Chapter 5

Behavioral Chaos: Beyond the Metaphor

Steven S. Robertson, Avis H. Cohen, and Gottfried Mayer-Kress

The field of dynamics is not new (Abraham and Shaw, 1982). Applied dynamics, which concerns itself with natural phenomena and their time-dependence, dates back at least to Johannes Kepler, who studied planetary motions near the beginning of the seventeenth century. This particular problem is highly nonlinear, and has since been shown to lack a general solution even for a simple three-planet system. Mathematical dynamics, beginning with Isaac Newton in the 1660s, includes the theory of differential equations which formally embodies the concepts of change. More recently, especially as a result of the work of Henri Poincaré, geometric methods in which the time-dependent changes in a dynamic system are conceptualized as motion on (sometimes complex) surfaces has received increased attention. Experimental dynamics, dating back to Galileo at the beginning of the seventeenth century, explores the properties of dynamic systems with mechanical and, today, computer models. In fact, it is the power of computers which promises to unlock many of the fundamental secrets of nonlinear dynamics to modern experimentalists (e.g., Lorenz, 1963).

Dynamic systems theory has recently caught the conceptual fancy of scientists trying to understand behavioral and psychological organization and development. Why is that? It seems there are a number of reasons. At the core is the fact that the organisms we seek to understand *are* dynamic systems. They exhibit time-dependent changes on scales spanning at least ten orders of magnitude, from fractions of a second to decades. Much of our work is an attempt to describe, model, and predict this activity, including the factors that control it. Above all, we are interested in organization and pattern, whatever our level of analysis or temporal scale. We study these complex systems and try to discover where the organization comes from, what is responsible for its stability and instability, and how structural reorganizations come about. We also share the belief that, to some extent, there is an underlying lawfulness to the phenomena we study. These are, of course, the concepts and aims of general theories of dynamic systems.

There is a danger, however, in this new fascination with dynamic theory. The danger lies in the temptation to naively adopt a new terminology or set of metaphors, to merely redescribe the phenomena we have been studying for so long, and then conclude that we have explained them. Because dynamic concepts and theory are seductive, we may mistake translation for explanation. As a first step in going beyond mere redescription, we should apply the theory to particular phenomena and pursue the application in as rigorous a fashion as we can, without trying to sidestep the theory's mathematical foundations. But the application should not be blind. We must demonstrate at each step that the phenomenon exhibits the characteristics assumed or predicted by the theory. In this process we must acknowledge the constraints imposed

In L.B. Smith and E. Thelen (Eds.), A Dynamic Systems Approach to Development: Applications. Cambridge, MA: MIT Press, 1993.

by our level of analysis (e.g., overt behavior) and by limitations in the quality and quantity of our measurements. Ultimately, the goal should be to build a dynamic model of the mechanism controlling the observed phenomenon from which new observable properties can be deduced and subjected to experimental test.

The purpose of this chapter is to take a particular phenomenon—the cyclicity in spontaneous motor activity in the human neonate—and examine it closely from the theoretical framework of dynamic systems. We ask very specific questions of both a qualitative and quantitative sort, illustrate the analyses required to address those questions, and critically evaluate the results from both computational and psychobiological perspectives.

Cyclic Motor Activity

The specific behavioral phenomenon we will analyze is the cyclic fluctuation in spontaneous motor activity (CM), comprised of general movements of the limbs, trunk, and head, and more isolated or stereotyped movements, with a cycle time on the order of a few minutes or less. The cyclicity has been observed in a variety of species (Corner, 1977; Hamburger, Balaban, Oppenheim, and Wenger, 1965), including the human neonate (e.g., Robertson, 1982) (figure 5.1). Some properties of CM in the human infant, such as the dominant frequency of oscillation and the proportion of movement variance accounted for by the oscillation may be relatively independent of the average level of activity (Robertson, 1987).

Development of Cyclic Motor Activity

In the human, CM emerges at least by midgestation (Robertson, 1985). It is possible that the mechanism responsible for CM is operating much earlier, perhaps in the first trimester. Suggestive evidence for this comes from ultrasound observations of normal fetuses in which burst-pause organization is noted around the 12th postmenstrual week (deVries, Visser, and Prechtl, 1982). The clear burst-pause organization, which depends (by definition) on the existence of periods of total motor quiescence, quickly disappears and is replaced by longer periods of more continuous but fluctuating activity (deVries et al., 1982). An important unanswered question concerning the ontogeny of CM is whether the early burst-pause organization is generated by the same mechanism responsible for the oscillations in spontaneous activity which have been studied

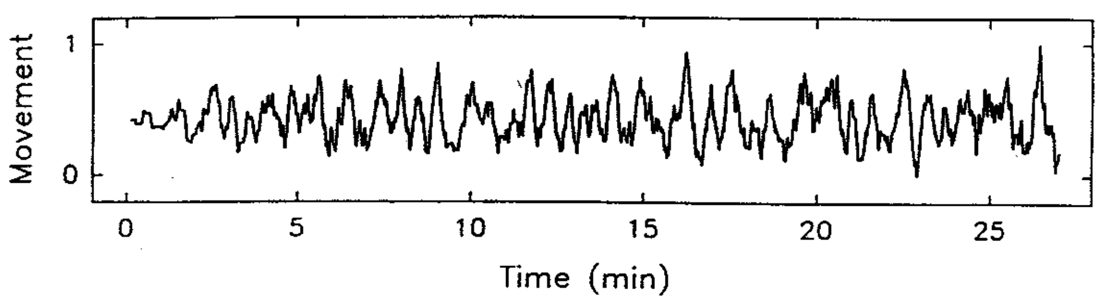

Figure 5.1
Movement time series from an awake infant studied 2 months after birth. After detrending, the data were rescaled to the interval [0, 1]. In this illustration, the time series has been smoothed with a 15-second moving average.

in the second half of gestation. During the second half of gestation, human CM changes very little in terms of properties that can be quantified using spectral analysis (Robertson, 1985). The frequency and relative strength of the dominant motility cycle, and the dispersion of movement variance around the dominant cycle in the frequency domain, all remain relatively stable during a time when other aspects of neurobehavioral organization (e.g., sleep and wake states) undergo dramatic changes.

