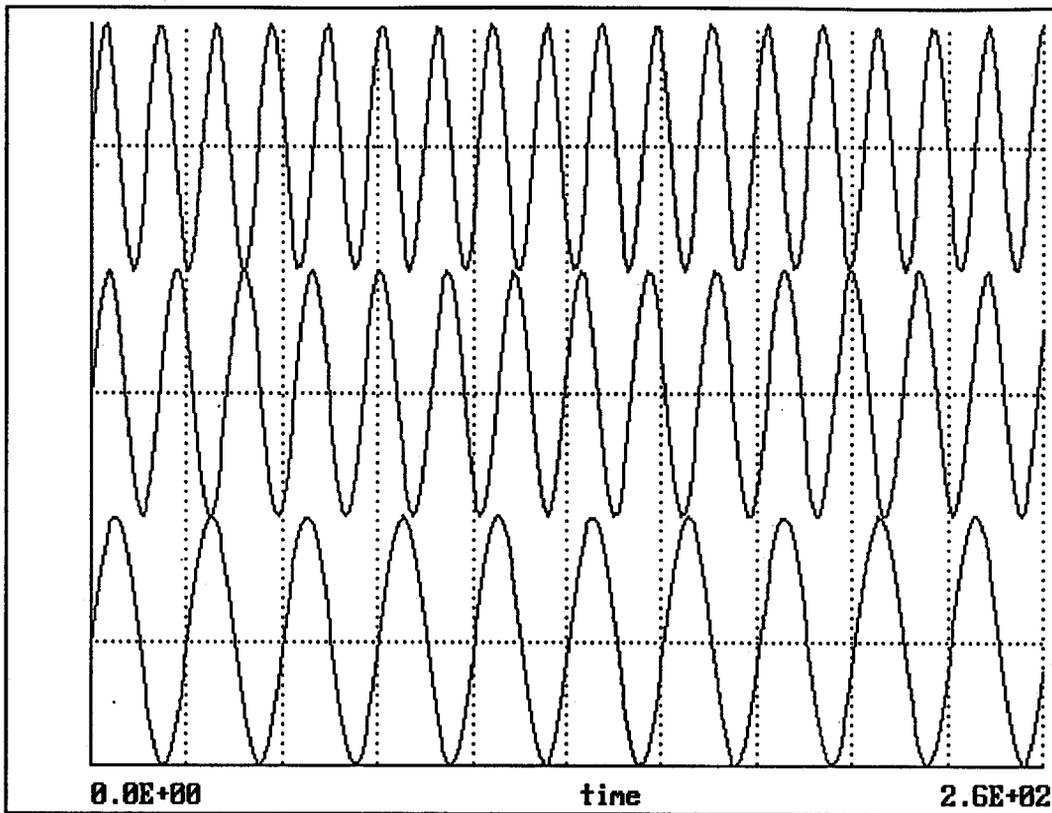


Figure.22 - The 3 torus time-series used with the multi-channel reconstruction method.

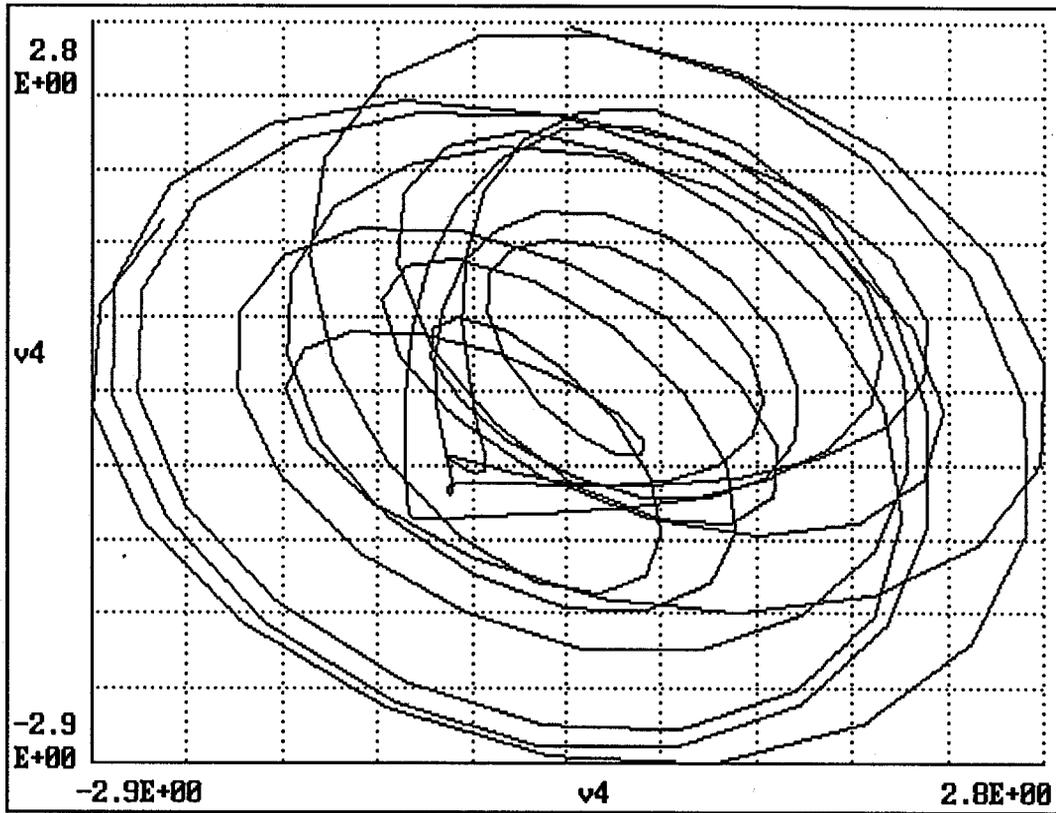


Figure 23 - Phase-space attractor for the 3-torus time-series used with the multi-channel reconstruction.

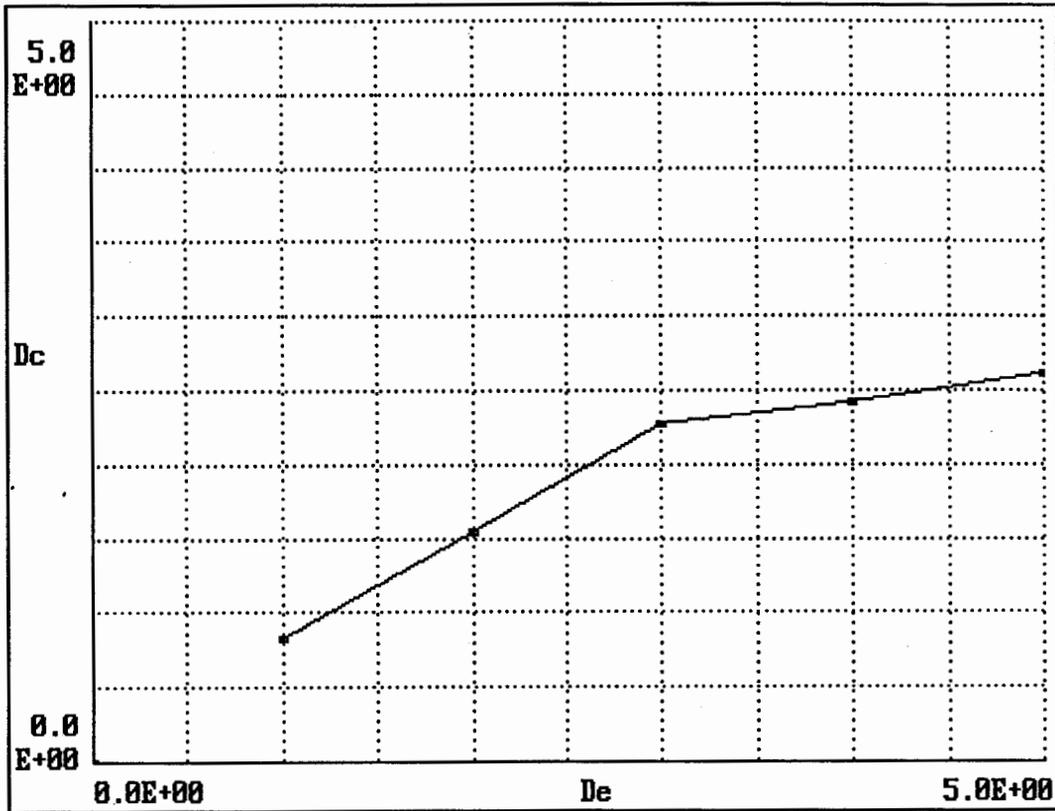


Figure 24 - Correlation dimension vs. embedding dimensions 1 to 5, for the multi-channel reconstruction method using the 3-torus data.

## 4 ANALYSIS OF BREC SPIKE WAVEFORMS

### 4.1 INTRODUCTION

With both the Typical and the Atypical BREC patterns, there are associated characteristic interictal EEG spike waveforms. Morphologically, these waveform consists of a single approximately sinusoidal main pulse or spike of relatively high amplitude, with a period of from 80 to 120 ms. Sample spike waveforms that have been clinically diagnosed as Typical and Atypical are shown in Figure 25 and 26 respectively. This main pulse may be preceded and followed by lower amplitude oscillations. It has been demonstrated that a discrimination between the Typical and Atypical cases may be made on the basis of characteristics of these components of such EEG spikes (Bencivenga, 1987; Wong, Bencivenga, and Gregory, 1988).