Some evidence regarding factors that can alter CM comes from studies of diabetic mothers. Fetal CM is disrupted by maternal diabetes early in the third trimester (Robertson and Dierker, 1986) and perhaps earlier, although this has not been investigated. The oscillations in spontaneous activity are a bit slower, and the differences are related to the control of the mother's diabetes. It is not known, however, whether the effects are acute in nature, reflecting short-term alterations in maternal metabolism, or whether they are manifestations of a different developmental trajectory related to very early disruptions in fetal neural organization caused by maternal diabetes. In either case, CM appears normal by the end of gestation, suggesting that the underlying mechanism is developmentally buffered from the earlier effects of maternal diabetes.

Birth, in itself, does not appear to induce any changes in CM. The cyclic organization of spontaneous newborn activity measured during active sleep is quantitatively similar to fetal CM in the last month of gestation (Robertson, 1987). These findings are consistent with other evidence of neurobehavioral continuity across birth (Prechtl, 1984). Similar results were obtained from newborn infants of diabetic mothers, including those with clinical evidence of exposure to an altered metabolic environment in utero (Robertson, 1988). Furthermore, the characteristic cyclicity is observed during all sleep and wake states, and the quantitative properties of CM estimated by spectral analysis (frequency, strength, etc.) are similar across wake states, which vary enormously in the average level of activity (Robertson, 1987, 1988).

The functional significance of CM is unknown. However, the consequences of movement and the underlying neural and neuromuscular activity on the development of muscles, bones, joints, and neural connectivity have been well documented (e.g., Drachman and Sokoloff, 1966; Harris, 1981). There is also some evidence that periodic activation, compared to continuous stimulation at the same average rate, may facilitate normal synapse elimination in neonatal rat muscle (Thompson, 1983). It has been speculated that cyclicity per se may have beneficial consequences by balancing the benefits of activity and inactivity (Robertson, 1989). It has also been speculated that CM may regulate interactions between the young infant and its physical and social environment, either directly at the behavioral level, or indirectly through the central modulation of perceptual and cognitive processes (Robertson, 1989). In any case, as we discuss below, a good dynamic model of CM might, through its structure or its parameters, suggest ways in which to begin studying the functional significance of CM in a systematic manner.

Characteristic Properties of Cyclic Motor Activity

The investigation of human fetal and neonatal CM reviewed above has revealed a fundamental property of the oscillations in spontaneous motor activity: they are sustained. They are a nearly ubiquitous, constant feature of spontaneous activity during a period of significant developmental changes and across qualitatively different behavioral states, and they persist under at least some pathologic conditions. This is not to say that the quantitative properties of CM measured by spectral analysis are fixed, but rather that they remain within well-defined limits under a range of circumstances. The oscillations appear to be intrinsic to the developing motor system and not normally driven or elicited by extrinsic factors.

Although the stable presence of CM is striking, it is matched by another equally robust property of the oscillations in spontaneous activity: they are irregular. A look at any movement time series immediately reveals this second, seemingly contradictory property of CM. The irregularity in CM has tended to be overlooked for two important reasons. First, the methods that have been used to analyze CM (spectral analysis), although powerful and capable of yielding quantitative estimates of some central characteristics of oscillatory phenomena (e.g., their frequency and strength), are designed to extract the linear contributions of regular, periodic processes in complex waveforms. Irregularity is reflected only indirectly as the finite width of the peak in the movement spectrum, or the presence of broadband power across portions of the frequency domain. Second, through our reliance on linear methods of analysis we have come to associate regularity with lawfulness and irregularity with noise in measurements or the environment. The theory of nonlinear dynamics, however, which focuses explicitly on the sequential, point-to-point fluctuations in the behavior of the system offers another alternative. It is possible, in principle, to account for both the stability and the minute-to-minute irregularity of CM with the same dynamic model, eliminating the need to postulate a causal role for stochastic sensory inputs or measurement error (Guckenheimer, 1982). It is to this task that we now turn.

State Space

The literature on dynamic systems has expanded very rapidly in recent years and there are now many useful texts and collections of papers. The four-volume series on dynamics by Abraham and Shaw (1982, 1983, 1985, 1988) is an exceptionally clear visual presentation of dynamic theory, which is summarized in Abraham and Shaw (1987). Other useful texts requiring varying degrees of mathematical expertise which emphasize chaotic dynamic systems are Glass and Mackey (1988), Guckenheimer and Holmes (1983), Schuster (1988), Thompson and Stewart (1986), the Springer Series in Synergetics edited by Haken (e.g., Haken, 1983), and the proceedings of a recent conference at the New York Academy of Sciences (Koslow, Mandell, and Shlesinger, 1987).

Before analyzing the dynamic properties of CM, it will be useful to introduce some of the basic concepts, definitions, and terminology that are required. In order to do this, assume (1) that we know the important variables that describe the mechanism responsible for CM and its evolution in time: these are the state variables in the sense that specifying the (time-dependent) value of each of these variables fully describes the current state of the system; (2) that we know how the state variables are related to one another in time, and how subsequent values of each depend on the previous values of all of them, and (3) that we know the parameters which weight or modify the time-dependent relationships among the state variables.

With a knowledge of the state variables for this dynamic system, we can construct a state space in which the coordinate axes are the state variables. The state of the system at any given time is then represented as a position in this (possibly high-dimensional) state space. The temporal evolution of the system is represented as a trajectory in the state space. Knowledge of how the state variables are related to one another in time allows us to trace the trajectory followed by the system from any arbitrary starting point. The collection of trajectories followed by the dynamic system is the phase portrait of the system. Finally, knowledge of the parameters involved in the functional relationships among the state variables permits us to specify the region of state space visited by the system and the precise nature of its temporal evolution in that region.

One (or more) region of the state space for a dynamic system may contain an attractor. Trajectories which begin in a certain region of state space (the attractor's basin of attraction or inset) lead eventually to a more restricted region of state space (or limit set). Similarly, small deflections of trajectories from an attractor which do not take the system out of the basin of attraction will be followed by a return to the attractor. Thus the theoretical concept of an attractor in the state space of a dynamic system embodies the experimental notions of stationary asymptotic behavior and stability following the transient response to a brief perturbation.

The types of attractors that have been studied in theory and identified experimentally include the following. Fixed point attractors are stable steady states in state space, and represent the absence of change in all of the state variables and the tendency to return to that identical state after small perturbations. Limit cycles are attracting loops in state space, and represent perfectly periodic (not necessarily sinusoidal) behavior of the system. An attracting torus (doughnut) in state space, toward which trajectories converge but then wind around without ever exactly repeating, represents quasiperiodic behavior due to the presence of two (or more) incommensurate frequencies. Finally, there are chaotic attractors where trajectories diverge in some direction(s) but converge in others. The exponential divergence (in at least one direction) of nearby trajectories on a chaotic attractor results in the characteristic long-term unpredictability in the behavior governed by chaotic dynamics. The stretching is combined with folding in other directions, however, so that the diverging trajectories remain in a bounded region of state space and lie on a well-defined (although topologically complex) surface. It is the deterministic irregularity of chaotic dynamics that makes it a seductive metaphor for describing the sustained but somewhat unpredictable fluctuations in spontaneous motor activity. The seduction comes from the possibility of accounting for both the stability of CM and its characteristic irregularity with the same dynamic model. The question is how we can test the hypothesis that CM is best described by chaotic dynamics and how we can carry the analysis beyond the simply descriptive.

Reconstructing a Phase Portrait

We can begin to go beyond the metaphoric description of CM as chaos by attempting to find an attractor in CM state space, determine its geometry, and measure the rate at which nearby trajectories diverge. These are reasonable and concrete first steps which, if successful, would need to be followed by additional efforts to characterize the properties of the attractor and the underlying dynamics in more detail, including studies of structural instabilities in the CM attractor and the relevant control parameters.

The first problem, as in many complex systems, is that the state variables for CM are unknown. Hence the state space for CM is undefined, as are the trajectories generated by the unknown dynamics. We therefore begin from a position of fundamental ignorance of the system. What we do have are data. In the case of CM, we have a readout of the system in the form of a movement time series. Intuitively, the details of this time series—the evolution of motor activity over many minutes—

should reflect in some (perhaps very complicated) way the underlying dynamics of CM (Packard, Crutchfield, Farmer, and Shaw, 1980; Takens, 1981). The theorems proved by Takens (1981) demonstrate that generically one can reconstruct a phase portrait of a dynamic system from a single time series of measurements on that system. The measurements need not be made on one of the state variables, as long as the measured variable reflects in a reasonably smooth way (at least through the second derivative) the trajectories on the unknown attractor.

Using Takens's method, a phase portrait is reconstructed in the following manner. Let $x_1, x_2, \ldots, x_n$ be the experimental time series, and let m be the dimension of the unknown state space. Construct vectors $\mathbf{y}_i$ whose components are successive values of the time series beginning with x_i. In general, $\mathbf{y}_i = [x_i, x_{i+t}, x_{i+2t}, \ldots, x_{i+2mt}]$, where t is a fixed time delay. The vectors $\mathbf{y}_i$ define points in a $2m + 1$ dimensional embedding space where the coordinate axes are measurements made at integer multiples of the time delay, t. The sequence $\mathbf{y}_i, \mathbf{y}_{i+1}, \ldots$ therefore defines a trajectory in the embedding space (figure 5.2). Takens's theorems show that a phase portrait reconstructed in this