The general approach adopted in the present work was suggested by the results of a study of BREC spike waveforms by Bencivenga (1987). The approach taken by Bencivenga (1987) involved a non-parametric classification strategy: A decision procedure or tree is constructed, with the nodes being decisions regarding a set of observable features of the EEG waveform in the vicinity of the spike. These features were selected on the basis of studies made by Gregory and Wong (1984), and observations of medical staff at the B.C. Children's Hospital. Included are polarity reversals of the spike, number and spatial extent of spikes, synchronization of spikes

across channels, presence of after-waves, spatial location and amplitude of spikes, and coordinates of the estimated focus. Of these variables, the last three, transverse location of the focus, and location and amplitude of the negative portion of the spike, were found to be sufficient to permit classification with a misclassification rate of 20%. This result was interpreted as suggesting that the Typical and Atypical cases were associated with different underlying processes (Wong et. al. 1988).

This strategy suggested the analysis used in the present study: calculating estimates of the correlation dimension for the time-series comprised of the spike waveform, along with preceding and following electrical activity. It was speculated that correlation dimension might capture a significant portion of the information associated with the variables studied by Bencivenga (1987), and thus would differ in magnitude for the Typical and the Atypical time-series.

Initial examination of the data suggested that Typical spikes appeared to be an in-phase continuation of preceding spontaneous activity to a greater extent than Atypical spikes. The spike thus would appear to be more closely related to the preceding activity, for the Typical waveforms (Figure 25) than for the Atypical waveforms (Figure 26). On the basis of this observation it could be argued that the mechanism underlying the generation of the waveforms is more complex for the Atypical waveforms than for the Typical waveforms.

The complexity of the hypothesized generating mechanism responsible for the spike waveforms can be estimated by the correlation dimension. More specifically, the correlation dimension is related to the minimum number of variables needed to describe the behaviour of the system; that is, the number of degrees of freedom at work within the system. If for example two systems were found to have unequal correlation dimensions, it would suggest that the two systems were comprised of mechanisms of unequal degrees of complexity. Wong et al.'s (1988) conclusion is that Typical and Atypical BREC involves different underlying mechanisms, and that the mechanisms associated with Atypical BREC might be more heterogenous relative to those associated with Typical BREC, resulting from an interaction between the spike event and existing neurological abnormalities.

This statement suggests the hypothesis for the present study with respect to the BREC data, that there might be a corresponding difference in the correlation dimensions for the two types of BREC, and that the correlation dimension for Atypical waveforms would be larger than the correlation dimension for Typical waveforms, reflecting the differential complexities of the underlying systems.

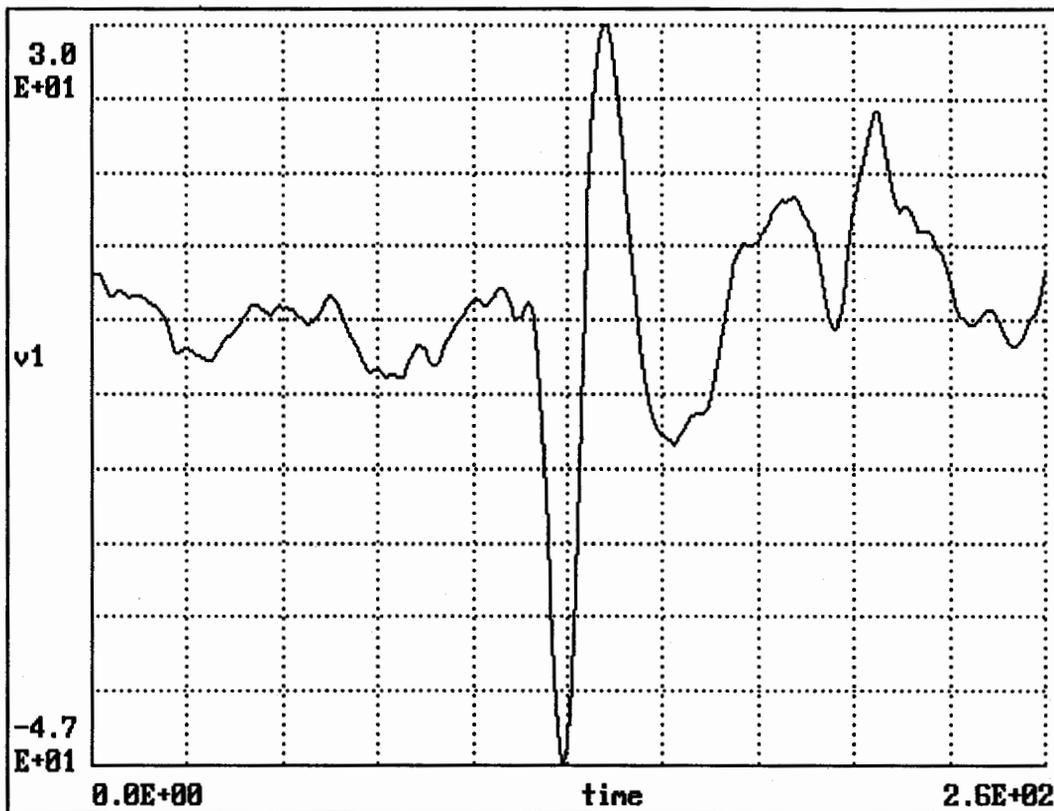


Figure 25 - Example of a Typical spike waveform; 256 data-points sampled at 200 points per second, for a time-span of 1.28 seconds (patient no. 1, channel T4).



Figure 26 - Example of an Atypical spike waveform; 256 data-points sampled at 200 points per second, for a time-span of 1.28 seconds (patient no. 12, channel T3).

There is no clear evidence available as to the extent, spatial and temporal, of the brain areas involved in the generation of the spike waveform. Wong (1989) however has conjectured that Atypical spikes are generated by a more spatially extended system than Typical spikes. In this experiment, spike classification will be carried out using 3 approaches in an attempt to bracket this uncertainty. First, spikes from a single channel common to all cases will be used. Second, spikes from a single channel corresponding to the maximum peak-to-peak spike amplitude for a particular case will be analyzed. Third, in an effort to utilize spatially extended information, a multi-channel reconstruction method will be used, with two different channel ensembles.

## 4.2 EXPERIMENT 3.1: Single Common Channel

### 4.2.1 METHOD

The data consisted of 44 data sets of averaged interictal spike waveforms. The data sets were originally classified as Typical or Atypical, on the basis of presence or absence of neurological and behavioural problems (Wong, Gregory, and Farrell, 1985). These data sets consisted of 29 that had been classified as Typical, and 15 that had been classified as Atypical. The data of each patient represented an average of between 6 and 20 individual spike episodes, measured at 21 points on the scalp using the international 10-20 system. In the original recording of these interictal spikes, filter

settings were 1 Hz high-pass and 70 Hz low-pass. Digitization rate was 200 samples per second, and the length of each record was 256 points. The time-interval was therefore 1.28 seconds.

The CORDIM algorithm was used to calculate correlation dimensions for these time-series. The parameters used in the computation of correlation dimension are summarized as follows:

Length of time-series: 256 points

Digitization: 200 samples per second

Original Filter settings: 1 Hz high-pass; 70 Hz low-pass

Averaging: each time-series is the average of 6 to 20 spike episodes

Lag: computed using LAGFIND algorithm; values listed in Table 1

Embedding dimension: 4

Number of data points used for computation: 32 points, equally spaced along the 256 point time-series.

The choice of embedding dimension was made on the basis of pilot work with the same data files used in this analysis. As embedding dimension was increased, and with the relatively small number of points available in each time-series (256), it was found that the scaling region, that is, the straight-line portion of the ogive curve of  $\ln C(r)$  vs.  $\ln r$ , decreased in length. At embedding dimensions over 6 the length of this

scaling region was in some cases too short to permit an accurate estimate of slope to be made. At an embedding dimension of 4 the scaling region was sufficiently long in all cases for accurate estimates of slope to be made.

Examination of the 44 files revealed the following distribution of large amplitude spike activity:

- In 24 files (55%), the spike amplitudes were found to be greatest over the right temporal region at electrodes F8, C4, T4, P4, and T6, with the largest number of spikes located at electrode T4.
- In 17 files (39%), the spike amplitudes were greatest over the left temporal region at electrodes T3, C3, P3, T5.
- In 2 files (5%) the spike amplitudes were greatest in the occipital region at electrodes FP1 and FP2.
- In 1 file (2%) the spike amplitudes were greatest over the frontal region at electrode F3.

On the basis of this distribution of spike amplitudes, in this experiment a single common channel was analyzed, electrode T4. For all 44 cases, the time-series corresponding to electrode T4 was analyzed using the CORDIM algorithm. For each individual time-series, the value of lag was computed using the LAGFIND algorithm.

The first step of the analysis involved constructing a phase-space attractor for each time-series. A sample at-

tractor is shown in Figure 27. It should be noted that such representations of attractors are only 2 dimensional projections of an object that is located in a hyperspace of dimensionality equal to the value of embedding dimension being used. A prominent feature of the geometry of this attractor is that it is highly non-uniform: the orbits are concentrated in a relatively small volume near the centre of gravity of the attractor. This characteristic non-uniformity is shared by most of the attractors generated from the spike waveforms used in this study. There is an implication of this non-uniformity for the correlation dimension calculation. The scaling region, the straight-line portion of the graph of  $\ln C(r)$  vs.  $\ln r$ , may be shorter than would ideally be desired for an accurate estimate of the slope of this line to be determined. The more ideal time-series is one which results in an attractor which visits most regions of the embedding phase-space with roughly equal frequency. The simulated spike waveforms used in Experiment 2 were in fact closer approximations to this ideal.

#### 4.2.2 RESULTS

The results obtained in the analysis of the 44 BREC files is summarized in Table 1. Table 1 lists the patient number, diagnosed BREC type, value of lag computed with the LAGFIND algorithm, and the calculated value of correlation dimension.