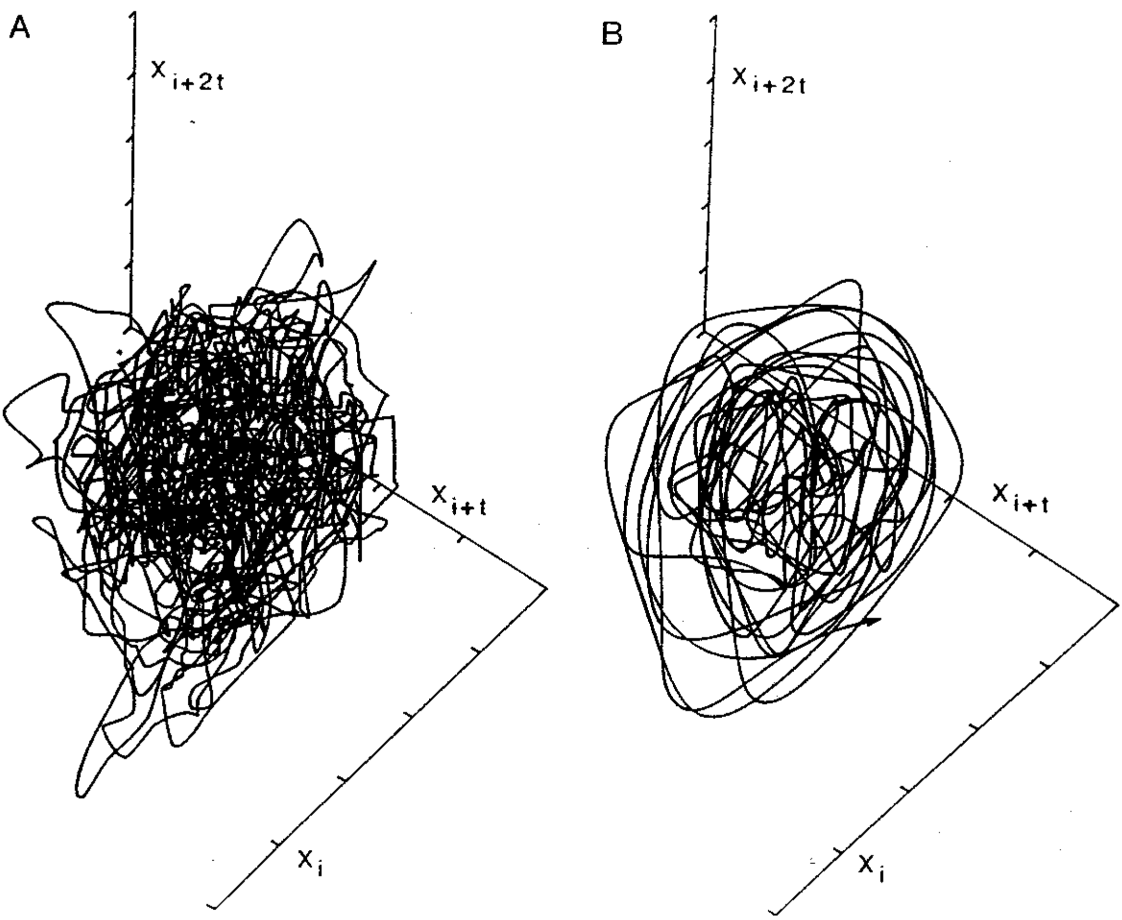

Figure 5.2
(*A*) The time series in figure 5.1 embedded in three dimensions (time delay = 15 seconds) and projected onto two dimensions. Successive points along the trajectory have been connected with a cubic spline. The view is from below the plane formed by the x_i- and x_{i+t}-axes (both increasing to the *left*, x_i below x_{i+t}), looking obliquely along the x_{i+2t}-axis. Scale is arbitrary, but is the same for each axis. (*B*) When the rapid fluctuations are suppressed with a low-pass filter (frequency response = .1 at 3 cycles/min) before embedding, structure not evident in *A* becomes visible.

way is an image of the unknown attractor which preserves many of its geometric and dynamic properties. Thus it is possible, beginning with a single time series consisting of a readout of the unknown dynamic system, to reconstruct a phase portrait from which properties such as the dimension of the unknown attractor can be estimated.

There are at least two practical problems that must be faced, however. First, since the dimension of the unknown attractor, m, is unknown, the dimension of the embedding space, $2m + 1$, is not known a priori. In practice, one usually begins with a small embedding dimension and increases it one step at a time, estimating the property of interest (e.g., the attractor's dimension) at each step until the estimates converge (e.g., a loop has one dimension whether it is drawn on a two-dimensional surface, in a three-dimensional volume, etc.).

The second practical problem is how to select the time delay, t, for reconstructing the phase portrait. In theory (Takens, 1981) it does not matter generically, but in practice it does. With very small delays, the components of the vectors $\mathbf{y}_i$ will be nearly the same, and so the trajectories in the embedding space will all be compressed into a long, thin volume equivalent to a diagonal in $2m + 1$ dimensions. The goal, of course, is to have the reconstructed phase portrait be as uniformly distributed as possible so that its properties can be accurately measured. If the attractor is, in fact, low-dimensional, it may be possible to inspect three-dimensional plots (or two-dimensional projections) of the reconstructed phase portrait for different time delays and select the time delay which gives the least compression of trajectories. For higher dimensions this is not possible. One method, which is described in more detail below, is to use the time delay that minimizes the redundancy in observing pairs of points separated by that amount of time (Fraser and Swinney, 1986). This has the effect of maximizing the independence of the coordinate axes in the embedding space.

Existence of an Attractor

The logical first step is to determine whether an attractor exists in the state space of CM. Given the potentially large number of degrees of freedom in the neuromotor system generating CM, is there a relatively low-dimensional attractor on which the system's trajectories lie? One could argue that since CM is sustained and the fluctuations are bounded, there is an attractor, perhaps even a noisy fixed-point attractor. This kind of argument, however, is not precise enough to evaluate the explanatory power of dynamic systems theory.