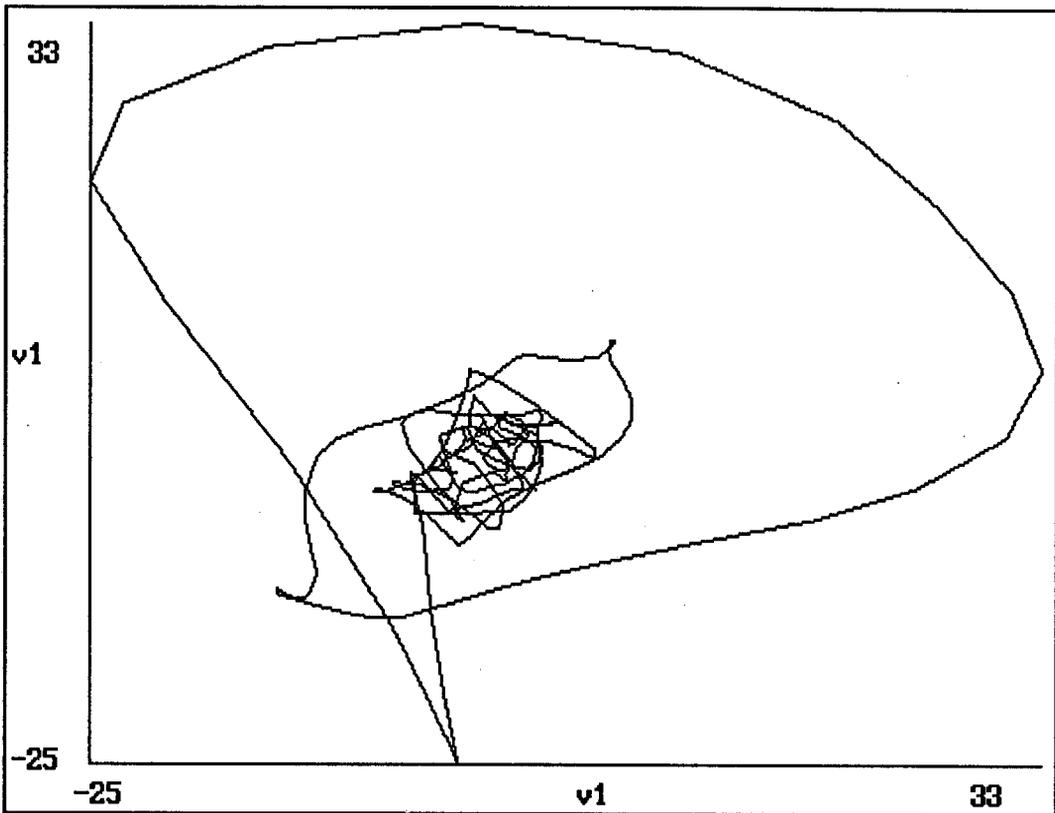


Figure 27 - Phase-space attractor corresponding to the spike waveform for patient no. 1, channel T4. A lag of 8 was used for this construction.

TABLE 1 - Summary: Exp. 3.1 - Common Channel T4

Type 1 = typical; type 2 = atypical

Patient	Type	Lag	Dc
1	1	10	1.857
2	2	27	1.509
3	1	8	2.536
4	1	10	1.982
5	1	8	1.598
6	1	7	2.273
7	1	30	1.817
8	1	11	1.342
9	1	14	1.426
10	1	10	1.348
11	2	9	1.723
12	2	10	2.072
13	2	10	2.081
14	2	30	1.171
15	2	11	1.471
16	1	22	2.299
17	2	10	1.847
18	1	7	2.465
19	1	30	1.73
20	1	10	2.158
21	1	14	1.15
23	1	30	1.274
24	1	16	1.432
25	2	12	2.00
26	1	22	1.833
27	1	9	2.258
28	1	10	1.555
29	1	13	1.363
30	1	10	1.46
31	1	10	1.274
33	2	30	1.525
34	2	11	2.357
35	1	7	1.619
36	2	24	1.778
38	1	13	1.904
39	1	8	1.47
42	2	6	2.406
43	1	7	2.209
44	2	30	1.521
48	2	10	1.779
50	1	12	1.51
53	1	30	1.872
54	1	7	2.398
60	2	30	1.792

Since the values of correlation dimension were found to be approximately normally distributed, a Student's t test was conducted on the correlation dimension values. A point-biserial correlation was also computed, correlating the assigned classification dichotomy with the correlation dimension continuum. The results of these calculations are shown in Table 2.

TABLE 2 - Statistics: Exp. 3.1 - Common Channel T4

Group	Mean	N
Typical	1.773	29
Atypical	1.802	15
Standard Error:	0.123	
t = -0.238, df = 42 (p > 0.25)		
Point-biserial r = 0.0368 (p > 0.1)		

The difference between the Typical and Atypical groups on correlation dimension is not significant in terms of Student's t and point-biserial correlation.

#### 4.2.3 DISCUSSION

The non-significant results are probably attributable to the use of the single common channel for all cases, regardless of the actual location of spike activity for a case. For 20 of the 44 cases, there was little or no evidence of a spike

waveform on channel T4. Analyzing channel T4 in such a case would amount to computing the dimensionality of whatever relatively low amplitude background activity was present. These low amplitude signals could be expected to have a lower signal-to-noise ratio. Here, noise can be operationalized as electrical activity unrelated to the target phenomenon, myoelectric signals, and spurious electrical activity induced into the subject and electrode lead wires from the environment. The dimensionality of such a time-series could be expected to be not systematically different for the 2 groups.

An implication of these findings is that the differences between Typical and Atypical cases in terms of the dynamical system generating the spike waveforms are not localized to the same area of the scalp, under electrode C4, for the 44 cases studied.

An improvement over the present method might be to analyze a unique channel, the channel with the maximum amount of spike activity, for each case. Using the maximum amplitude channel should ensure that in all cases the analyzed channel has a relatively high signal to noise ratio. This modification is implemented in Experiment 3.2.

#### 4.3 EXPERIMENT 3.2: Maximum Amplitude Channel

##### 4.3.1 METHOD

For each of the same 44 files that were used in experiment 3.1, the channel with the largest peak-to-peak spike

amplitude was identified. The identification was carried out using algorithm SELECT1, shown in section A4. All other conditions were identical to those in experiment 3.1.

#### 4.3.2 RESULTS

The results obtained in the analysis of the 44 BREC files is summarized in Table 3, and the analysis of the statistical significance of these results is shown in Table 3. The difference between the Typical and Atypical groups on correlation dimension is significant in terms of both the Student's t-test, and the point-biserial correlation.

Figure 28 shows the 44 data files ranked on correlation dimension, and indicating the original diagnostic classification. The 44 files appear to be roughly separable into two groups. The first group, with the lower values of correlation dimension, contains 20 files, of which 19 are Typical, and one is Atypical. The second group, with the higher values of correlation dimension, contains 24 files, of which 10 are Typical and 14 are Atypical. This observation was tested using cluster analysis on the correlation dimension values and the diagnosed BREC types.

TABLE 3 - Data Summary: Exp. 3.2 - Maximum Amplitude Channel

=====  
 Type 1 = typical; type 2 = atypical (prior diagnosis)  
 =====

Case	Type	Lag	Channel	Dc
1	1	8	13	1.745
2	2	10	10	1.756
3	1	7	15	2.425
4	1	30	13	1.279
5	1	9	13	1.957
6	1	7	13	1.507
7	1	30	9	1.314
8	1	30	9	1.33
9	1	13	13	1.41
10	1	9	13	1.311
11	2	9	12	1.723
12	2	9	9	2.094
13	2	10	9	2.029
14	2	21	5	1.227
15	2	11	12	1.949
16	1	6	9	2.173
17	2	10	12	1.918
18	1	7	18	2.451
19	1	13	13	1.782
20	1	30	13	1.397
21	1	10	9	1.445
23	1	30	10	1.541
24	1	10	9	1.272
25	2	8	9	2.203
26	1	9	13	1.502
27	1	8	9	1.618
28	1	11	14	1.196
29	1	30	15	1.26
30	1	10	13	1.574
31	1	10	13	1.257
33	2	9	1	1.768
34	2	6	18	2.502
35	1	7	12	1.875
36	2	11	10	1.945
38	1	9	13	1.621
39	1	8	12	1.431
42	2	6	12	2.406
43	1	7	13	2.265
44	2	9	9	2.126
48	2	10	12	1.812
50	1	9	13	1.794
53	1	30	9	1.144
54	1	8	13	1.795
60	2	9	3	2.04

TABLE 4 - Statistics: Exp. 3.2 - Maximum Amplitude Channel

Group	Mean	N
Typical	1.609	29
Atypical	1.967	15

Standard Error: 0.110

$t = -3.25$ ,  $df = 42$ ; ( $p < 0.005$ )

Point-biserial  $r = 0.448$  ( $p < 0.005$ )

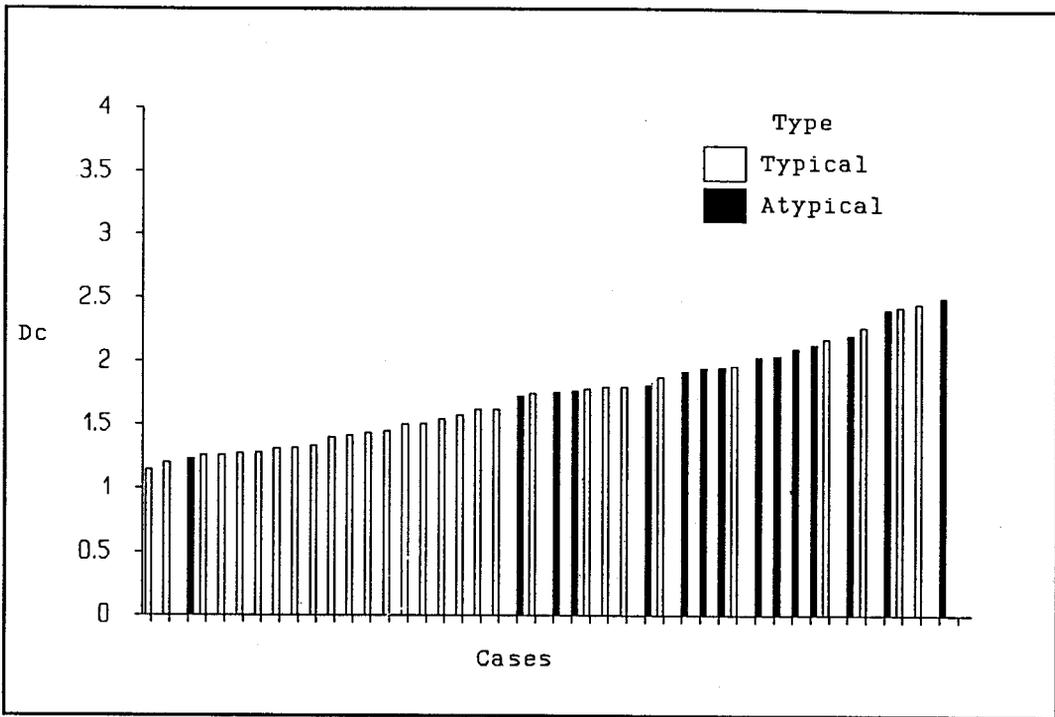


Figure 28 - Cases are sorted by correlation dimension, and labelled by clinical diagnosis of BREC type.

## CLUSTER ANALYSIS

A hierarchical cluster analysis was carried out on the results of Experiment 3.2 using BMDP-2M. The options were no standardization, and centroid amalgamation rule. The results, shown in Table 5, indicate the presence of two main clusters. The first cluster has 19 Typical files and 1 Atypical File. This cluster corresponds to the 20 cases in the left portion of the graph in Figure 27. The second cluster has 10 Typical files and 14 Atypical files. This cluster corresponds to the 24 files in the right portion of the graph in Figure 28.

A K-means cluster analysis was carried out on the correlation dimensions using BMDP-KM. This analysis was performed using 2 and 3 initial clusters. The results using 2 initial clusters are shown in Tables 6a, 6b, and 6c. Cluster 1 has 10 Atypical and 6 Typical files while cluster 2 has 5 Atypical and 23 Typical files. Typical files appear to be proportionately better grouped than Atypical files. The results for 3 initial clusters are shown in Tables 7a, 7b, and 7c. Cluster 1 has 11 cases, 7 Atypical and 4 Typical, cluster 2 has 16 cases, 7 Atypical, and 9 Typical, and cluster 3 has 17 cases, 1 Atypical and 16 Typical. Cluster 1 thus contains predominantly Atypical cases, cluster 3 contains mainly Typical cases, while cluster 2 contains both cases in approximately equal numbers.



TABLE 6a - K-Means Cluster Analysis; 2 Groups

Cluster	Size	Mean	Std. Dev.
1	16	2.147	0.208
2	28	1.493	0.217

TABLE 6b - Mean Squares; 2 Groups

Source	df	Mean Squares	F-ratio
Between	2	4.357	95.275
Within	41	0.046	

TABLE 6c - Classification by Cluster; 2 Groups

Cluster	1	2
Typical	6	23
Atypical	10	5

TABLE 7a - K-Means Cluster Analysis; 3 Groups

Cluster	Size	Mean	Std. Dev.
1	11	2.247	0.173
2	16	1.789	0.119
3	17	1.343	0.116

TABLE 7b - Mean Squares; 3 Groups

Source	df	Mean Squares	F-ratio
Between	2	2.773	155.5
Within	41	0.018	

TABLE 7c - Classification by Cluster; 3 Groups

Cluster	1	2	3
Typical	4	9	16
Atypical	7	7	1

#### 4.3.3 DISCUSSION

Using the channel of maximum peak-to-peak spike amplitude, there is a significant difference between the Typical and Atypical groups. This finding suggests that the differences between Typical and Atypical cases in the dynamical systems underlying the spike waveforms are easily detected in the recordings of a single channel. This channel is one which represents the maximum peak-to-peak amplitude of the spike waveform.

A possible explanation for the success of this channel selection criterion is that when the channel of maximum spike amplitude is used, the spike, and presumably the related surrounding electrical activity, are significantly larger in amplitude than other signals which may be present in the EEG record, and which are unrelated to the spike event. Such signals may be considered to be noise for the present purpose of analyzing the signal of interest, the spike phenomenon. The channel of maximum spike amplitude may then be considered to be the channel of maximum signal to noise ratio. By analyzing such a channel, it is the spike event, comprised of the spike and related surrounding activity, which is being preferentially analyzed. The contribution to this analysis of the noise components is effectively minimized.

The results of this experiment support the hypothesis that the generating mechanisms responsible for Atypical waveforms are more complex than those responsible for the

Typical waveforms, and that this differential complexity can be seen in the corresponding correlation dimension values.

Cluster analysis supports the notion that there are two clusters in the data. The Atypical files are best separated, with 14 of the 15 Atypical files being located in one of the clusters. Typical files are less well separated, with 19 Typical files in one cluster, and 10 Typical files in the other cluster.

It should be recalled that the original classification of these files as Typical and Atypical was made on the basis of clinical, neurological, and behavioural diagnostic criteria. The accuracy of these original classifications can only be verified by longitudinal studies of the patients involved. One interpretation of the cluster analysis results involves making the assumption that the correlation dimension analysis is inherently a more accurate criterion for classifying the files than the original diagnostic measures. On this assumption, the Atypical files were originally more accurately classified than the Typical files. The one anomalous Atypical file, according to the cluster analysis, should then in fact be considered to be a Typical file. Similarly, the 10 anomalous Typical files, according to the cluster analysis, should in the same way be considered to be Atypical files. To reiterate, however, only through the results of follow-up studies will it be possible to decide between the clinical and the correlation dimension based classifications.

The results of the K-means cluster analysis using 2 initial means show that Typical files are better grouped than Atypical files. Under the assumption that correlation dimension is a better predictor of BREC type than the clinical diagnostic measures, this finding would suggest that Typical files were originally better classified than Atypical files. Such a result however contradicts the results of the hierarchical cluster analysis. The contradiction may be the result of differences between the analytic procedure used by the two programs. No other explanation suggests itself.

The results of the K-means cluster analysis using 3 initial means similarly appear to show that Typical files are better grouped than Atypical files. The motivation for this particular analysis was to explore the possibility that there are 3 groups in the original data: one group correctly originally classified as Typical, one group correctly originally classified as Atypical, and one group that was incorrectly originally classified. The F-ratio for this analysis is higher than for the 2 initial group analysis. This would appear to suggest that the 44 files are more distinctly separable into these 3 groups, than into 2 groups. If this is the case, then it would appear to support the notion that there exists in the data a group of files that were originally incorrectly classified.

#### 4.4 EXPERIMENT 3.3 Maximum Peak/RMS Ratio Channel

##### 4.4.1 INTRODUCTION

An extension of experiment 3.2 would be to attempt to determine whether the criterion used in that experiment is in fact optimal. Related but different criteria could be tested to determine if effective discrimination depends solely on maximum spike amplitude. An example of such a related criterion might be the ratio of peak to peak amplitude, to the root-mean-square (RMS) amplitude of the waveform. The RMS amplitude is equivalent to the standard deviation of a time-series. This criterion will be tested in the present experiment.

A successful outcome, a significant between groups difference, in the present experiment would support the hypothesis that the most significant criterion for discriminating between the two groups is the amplitude of the spike itself, relative to the amplitude of the surrounding activity. Since the surrounding electrical activity contribute the majority of the data points in the spike waveform time-series, a large ratio of spike peak value to overall RMS amplitude suggests that the waveform consists of both a large spike and a low level of surrounding electrical activity.

An unsuccessful outcome of the present experiment would suggest that characteristics of the spike itself are not sufficient to discriminate between the Typical and Atypical groups. It would suggest that both the spike and the sur-

rounding activity are important when attempting to discriminate between these groups.

#### 4.4.2 METHOD

For each of the same 44 files that were used in experiment 3.1, the channel with the largest ratio of peak-to-peak amplitude to RMS was identified. For each channel within a data set, two values were calculated, the peak-to-peak amplitude and the RMS amplitude (standard deviation). Using these values, the ratio of the peak-to-peak amplitude to the RMS amplitude was calculated. The channel within a data set with the largest ratio was selected as the target channel for the calculation of the correlation dimension. The identification of the target channel was carried out using algorithm SELECT2, shown in section A5. All other conditions were identical to those in experiment 3.1.

#### 4.4.3 RESULTS

The results obtained in the analysis of the 44 BREC files is summarized in Table 8, and the analysis of the statistical significance of these results is shown in Table 9. The difference between the Typical and Atypical groups on correlation dimension is significant both in terms of the Student's t-test, and the point-biserial correlation.