If we had a clear picture (visual or abstract) of the reconstructed phase portrait of CM, we could in principle make two important inquiries. First, we could ask whether trajectories that begin some distance away converge toward the phase portrait, which would be evidence that it is an attractor. In studies with human infants, however, we are not at liberty to manipulate such initial conditions (where a trajectory begins). This would correspond to stopping and restarting the mechanism responsible for CM, which we know is sustained under normal conditions. In an animal model, on the other hand, it might be possible to control initial conditions, perhaps pharmacologically.

A more plausible approach is to examine the way ongoing CM responds to perturbations. Do trajectories in the reconstructed phase portrait, when deflected by an experimental perturbation, return to the original phase portrait? If so, it would be evidence for a CM attractor. Although there is no fundamental theoretical difference between manipulating initial conditions and studying the response to single perturbations, the latter is much more tractable experimentally.

Obtaining evidence for the existence of an attractor is clearly important in this endeavor to apply dynamic systems theory to CM. But the kinds of experiments just described require a clear picture of the phase portrait in order to determine whether trajectories converge toward it from different initial conditions, or return to it following a perturbation that deflects them away. So the practical first step is to get a clear picture of the phase portrait.

Dimension

The first level of knowledge needed to characterize an attractor is its dimension (Farmer, Ott, and Yorke 1983). Intuitively, an attractor's dimension is a measure of its geometric complexity. For example, a fixed point has a dimension of 0, a limit cycle (a loop in state space) has a dimension of 1, and a doubly periodic attractor (a torus in state space) has a dimension of 2. The attractor's dimension also places a lower bound on the number of variables that will be needed to model the dynamics of the system on the attractor.

There are many ways to define the concept of dimension (e.g., Eckmann and Ruelle, 1985; Farmer et al., 1983; Grassberger and Procaccia, 1983; Mandelbrot, 1982; Mayer-Kress, 1986). The Kolmogorov capacity measure is one of the simplest and intuitively straightforward. If one determines how many very small cubes (hypercubes, really, since they may be more than three-dimensional) with edge length s it takes to completely cover the attractor, then the capacity dimension is defined as:

$$(1) \quad d_c = \lim_{s \to 0} \frac{\log N(s)}{\log (1/s)}$$

where $N(s)$ is the number of cubes. For a fixed point, $N(s) = 1$, so $d_c = 0$. Similarly, for a loop of length L, $N(s) = L/s$, so the limit as $s \to 0$ yields $d_c = 1$, which agrees with our intuition. When this definition is applied to other attractors, however, the value of d_c obtained is not an integer. These strange attractors are fractal objects with complex and sometimes beautiful geometry. As many have argued (e.g., West and Goldberger, 1987), fractal geometry is not merely a curiosity or an invention of mathematicians, but appears to characterize many objects in the natural world, from lungs to coastlines. Noninteger dimension is frequently a characteristic of chaotic attractors, although not necessarily (Grebogi, Ott, Pelikan, and Yorke, 1984).

The Kolmogorov capacity is a metric dimension (Farmer et al., 1983) in the sense that only the basic concept of distance is required for its definition. This contributes to its intuitive appeal but leads to problems when the Kolmogorov capacity or other metric dimensions are used to characterize highly nonuniform chaotic attractors. A typical feature of many chaotic attractors is the tendency for certain regions on the attractor to be visited by the system more frequently than others. Rather than moving gradually from one region to another, trajectories may linger in one region and then move rapidly to another. Or trajectories may evolve smoothly, but return to some regions more often than others as a result of the stretching and folding that characterize chaotic attractors. The highly nonuniform fluctuations in spontaneous motor activity would suggest that, if a CM attractor exists, we might expect it to have this property.

The information dimension, in contrast, is a probabilistic measure of dimension (and a generalization of the Kolmogorov capacity) which takes into account the relative probability of finding the system on each small region of the attractor (Farmer et al., 1983):

$$(2) \quad d_I = \lim_{s \to 0} \frac{\log I(s)}{\log (1/s)}$$

where

$$(3) \quad I(s) = -\sum_{i=1}^{N(s)} P_i \log P_i$$

and P_i is the probability of finding the system in the ith cube. The information dimension, and other probabilistic definitions of dimension, therefore capture more of the dynamics of the system whose stable behavior is characterized by a nonuniform attractor.

In most cases, however, it is extremely difficult to implement the box-counting procedures necessary to calculate the information dimension from a particular data set. Therefore, because of the fundamental importance of knowing an attractor's dimension, there has been a substantial effort to develop measures of dimension that are computationally tractable (Mayer-Kress, 1986). At the core of some of the more useful methods (e.g., Grassberger and Procaccia, 1983) is the following intuitive notion: if points are distributed in a d-dimensional volume, then the number of points within a distance r of some reference point will be proportional to r^d. For example, if the points are distributed along a line (such as a limit cycle in state space), then the number of points within a distance r of some reference point will simply be proportional to the distance r. Similarly, if the points are distributed on a two-dimensional surface, the number of points within a distance r will be proportional to the area swept out by the distance r, i.e., proportional to r^2. Thus an intuitive generalization is:

$$(4) \quad N(r) = cr^d$$

which leads to the very useful result that

$$(5) \quad \log N(r) = \text{constant} + d \log r$$

That is, the slope of log $N(r)$ vs. log r should be an estimate of the dimension, d.

Grassberger and Procaccia (1983) define the correlation integral, $C(r)$, as the average value of $N(r)$ across all the points on the reconstructed phase portrait. They show that for a sufficiently large number of points and sufficiently small r, a scaling region will exist in which log $C(r)$ vs. log r is linear. The correlation dimension is defined as the slope in that linear scaling region. Furthermore, the correlation dimension provides a lower bound on the information dimension defined above.

There are a number of important caveats, however (Guckenheimer, 1984; Mayer-Kress, 1987). Obviously, the relationship will not hold when r is large enough to extend beyond the attractor. More important, the scaling relationship will not hold when the distance r, which defines the neighborhood of the reference point, is less than the separation of the points in state space. Thus, we would expect that only for intermediate values of r might there be a scaling region in which this relationship holds. This is not a trivial matter, since the definition of the correlation dimension holds in the limit of very small distances. Two factors conspire to destroy the scaling region when working in a state space reconstructed from an experimental time series. First, poor resolution in the data means large distances between points in the reconstructed state space. Second, noise in the measurements is space-filling at small distances in the state space and so will mask the scaling relationship.

The second caveat is the number of data points required for a reliable estimate of the correlation dimension. Although it can sometimes be done with less (Kurths and Herzel, 1987), 10^d data points are recommended. For anything except low-dimensional attractors, this constraint is crippling for most of us in the behavioral and developmental sciences, and emphasizes the practical importance of experiments designed to measure changes in dimension rather than their absolute value.

The third caveat is that considerable information about the dimensional complexity of the attractor may be lost in the averaging process proposed by Grassberger and Procaccia (1983). For a highly nonuniform attractor, the local dimension evaluated at a single reference point may vary considerably for reference points in different regions of the attractor. It may therefore be useful to estimate the local dimension of an attractor for a sequence of reference points along a trajectory to obtain information about the temporal evolution of the dimensional complexity. Clearly, however, the demands for many data points and computational effort are multiplied approximately by the number of reference points used, a situation which for most of us studying behavioral organization is prohibitive.

Estimating the Dimension of a Cyclic Motor Activity Attractor

As noted earlier, we have no direct evidence that a CM attractor exists. But if one exists, its dimension should be preserved in a phase portrait reconstructed from a suitable readout of the underlying dynamic system, subject to the additional constraints of resolution and noise just discussed. In the following therefore, we illustrate the analysis of a reconstructed phase portrait for evidence of a finite-dimensional attractor.

The movement time series shown in figure 5.1 was obtained from an infant during approximately 27 minutes of awakeness 2 months after birth. The infant was studied in a sound-attenuated laboratory, and had an interesting but static visual environment. Body movement was detected by sensors beneath the infant. The sensor output was digitized on line at 50 Hz and stored for later processing. Thresholds were determined (off-line) which represented the range of sensor output due to respiration when the infant was not otherwise moving. The movement time series was formed by counting the number of 0.1-second intervals that contained suprathreshold sensor activity in each second during the 1630 seconds during which the infant was awake and not fussy. Linear and very slow curvilinear trends (less than 0.11 cycle/min) were removed before subsequent analysis.

The time series was then embedded in an embedding space of successively higher dimension using the method of delays (described earlier) proposed by Takens (1981) and FORTRAN routines written by Schaffer, Truty, and Fulmer (1989). The time delay was set at 15 seconds based on the first minimum of mutual information, as suggested by Shaw (Fraser and Swinney, 1986). Mutual information is a general measure, based on information theory, of the extent to which the values in a time series can be predicted by earlier values. It is not limited to linear dependence like the autocorrelation function. Fraser and Swinney (1986) argue that when a phase portrait is reconstructed from a time series using a time delay that minimizes mutual information, the trajectories in the reconstructed state space will be maximally separated, which in turn will facilitate their analysis.

For each embedding dimension, the correlation dimension of the reconstructed phase portrait was estimated using the procedure proposed by Grassberger and Procaccia (1983) described above, and the computational routines of Schaffer et al. (1989). Figure 5.3 shows the results. For computational simplicity, the original data are rescaled to the interval [0, 1]. The value of r is the usual Euclidean distance in the state space reconstructed from the rescaled data. Panel A plots $\ln C(r)$ vs. $\ln r$ separately for each embedding dimension, for $\ln r$ between -4.7 and 0. As $\ln r$ increases, the number of points in the neighborhood defined by the distance r increases, as expected, until all points are included when $r = 1$ ($\ln r = 0$).

To aid in finding the scaling region where $\ln C(r)$ increases in an approximately linear fashion with $\ln r$, the slopes of the small line segments connecting successive values of $\ln C(r)$ are plotted against $\ln r$ in panel B (again, for each embedding dimension). For other than the smallest values of $\ln r$ (corresponding to distances smaller than the resolution of the original measurements) it is apparent that the slopes increase as the time series is embedded in successively higher dimensions. For $\ln r$ near -2 (marked by the vertical bars), the slopes of the small line segments (in panel A) do not systematically increase or decrease with changes in $\ln r$, and for some embedding dimensions (e.g., $d_e = 3$) they are relatively constant. This suggests the existence of approximately linear scaling regions in the plot of $\ln C(r)$ vs. $\ln r$, which are marked in panel C. Furthermore, after an embedding dimension of 3, the slopes in this region around $\ln r = -2$ stop increasing. This suggests, although not strongly, that the correlation dimension might be converging to a value between 3 and 4. It is worth pointing out that various algorithms have been developed to automate the process of finding a scaling region of acceptable size (e.g., Albano, Mees, deGuzman, and Rapp, 1987; Mayer-Kress, Yates, Benton, Keidel, Tirsch, Poppl, and Geist, 1988).