TABLE 8 - Data Summary: Exp. 3.3 - Max. Peak/RMS Ratio  
Channel

=====  
Type 1 = typical; type 2 = atypical (prior diagnosis)  
=====

Case	Type	Lag	Channel	Dc
1	1	6	7	2.247
2	2	7	15	1.683
3	1	7	15	2.425
4	1	8	7	1.719
5	1	7	18	1.882
6	1	7	13	1.551
7	1	30	11	1.626
8	1	13	11	1.618
9	1	7	7	2.235
10	1	9	13	1.415
11	2	8	18	2.262
12	2	9	9	2.094
13	2	10	9	2.126
14	2	7	17	2.1
15	2	9	18	2.365
16	1	5	14	2.369
17	2	9	13	1.966
18	1	7	18	2.287
19	1	9	11	1.854
20	1	12	8	1.765
21	1	9	8	1.551
23	1	8	14	2.184
24	1	6	15	2.078
25	2	8	9	2.278
26	1	9	13	1.577
27	1	8	14	2.045
28	1	7	4	2.287
29	1	6	5	1.926
30	1	8	8	2.301
31	1	8	4	1.786
33	2	7	5	1.999
34	2	6	18	2.565
35	1	6	7	2.239
36	2	10	15	2.447
38	1	6	11	2.185
39	1	7	18	2.243
42	2	6	12	2.485
43	1	7	13	2.369
44	2	8	19	2.469
48	2	8	9	2.312
50	1	7	19	2.5
53	1	30	11	1.988
54	1	8	13	1.755
60	2	7	5	2.443

TABLE 9 - Statistics: Exp. 3.3 - Max. Peak/RMS Ratio Channel

Group	Mean	N
Typical	2.000	29
Atypical	2.240	15

Standard Error: 0.093

$t = -2.58$ ,  $df = 42$ ; ( $p < 0.05$ )  
 Point-biserial  $r = 0.370$  ( $p < 0.05$ )

#### 4.4.4 DISCUSSION

Although the results of this experiment indicate that there is a significant difference between the Typical and Atypical groups, the level of significance is lower than in experiment 3.2. The criterion used in experiment 3.2 appears to be more effective in discriminating between the two groups. On the one hand, the criterion used in the present experiment, the ratio of peak to RMS, should discriminate to some extent between the Typical and Atypical groups. By this criterion, the analyzed channel should contain a significant amount of spike activity. On the other hand, the selected channel will at the same time have a small value of RMS amplitude. In effect, the selected channel will have relatively less electrical activity both before and after the spike. It is in fact this electrical activity which would appear to contribute the additional complexity which then permits an effective discrimination to be made between the Typical and Atypical groups. The criterion of maximum peak to RMS ratio suppresses

the contribution of this electrical activity surrounding the spike. There appears to be support for the hypothesis that the Typical and Atypical groups can be best separated if both the characteristics of the spike and of the surrounding activity are analyzed.

#### 4.5 EXPERIMENT 3.4: Multichannel Reconstruction I

##### 4.5.1 INTRODUCTION

This approach to the reconstruction of a phase-space attractor is an alternative to the method of lags used in experiments 3.1, 3.2 and 3.3. In the method of lags, vectors defining the attractor are each composed of multiple points taken from the original single time-series, with an interval between the points equal to the lag parameter. In the present method, vectors defining the attractor each have as components multiples of points recorded simultaneously at several spatial locations. It has been conjectured (Eckmann and Ruelle, 1985) that this method will result in an adequate embedding of the multiple time-series in a phase-space. Dynamical properties of the resulting attractor, such as fractal dimensionality, will thus reflect the dynamics of the system generating the time-series.

Dvorak (1990) suggests that the multichannel reconstruction method avoids a source bias inherent in Taken's method of lags. Specifically, it is suggested that there is an upward bias in correlation dimension due to the finite resolu-

tion with which experimental data is typically recorded. This finite resolution leads to a rounding error which is effectively a random noise component that is combined with the signal of interest. The dimensionality of this random noise component increases proportionately with embedding dimension, producing an upward bias in correlation dimension for the digitized recording.

Destexhe, Sepulchre, and Babloyantz (1988) compared the multichannel reconstruction method, the method of lags, and the combination of singular value decomposition (SVD) and method of lags suggested by Broomhead and King (1986). Destexhe et al. (1988) used time-series associated with cardiac rhythms, Creutzfeld-Jakob disease, alpha waves, deep sleep, epileptic seizures, the Lorenz system of differential equations, and random noise. Destexhe et al. (1988) found that the method of lags, and Broomhead and King's (1986) combination method gave similar values for correlation dimension, with however the longer scaling region for the singular value decomposition method.

Palus et al. (1992) used the multichannel reconstruction technique with 16 EEG channels recording activity during the course of ethanol consumption in a human subject. Estimates of correlation dimension obtained agreed with a second variable which Palus et al. (1992) calculated, and which they termed linear complexity. Standard singular value decomposition of the data matrix yielded a set of significant

eigenvalues,  $\sigma_i$ . Linear complexity was defined as  $-N / \sum \log \sigma_i$ , where  $N$  was defined only as a normalization constant, but presumably could be equal to the number of significant eigenvalues used in the summation. Over the course of the experiment, changes in the value of linear complexity appeared to be similar to changes in the value of correlation dimension.

While the multichannel reconstruction method avoids the problem of choosing an appropriate lag parameter, according to Destexhe et al. (1988) an equivalent problem appears, choosing an inter-site distance for the recording of the multiple channels. The underlying goal here is the same as that for choosing the proper lag: to generate a series of points for the components of a vector that are essentially independent. A factor that can be both an advantage and a disadvantage appears with the multichannel reconstruction method: An advantage is that the multiple simultaneously recorded channels make use of more information that is in any case available in the typical multichannel EEG recording paradigm, than does the method of lags. This is an advantage when the system under investigation is spatially distributed, and a disadvantage when the target system is spatially localized. In this latter case, the computed value of correlation dimension will be less reflective of the dynamics of the target system.

Nan and Jinghua (1988) suggest that when segments of different dimensionality are analyzed together the value of cor-

relation dimension for the composite will approach that of the lowest value for the segments. The present situation is not entirely of the type referred to by Nan and Jinghua (1988) but it might be conjectured that in general when segments of differing dimensionality are analyzed together in any configuration, as contiguous segments, or as multiple channels in the multichannel reconstruction, the resulting value of correlation dimension may not accurately reflect the dynamics of any one segment in the ensemble.

Destexhe et al. (1988) further found that if the value of correlation dimension is less than 4, the three algorithms that they studied yielded similar results.

#### 4.5.2 METHOD

The ensemble of channels used were T3, C3, Cz, C4, and T4, lying approximately over the area of the rolandic fissure. These channels were selected on the basis of the observation that out of the 44 files that were examined, all showed some spike activity over the central and temporal regions. Further, 42 files showed significant spike activity over these areas. Synchronous data points from these 5 channels were used to construct a set of 256, 5-dimensional vectors, in the following way. The first time-point from each source channel, for a total of 5 time-points, became the first 5 sequential time-points of the output file. This process was repeated for the second time-point from each channel, and so on, up to the

last or 256th time-point of each source channel. The output file now consisted of 5 times 256 or 1280 time-points. The 5 source channels were essentially multiplexed onto a single output file.

Thus, each set of 5 sequential time-points in the output file defined a 5-component vector. This set of 256 vectors thus defined an attractor in a 5-dimensional phase-space. The correlation dimension for this attractor was computed using the CORDIM algorithm. Embedding dimension was 5, dictated by the use of 5 channels and thus 5 components to each phase-space vector. The value of lag was set at 1, owing to the method used to amalgamate the 5 channels.

#### 4.5.3 RESULTS

The results of the analysis are shown in Table 10, and the statistical calculations are shown in Table 11. The difference in correlation dimension between the Typical and Atypical groups of files was found to be not significant in terms of both the t-test and the point-biserial correlation.

TABLE 10 - Data Summary: Exp. 3.4 - Multichannel  
Reconstruction

=====  
Type 1 = typical; type 2 = atypical  
=====

Case	Type	Lag	Channel	Dc
1	1	1	13	3.186
2	2	2	10	2.755
3	1	2	15	3.091
4	1	2	13	3.073
5	1	1	13	3.175
6	1	2	13	3.097
7	1	2	9	2.972
8	1	2	9	3.329
9	1	2	13	2.854
10	1	2	13	3.352
11	2	2	12	3.036
12	2	2	9	2.632
13	2	2	9	2.92
14	2	2	5	1.893
15	2	30	12	2.81
16	1	30	9	3.34
17	2	2	12	3.095
18	1	1	18	3.374
19	1	2	13	2.614
20	1	2	13	2.752
21	1	1	9	2.718
23	1	2	10	2.77
24	1	2	9	2.88
25	2	2	9	3.354
26	1	1	13	3.263
27	1	1	9	3.107
28	1	2	14	2.471
29	1	1	15	2.948
30	1	26	13	2.772
31	1	2	13	3.05
33	2	30	1	2.677
34	2	30	18	3.445
35	1	2	12	3.439
36	2	1	10	3.278
38	1	2	13	3.058
39	1	2	12	3.102
42	2	2	12	3.569
43	1	2	13	3.467
44	2	2	9	2.998
48	2	1	12	2.989
50	1	2	13	2.897
53	1	2	9	2.931
54	1	2	13	3.111
60	2	2	3	2.603

TABLE 11 - Statistics: Exp. 3.4 - Multichannel Reconstruction

Group	Mean	N
Typical	3.04	29
Atypical	2.94	15
Standard Error: 0.100		

t = 1.04, df = 42; (p > 0.05)  
 Point-biserial r = 0.158 (p > 0.05)

#### 4.5.4 DISCUSSION

The non-significant result of this experiment are likely due to the particular set of channels chosen for the reconstruction. Examination of the data sets revealed that in all cases the focus of spike activity was clearly localized to one hemisphere or the other. In no case was there evidence of spike activity in both hemispheres in one file. Thus the particular ensemble of channels used in this analysis included, in the case of all files, channels with strong spike activity as well as channels with no spike activity. As Nan and Jinghua (1988) point out, analysis of data comprised of segments with evidence of dynamic activity and segments without evidence of such activity may result in a correlation dimension value equal to the minimum dimensionality of the comprising segments. The dimensionality of the dynamical system of interest is in a sense diluted by the presence of the unrelated time-segments. A follow-up to this experiment could make use of an ensemble of 4 or 5 channels centred on the central-temporal area of maximum spike amplitude, either

in the left or in the right hemisphere, rather than over both hemispheres as was done in the present experiment

#### 4.6 EXPERIMENT 3.5 - Multichannel Reconstruction II

##### 4.6.1 METHOD

In this experiment, the multichannel reconstruction was carried out using a set of 5 channels from either the left or the right temporal area. The selection of left or right was made on the basis of which side had the spike activity of greatest peak-to-peak amplitude.