Returning to panel A, straight lines are fitted to the plots of $\ln C(r)$ vs. $\ln r$ in the presumptive scaling region. These are shown in panel C. The slopes of these best-fit (by least squares) lines, which are the estimates of the correlation dimension for each embedding dimension, are plotted in panel D along with their 95% confidence intervals. The apparent convergence of these estimates suggests that the phase portrait reconstructed from the movement time series has a finite dimension with a value of approximately $3.1 \pm .3$ (the value for $d_e = 5$). For noise (or a high-dimensional phase portrait) the estimates of the correlation dimension would continue to increase as the embedding dimension increased (as in figure 5.4 below). To the extent that the reconstructed phase portrait is an image of the unknown CM attractor which preserves its dimension, these results are consistent with the existence of a relatively low-dimensional attractor governing the dynamics of spontaneous motor activity.

Before discussing the possible implications of this finding, it is appropriate to draw attention to the limitations of the results. First, the scaling region is of very small size, which decreases our confidence about its existence as well as limiting the accuracy of the estimate of the slope of $\ln C(r)$ vs. $\ln r$ in that region, as seen in panel C and reflected in the error bars in panel D. Second, the slopes in the region around $\ln r = -2$ are not constant, but fluctuate somewhat. Both of these limitations are due in part to the relatively poor resolution in the original data and almost certainly to the presence of noise, or fluctuations in the movement data that do not directly reflect the dynamics responsible for CM.

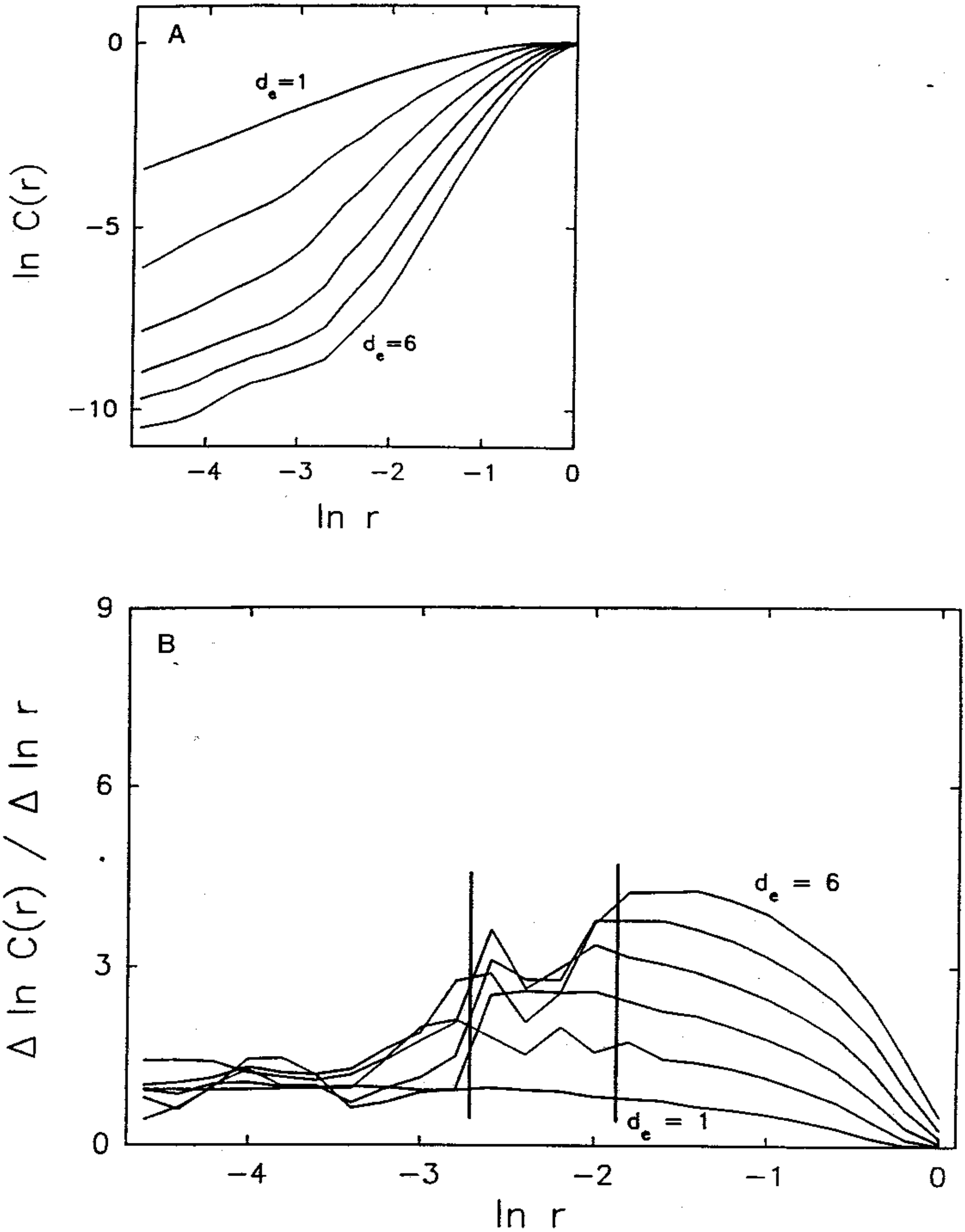

Figure 5.3
Dimension calculations for the data in figure 5.1 (unsmoothed). (*A*) Correlation integral vs. distance for embedding dimensions 1 through 6. (*B*) Slope of the line segments connecting adjacent estimates of the correlation integral in *A*, plotted against distance for each embedding dimension. *Vertical bars* mark the scaling region. (*C*) As in *A*, but with points not connected. The *solid lines* are the least-squares regressions of ln $C(r)$ on ln r in the scaling region. (*D*) Estimates of the correlation dimension, and their 95% confidence intervals, based on the slope of the regression lines in *C*, plotted for each embedding dimension.