For files with maximum activity over the left temporal area, the channels used were F3, T3, C3, T5, and P3. For files with maximum spike amplitude over the right temporal area, the channels were F4, C4, T4, T6, and P4. The processing of these channels was carried out in identically the same manner as for the rolandic channels in Experiment 3.4. The 5 source channels were transformed into a single output file, with each successive set of 5 data-points defining a 5-dimensional vector. Lag was set to 1 as in experiment 3.4.

##### 4.6.2 RESULTS

The results of the analysis are shown in Table 12, and the statistical calculations are shown in Table 13. The difference in correlation dimension between the Typical and Atypical groups of files was found to be not significant.

TABLE 12 - Data Summary: Exp. 3.5 - Multichannel  
Reconstruction

=====  
Type 1 = typical; type 2 = atypical  
Side 1 = left temporal area; side 2 = right temporal area.

Case	Type	Side	Dc
1	1	2	2.514
2	2	1	1.125
3	1	1	2.672
4	1	2	1.816
5	1	2	2.785
6	1	2	2.398
7	1	1	2.317
8	1	1	2.439
9	1	2	2.176
10	1	2	1.277
11	2	2	3.033
12	2	1	2.64
13	2	1	1.952
14	2	1	1.907
15	2	2	2.76
16	1	1	3.04
17	2	2	1.852
18	1	2	2.503
19	1	2	2.6
20	1	2	2.524
21	1	1	2.087
23	1	1	2.392
24	1	1	2.42
25	2	1	2.654
26	1	2	2.544
27	1	1	2.73
28	1	1	1.7
29	1	1	1.298
30	1	2	1.404
31	1	2	1.719
33	2	1	2.495
34	2	2	2.751
35	1	2	2.85
36	2	1	2.589
38	1	2	1.731
39	1	2	1.645
42	2	2	2.62
43	1	2	2.924
44	2	1	2.559
48	2	2	2.247
50	1	2	2.797
53	1	1	1.678
54	1	2	2.074
60	2	2	2.475

TABLE 13 - Statistics: Exp. 3.5 - Multichannel Reconstruction

Group	Mean	N
Typical	2.24	29
Atypical	2.38	15
Standard Error:	0.160	

$t = -0.84$ ,  $df = 42$ ; ( $p > 0.05$ )  
 Point-biserial  $r = 0.128$  ( $p > 0.05$ )

#### 4.6.3 DISCUSSION

In general terms the non-significant results are probably due once again to a selection of channels for the analysis that do not equally contribute signals related to the dynamical system underlying the spike waveforms.

For the multichannel reconstruction method of analysis to succeed in the present application of differentiating between Typical and Atypical BREC generator systems, it is likely necessary that all channels used in the analysis be closely coupled throughout the time-evolution of the dynamical system of interest (Gershenfeld, 1988). Where this condition is not satisfied, and some channels are not closely involved in the spike generating mechanism, the correlation dimension value cannot accurately reflect the dynamics of a single target system. When the electrical activity recorded at the several channels reflects loosely-coupled or autonomous activity of several dynamical systems, correlation dimension will reflect the number of such dynamical systems in operation, rather than the level of complexity of any one of these systems. Gershen-

feld (1988) points out that a significant goal of dimensional analysis is the discrimination between an n-dimensional system and n 1-dimensional systems.

#### 4.7 GENERAL DISCUSSION

No significant difference between the Typical and Atypical groups was found when the same channel, C4, or the same set of channels, T3, C3, Cz, C4, T4, was used for all cases. To reiterate, an implication of these findings is that the differences between the 44 Typical and Atypical cases studied in this work are not localized to a region covered either by a single common channel, or strictly by a single common set of channels. No significant difference was found when the analysis utilized a group of channels which differed between cases in terms of laterality, according to the side of the scalp which contained the channel of maximum spike amplitude. These channels surrounded the channel of maximum spike amplitude.

A possible implication of these findings is that the differences between Typical and Atypical systems within a single file are not distributed as widely as the area covered by the sets of channels used in the analysis. Alternatively, the selection criterion used to select channels might not be appropriate. Finding a better criterion for selecting such channels is a problem which remains to be explored.

A significant difference was found in two cases. The

most significant difference was found when, for each data set, a single channel corresponding to maximum spike amplitude was analyzed. A lesser, but still significant difference was found when the analyzed channel was selected on the basis of maximum ratio of peak-to-peak amplitude to RMS amplitude. This lesser difference suggests that both the spike and the electrical activity before and after the spike are together important in effectively discriminating between the Typical and Atypical groups.

Correlation dimension analysis of a single channel inherently gives information about the temporal distribution of the underlying dynamical systems. The analysis provides an estimate of the lower bound on the number of variables involved in the underlying dynamical system, over the time-course of the recorded time-series. Thus the results of experiments 3.2 and 3.3 appear to support the conclusion that Atypical cases are associated with dynamical systems of greater complexity, than Typical cases, involving a greater number of variables, over the time-course of the spike waveform.

This conclusion is consistent with a model of Typical and Atypical BREC by Wong (1989), proposed on the basis of results from dipole localization studies. According to Wong's (1989) model, both Typical and Atypical BREC involve a common neural generator site, while the Atypical BREC case involves additional surrounding neural tissue that interacts with the

common generator locus. Such interaction, Wong (1989) suggests occurs mainly in the interval after the main spike. The correlation dimension is well suited to tracking such a difference in the time-evolution between the Typical and Atypical cases, since it is sensitive to the changing pattern of activity throughout the 1.28 second epoch surrounding the spike maximum. Were this pattern of activity to reflect the operation of only a small number of variables, then correlation dimension would be correspondingly low. If the pattern of activity reflected the successive engagement of a greater number of variables, then correlation dimension would be correspondingly higher.

In terms of spike topography, the differences between Typical and Atypical cases was apparently well represented by the channel of maximum amplitude, implying that these differences may be localized to an area represented by this channel. This implication is supported by the non-significant difference found when a wider area surrounding the maximum amplitude channel was analyzed.

It may be noted that correlation dimension analysis applied to a single EEG time-series inherently looks at a pattern of behaviour over time, at a single location. In contrast, dipole analysis looks at spatial distribution of EEG behaviour at a single point in time. These methods might therefore be used together to answer questions about both the spatial and temporal distribution of EEG sources.

In the present analysis, only differences between the Typical and Atypical groups are obtained, rather than information about particular cases. It cannot be concluded therefore that the system generating a spike is spatially localized to an area represented by a single channel. It would seem to be a reasonable hypothesis however that the differences between the Typical and Atypical groups are spatially localized to an area which is of the approximate order of that represented by a single electrode.

Taken together, the methods of analysis used in experiments 3.1 to 3.5 thus provide some information about the spatial as well as the temporal distribution of the differences between the Typical and Atypical groups.

A variable which was not controlled for in this study was the number of records that were averaged to form the data files as analyzed. This averaging was done prior to the files becoming available to this study. Between 6 and 20 separate spike waveform recordings were averaged to generate the data files. Such differences in averaging have been shown to have an effect on the value of correlation dimension. Badii, Broggi, Derighetti, Ravani, Ciliberto, Politi, and Rubio (1988) found that correlation dimension increases with averaging. Intuitively, the averaging combines or more closely couples variables from the different time-series that enter into the average, thus increasing the dimensionality of the resulting averaged time-series.

Another variable which may be a source of ambiguity is the filter settings used for the original recordings of the spike waveforms. These settings were specified by the source of the data (B. C. Children's Hospital) as 1 Hz high-pass and 70 Hz low-pass. However, examination of the spectrum of several files selected at random revealed that in all cases the upper frequency point (the frequency at which the power spectrum decreased to one-half of the mean power level at intermediate frequencies) did not exceed approximately 30 Hz. Further, this upper frequency point varied from file to file, from approximately 17 Hz to 30 Hz. This observation may be the result of the differential averaging of the data files, or it may reflect unequal filter settings used in the original recordings. This ambiguity cannot be resolved without certain knowledge about the actual filter settings that were used for the EEG recordings.

For reliable estimates of correlation dimension, parameters such as the number of averaged records, and all filter settings, should be kept constant across files which are to be compared. If the correlation dimension technique were to be used as a diagnostic criterion, it is strongly recommended that a protocol be established which includes a definition of the value of these parameters for all recordings.

#### RECOMMENDED RECORDING PROTOCOL

The following is a suggested protocol for EEG recordings

for use with correlation dimension analysis:

- filter settings of 1 Hz high-pass and 70 Hz low-pass. The actual value of the upper cutoff frequency is, within limits, perhaps less important than that this frequency be constant across files which will be subsequently compared on their value of correlation dimension. A value of 70 Hz is reasonable in that this is a commonly used value in EEG recording.

- no averaging of the spike-waves. It is a conjecture based on intuition developed during the course of this study that averaging may minimize the effects of the different variables involved in generating the spike-wave waveform. If this were the case it might cause the correlation dimension values to converge artificially to some value not reflecting the contributions of all the system variables. It is recommended therefore that an epoch of the EEG recording be selected that is free of clearly identifiable artifacts such as movement and myoelectric artifacts.

- the selected EEG epoch should contain a number of consecutive spike-wave events. Such events are in fact known to commonly occur in such bursts (eg. Beaussart, 1972). It is therefore recommended that the EEG epoch selected for analysis contain a number of closely-spaced spike-wave events. In analyzing a number of spike-wave events, rather than just a single event as in the present study, the correlation dimension algorithm should yield a longer scaling region, that is,

a longer straight-line segment of the graph of  $\ln C(r)$  vs.  $\ln r$ . This longer scaling region will then make it possible to get a more reliable estimate of slope, the correlation dimension.

- the time-interval of the selected EEG epoch should be set to 4 seconds. The approximate frequency of the spike-wave events is 3 Hz. The 4 second interval would contain on the order of 10 spike-wave events, if the train of spike-waves is selected to contain minimal intervening time-gaps. In principle the analytical technique would benefit from having an even larger number of spike-wave events. The difficulty with attempting to analyze longer epochs is the non-stationary characteristic of the EEG. Statistical features of the EEG ever-increasingly diverge as longer epochs are considered. Again, on the basis of experience gained in this work it is felt that a time-interval of 4 seconds might represent a not unreasonable trade-off between a sufficient quantity of data and non-stationarity.

- the rate at which the EEG signal is sampled should be set at 200 samples per second, the rate used for the files used in the present work. A small value of sampling rate reduces the number of data points available to the analysis, while a large sampling rate incurs the penalty of increasing dependence between adjacent data points; essentially, as sampling rate is increased, adjacent data points contain little new information. A sampling rate of 200 samples per

second is probably a not unreasonable compromise. For a 4 second interval, this represents a total of 800 data points, which would seem to be a reasonable choice for the present comparative application of correlation dimension analysis.

- a general consideration would be that EEG recordings be made within similar patient conditions. Since the BREC spike-wave episodes commonly occur during sleep, it is recommended that EEG recordings for this analysis be made during sleep only, and if possible during similar sleep stages. If spike-waves occur during slow-wave sleep, this would be the preferred time for recording. In this way, it may be possible to minimize the influence of neural variables unrelated to the BREC spike-wave generation, and thus minimize their effect on correlation dimension.

#### SUMMARY

The results of the present correlation dimension analysis appear to agree with the clinically-based classification of BREC into the Typical and Atypical forms. Correlation dimension may thus be a useful diagnostic variable in classifying BREC type. The advantage of using correlation dimension is that a computer algorithm exists to process the spike-wave waveforms. It is suggested that if a consistent recording protocol were to be adopted for recording the spike-wave events, an even more significant difference between the Typical and Atypical BREC types would be found.

## APPENDIX I - Algorithms

A1: CORDIM - This algorithm calculates the correlation integral, over the selected values of scale length, and over the selected range of embedding dimensions.

Parameters:

y()            array holding original time-series  
length        length of time-series  
lag            value of lag  
s              selected sample of points on attractor  
mmin          minimum embedding dimension  
mmax          maximum embedding dimension  
rmin          minimum r value  
rmax          maximum r value  
C()            array holding values of correlation integral, C(r)

CORDIM (y(), length, lag, s, mmin, mmax, rmin, rmax, C())

Calculate value of LAG:

CALL LAGFIND (y(), lag, length)

FOR m = mmin TO mmax

  Compute maximum usable length of time-series:

  Nmax = length - ((m - 1) \* lag)

  Compute number of points averaged over:

  Navg = Nmax / s

  Compute all distances between points:

  CLEAR s()

  FOR i = 1 TO Navg

    FOR j = 1 TO Nmax

      FOR k = 0 TO (m-1)

        s(i,j) = s(i,j) + |y(i\*s+k\*lag) - y(j+k\*lag)|

      NEXT

    NEXT

  NEXT

  Compute C(r) for all scale lengths r:

  FOR r = rmin TO rmax

    sum = 0

    FOR i = 1 TO Navg

      FOR j = 1 TO Nmax

        IF s(i,j) < r THEN sum = sum + 1

      NEXT

    NEXT

    C(r) = sum / (Navg \* Nmax)

  NEXT

NEXT

A2: LAGFIND - This algorithm calculates the value of LAG.

Parameters

y()            the array storing the time-series  
length        length of the time-series  
lag            lag parameter.

LAGFIND (y(), lag, length)

Initialize the value of lag:

lag = 0

DO

  Increment the value of lag:

  lag = lag + 1

  Calculate y' = mean value of points along y1 axis:

  sum = 0

  FOR i = 1 TO length

    sum = sum + y(i)

  NEXT

  y' = sum / length

  Calculate ss1 = sum of squared deviations along axis r1:

  SS1 = 0

  FOR i = 1 TO length - lag

    SS1 = SS1 + (y(i + lag) - y(i))<sup>2</sup>

  NEXT

  Calculate ss2 = sum of squared deviations along axis r2:

  SS2 = 0

  FOR i = 1 TO length - lag

    SS2 = SS2 + (y(i + lag) + y(i) - 2y')<sup>2</sup>

  NEXT

LOOP UNTIL SS1 / SS2 > .8 OR lag = 30

A3: REGRESS - This algorithm calculates the slope of the regression line tangent to a selected segment of the graph of correlation integral vs. scaling length.

Parameters:

rmin            minimum value of ln r  
rmax            maximum value of ln r  
nreg            no. of points for regression calculation  
nsteps          total no. of data points  
y()            array containing data points  
dr              increment size for ln r  
auto            selects manual or automatic segment selection  
slope           slope of computed regression line

errmin      standard error of regression for regression line  
REGRESS (rmin, rmax, nreg, nsteps, y(), dr, auto, slope,  
errmin)

Set minimum regression error level to arbitrary high value:  
errmin = 100

Compute slope and regression error; repeat incrementing starting value rlo if optimum segment selection used (auto = true):

DO

    Initialize variables:

    j = 0: Sx = 0: Sy = 0: Sxx = 0: Syy = 0: Sxy = 0

    Compute end-points of regression line:

    rlo = rmin + dr

    rhi = rlo + dr \* nreg

    Compute slope and regression error:

    FOR k = 1 TO nsteps

        x = y(k, 1)

        y = y(k, 2)

        IF x >= rlo AND x <= rhi THEN

            j = j + 1

            Sy = Sy + y

            Sx = Sx + x

            Sxy = Sxy + x \* y

            Sxx = Sxx + x<sup>2</sup>

            Syy = Syy + y<sup>2</sup>

        END IF

    NEXT

    COVxy = Sxy - (Sy \* Sx) / j

    VARx = Sxx - (Sx \* Sx) / j

    VARY = Syy - (Sy \* Sy) / j

    rcoef = COVxy / VARx

    SSreg = rcoef \* COVxy

    regerror = SQR((VARY - SSreg) / (j - 2))

    Check if slope and error are within bounds:

    IF regerror < errmin AND rcoef > .4 THEN

        errmin = regerror

        slope = rcoef

    END IF

LOOP UNTIL rhi = rmax OR NOT auto

A4: SELECT1 - This algorithm finds the channel within a data file with the largest peak-to-peak amplitude.

Parameters:

y()                    array holding EEG data set  
 m                     no. of EEG channels, 21  
 length                length of each channel in data points, 256  
 channel                channel with maximum peak to peak amplitude

SELECT1 (y(), m, length, channel)

Find the channel with the largest difference between minimum and maximum amplitude:

```

ymax = 0
FOR j = 1 TO m
  yhi = 1E-12: ylo = 1E+12
  FOR i = 0 TO length - 1
    IF y(i, j) > yhi THEN yhi = y(i, j)
    IF y(i, j) < ylo THEN ylo = y(i, j)
  NEXT
  yamp = yhi - ylo
  IF yamp > ymax THEN ymax = yamp: channel = j
NEXT

```

A5: SELECT2 - This algorithm finds the channel within a data file with the maximum peak/RMS ratio.

Parameters:

y()                    array holding EEG data set  
 m                     no. of EEG channels, 21  
 length                length of each channel in data points, 256  
 channel                channel with maximum peak/RMS ratio

SELECT2 (y(), m, length, channel)

```

ratiomax = 0
FOR j = 1 TO m
  yhi = 1E-12: ylo = 1E+12: sum = 0
  FOR i = 0 TO length - 1
    IF y(i, j) > yhi THEN yhi = y(i, j)
    IF y(i, j) < ylo THEN ylo = y(i, j)
    sum = sum + y(i, j)
  NEXT
  peak = yhi - ylo
  mean = sum / length
  ss = 0
  FOR i = 0 TO length - 1
    ss = ss + (y(i, j) - mean) ^ 2
  NEXT
  rms = SQR(ss / length)
  ratio = peak / rms
  IF ratio > ratiomax THEN channel = j
NEXT

```

## BIBLIOGRAPHY

- Aicardi, J. and Chevrie, J. J. (1982). Atypical benign partial epilepsy of childhood. Developmental Medicine and Child Neurology, 24, 281-292.
- Albano, A. M., Muench, J., Schwartz, C., Mees, A. I., and Rapp, P. E. (1988). Singular-value decomposition and the Grassberger-Procaccia algorithm. Physics Review A, 38, 3017-3026.
- Albrecht, V. and Palus, M. (1991). The spectral dynamics and its applications in EEG. Biological cybernetics, 66 (1), 71.
- Arle, J. E. and Simon, R. H. (1990). An application of fractal dimension to the detection of transients in the electroencephalogram. Electroencephalography and Clinical Neurophysiology, 75 (4), 296.
- Babloyantz, A. and Destexhe, A. (1986). Low-dimensional chaos in an instance of epilepsy. Proceedings of the National Academy of Science, 83, 3513-3517.
- Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., and Rubio, M. A. (1988). Dimension increase in filtered chaotic signals. Physical Review Letters, 60 (11), 979-982.
- Basar, E., Basar-Eroglu, J., Roschke, J., and Schult, J. (1990). Strange attractor EEG as sign of cognitive function. In E. Roy John (Ed.), Machinery of the Mind, New York: Birkhauser Press.
- Beaussart, M. (1972). Benign epilepsy of children with rolandic (centro-temporal) paroxysmal foci. Epilepsia, 13, 795-811.
- Bencivenga, R. (1987). A Statistical Analysis of Electroencephalographic Spikes in Benign Rolandic Epilepsy of Childhood. M.Sc. Thesis, Department of Statistics, The University of British Columbia.
- Berliner, L. M. (1992). Statistics, probability and chaos. Statistical Science, 7 (1), 69-90.
- Beydoun, A., Garofalo, E. A. and Drury, I. (1989). Comparison between typical and atypical cases of benign rolandic epilepsy. Epilepsia, 30, 636-739.
- Blom, S., Heijbel, J., and Bergfors, P. G. (1972). Benign epilepsy of children with centro-temporal EEG foci.

- Prevalence and follow-up study of 40 patients. Epilepsia, 13, 609-619.
- Broomhead, D. S. and King, G. P. (1986). Extracting qualitative dynamics from experimental data. Physica 20D, 217-236.
- Casdagli, M., Eubank, S., Farmer, J. D. (1991). State-space reconstruction in the presence of noise. Physica 51D, 352-359.
- Chatterjee, S. and Yilmaz, M. R. (1992). Use of estimated fractal dimension in model identification for time series. Journal of Statistical Computation and Simulation, 41 (3/4), 129-141.
- Chatterjee, S. and Yilmaz, M. R. (1992). Chaos, fractals and statistics. Statistical Science, 7 (1), 49-68.
- DeCoster, G. P. and Mitchell, D. W. (1991). The efficacy of the correlation dimension technique in detecting determinism in small samples. Journal of Statistical Computation and Simulation, 39 (4), 221-229.
- Destexhe, A., Sepulchre, J. A., and Babloyantz, A. (1988). A comparative study of the experimental quantification of deterministic chaos. Physics Letters A, 132, 101-106.
- Ding, M., Grebogi, C., and Ott, E. (1989). Dimensions of strange nonchaotic attractors. Physics Letters A, 137 (4/5), 167-172.
- Dreifuss, F. E. (1983). Pediatric Epileptology: Classification and Management of Seizures in the Child. Boston: J. Wright.
- Dvorak, I. and Siska, J. (1986). On some problems encountered in the estimation of the correlation dimension of the EEG. Physics Letters A, 118 (2), 63-66.
- Dvorak, I. (1990). Takens versus multichannel reconstruction in EEG correlation exponent estimates. Physics Letters A, 151 (5), 225-233.
- Eckmann, J. P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. Review of Modern Physics, 57, 617-656.
- Essex, C. and Nerenberg, M. A. H. (1990). Correlation dimension and systematic geometric effects. Physics Review A, 42, 7065-7074.

- Essex, C. and Nerenberg, M. A. H. (1991). Comments on 'Deterministic chaos: the science and the fiction' by D. Ruelle. Proceedings of the Royal Society of London A, 435, 287-292.
- Farmer, J. D., Ott, E., and Yorke, J. A. (1983). The dimension of chaotic attractors. Physica 7D, 153-180.
- Ford, J. (1987). Directions in classical chaos. In Hao Bai-lin (Ed.), Directions in Chaos. Singapore: World Scientific.
- Frank, G. W., Lookman, T., Nerenberg, M. A. H., Essex, C., Lemieux, J., and Blume, W. (1990). Chaotic time-series analyses of epileptic seizures. Physica 46D, 427-438.
- Fraser, A. M. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. Physical Review A, 33 (2), 1134.
- Freeman, W. J. (1987). Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. Biological Cybernetics, 56, 139-150.
- Freeman, W. J. (1988). Strange attractors that govern mammalian brain dynamics shown by trajectories of electroencephalographic (EEG) potential. IEEE Transactions on Circuits and Systems, 35 (7), 780-783.
- Heijbel, J., Blom, S., and Bergfors, P. G. (1975). Benign epilepsy of children with centrotemporal EEG foci. A study of incidence rate in outpatient care. Epilepsia, 16, 657-664.
- Gallez, D. and Babloyantz, A. (1991). Predictability of human EEG: a dynamical approach. Biological Cybernetics, 4 (5), 381-391.
- Geladze, T. S., Toidze, O. S., and Lomashvili, N. D. (1983). Significance of the early diagnosis of benign childhood epilepsy with rolandic peaks. Zhurnal Nevropatologii i Psikhiatrii, 83(10), 1492-1496.
- Gershenfeld, N. (1988). An experimentalist's introduction to the observation of dynamical systems. In Hao Bai-lin (Ed.), Directions in Chaos, Vol. 2. Singapore: World Scientific.
- Gibson, J. F., Farmer, J. D., Casdagli, M., and Eubank, S. (1992). An analytic approach to practical state-space reconstruction. Physica 57D, 1-30.

- Grassberger, P. and Procaccia, I. (1983a). Characterization of strange attractors. Physical Review Letters, 50 (5), 346-349.
- Grassberger, P. and Procaccia, I. (1983b). Measuring the strangeness of strange attractors. Physica 9D, 189-208.
- Grassberger, P. (1986). Do climatic attractors exist. Nature, 323, 609-612.
- Grebogi, C., Ott, E., Pelikan, S., and Yorke, J. A. (1984). Strange attractors that are not chaotic. Physica 13D, 261-268.
- Gregory, D. and Wong, P. K. H. (1984). Topographic analysis of the centrotemporal discharges in benign rolandic epilepsy of childhood. Epilepsia, 25(6), 705-711.
- Hobbs, J. (1991). Chaos And indeterminism. Canadian Journal of Philosophy, 21 (2), 141-164.
- Jackson, E. A. (1990). Perspectives of Nonlinear Dynamics, Vol. 2. Cambridge: Cambridge University Press.
- Lefranc, M., Hennequin, D., and Glorieux, P. (1992). Improved correlation dimension estimates through changes of variable. Physics Letters A, 163 (4), 269-274.
- Leibert, W. and Schuster, H. G. (1989). Proper choice of the time-delay for the analysis of chaotic time series. Physics Letters A, 142, 107-112.
- Martinerie, J. M., Albano, A. M., Mees, A. I., and Rapp, P. E. (1992). Mutual information, strange attractors, and the optimal estimation of dimension. Physical Review A, 45 (10), 7058-7064.
- Mayer-Kress, G. and Layne, S. P. (1987). Dimensionality of the human electroencephalogram. In S. H. Koslow (Ed.), Perspectives in Biological Dynamics and Theoretical Medicine. Annals of the New York Academy of Sciences, 54, 62-87.
- Moon, F. C. (1987). Chaotic Vibrations: An introduction for applied scientists and engineers. New York: John Wiley and Sons Ltd.
- Nan, X. and Jinghua, X. (1988). The fractal dimension of EEG as a physical measure of conscious human brain activities. Bulletin of Mathematical Biology, 50, 559-565.
- O'Donohoe, N. V. (1985). Epilepsies of Childhood, 2nd edit-

ion. London: Butterworths.

- Osborne, A. R. and Provenzale, A. (1989). Finite correlation dimension for stochastic systems with power-law spectra. Physica D, 35 (3), 357-381.
- Palus, M., Dvorak, I., and David, I. (1992). Spatio-temporal dynamics of human EEG. Physica A, 185 (1/4), 433.
- Palus, M. and Dvorak, I. (1992). Singular value decomposition in attractor reconstruction: Pitfalls and precautions. Physica D, 55 (1/2), 221.
- Pijn, J. P., Van Neerven, J., Noest, A., and da Silva, F. H. (1991). Chaos or noise in EEG signals; dependence on state and brain site. Electroencephalography and Clinical Neurophysiology, 79 (5), 371-381.
- Pritchard, W. S. and Duke, D. W. (1992). Dimensional analysis of no-task human EEG using the Grassberger-Procaccia method. Psychophysiology, 29 (2), 182-192.
- Ramsey, J. B., Sayers, C. L., and Rothman, P. (1990). The statistical properties of dimension calculations using small data sets: some economic applications. International Economic Review, 31 (4), 991-1020.
- Rapp, P. E., Bashore, T. R., Martinerie, J. M., Albano, A. M., Zimmerman, I. D., and Mees, A. I. (1989). Dynamics of brain electrical activity. Brain Topography, 2 (1/2), 99-118.
- Romeiras, F. J., Bondeson, A., Ott, E., Antonsen, T. M., and Grebogi, C. (1987). Quasiperiodically forced dynamical systems with strange nonchaotic attractors. Physica 26D, 277-294.
- Roschke, J. and Aldenhoff, J. B. (1992). A Nonlinear Approach to Brain Function: Deterministic Chaos and Sleep EEG. Sleep, 15 (2), 95.
- Skarda, C. A. and Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. Behavioural and Brain Sciences, 10, 161-195.
- Smith, L. A. (1988). Intrinsic limits on dimension calculations. Physics Letters A, 133, 283-288.
- Smith, R. L. (1992). Comment: Relation between statistics and chaos. Statistical Science, 7 (1), 109-113.
- Soong, A. C. K. and Stuart, C. (1989). Evidence of chaotic

- dynamics underlying the human alpha-rhythm electroencephalogram. Biological Cybernetics, 62, 55-62.
- Takens, F. (1980). Detecting strange attractors in turbulence. Dynamical Systems and Turbulence. Lecture Notes in Mathematics, 898, 366-381.
- Tsay, R. S. (1992). Comment: Relation between statistics and chaos. Statistical Science, 7 (1), 113-114.
- Verity, C. M. (1988). When to start anticonvulsant treatment in childhood epilepsy: The case for early treatment. British Medical Journal, 297, 1528
- Watt, R. C. and Hameroff, S. R. (1987). Phase space analysis of human EEG during general anesthesia. In S. H. Koslow (Ed.), Perspectives in biological dynamics and theoretical medicine. New York: New York Academy of Sciences.
- Wong, P. K. H. (1989). Stability of source estimates in rolandic spikes. Brain Topography, 2 (1/2), 31-36.
- Wong, P., Bencivenga, R., and Gregory, D. (1988). Statistical classification of spikes in benign rolandic epilepsy. Brain Topography, 1(2), 123-129.
- Wong, P. K. H., Gregory, D., and Farrell, K. (1985). Comparison of spike topography in Typical and Atypical benign rolandic epilepsy of childhood. Electroencephalographic and Clinical Neurophysiology, 61, S47.