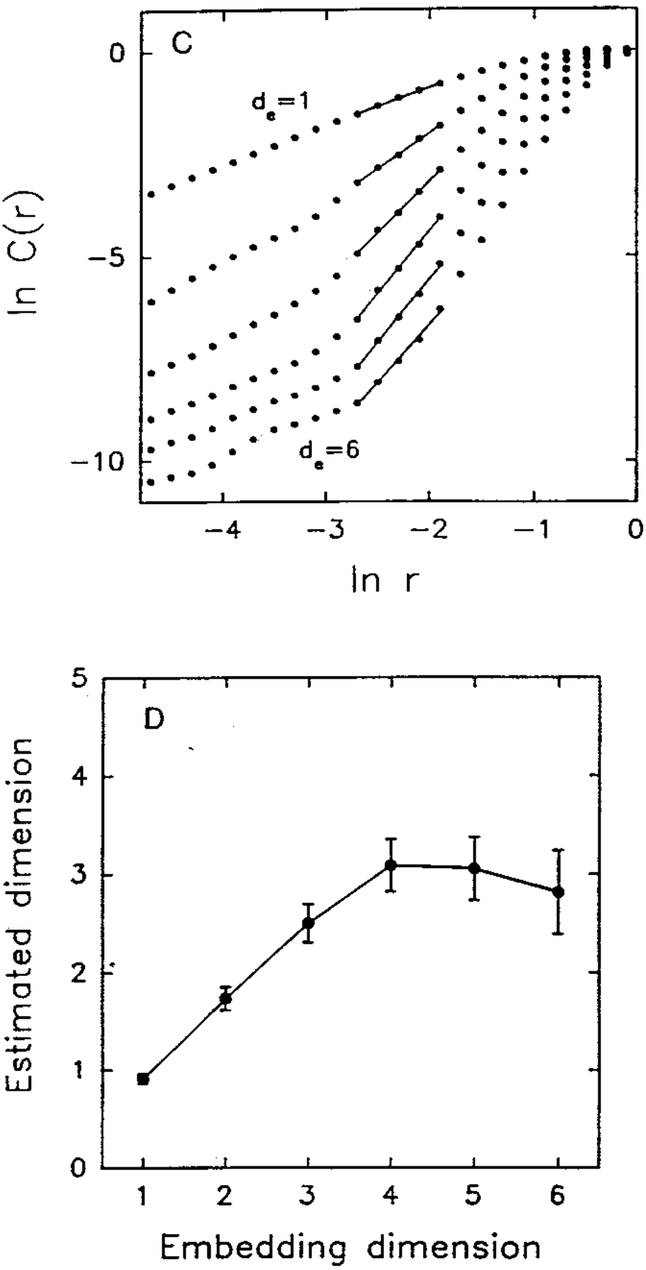

Figure 5.3 (continued)

These are not easy problems to solve. The limited resolution results from the decision to look during 0.1-second intervals for suprathreshold activity in the movement sensors. Looking more often would provide more resolution in the data, but might be of questionable validity at the behavioral level because few limb or body movements last less than 0.1 second. On the other hand, individual muscle contractions can be very brief, and might contain information about the dynamics underlying CM. In the data analyzed here, the presence of noise was minimized by setting the thresholds used to process the raw sensor signals wide enough to exclude most sensor activity due to respiration. But this inevitably also excludes small amounts of sensor activity due to body movement. When the thresholds were set closer, the estimates of the correlation dimension did not converge (figure 5.4). However, if the data from the narrow thresholds are smoothed using a 15-second moving average, the scaling region is large, the estimates of the slopes of $\ln C(r)$ vs. $\ln r$ have very small statistical error, and the estimates of the correlation dimension converge to a value of $3.27 \pm .03$ at $d_e = 5$ (figure 5.5). This value is within the error of the estimate ($3.1 \pm .3$) based on the unsmoothed data (see figure 5.3) in which wider thresholds were used in processing the raw sensor signal. Although this degree of smoothing should not affect the low-frequency events that have been characterized as CM, there may be higher-frequency motor events that are important for a more complete understanding of CM and that are lost in the smoothing procedure.

Implications of Finite Dimension

There are several implications of finding that the phase portrait reconstructed from a movement time series has a correlation dimension between 3 and 4. First, and possibly most important, is that the fluctuations in spontaneous motor activity are not simply due to noise (Rapp et al., 1986). If they were, the dimension estimates would fail to converge. That is, the slope of $\ln C(r)$ vs. $\ln r$ would continue to increase as the embedding dimension increases, since noise will be space-filling in the reconstructed state space. Thus, we are now more justified in talking about the underlying dynamics of CM as being deterministic. It is important to note in this context that the failure of the dimension estimates to converge would not be evidence that the fluctuations are random. As discussed earlier, there are severe data requirements in order to resolve the image of a high-dimensional attractor in the reconstructed phase portrait. Even with enough data, the presence of noise and poor resolution can be limiting factors.

Finding a correlation dimension between 3 and 4 also suggests that the potentially large number of degrees of freedom involved in the dynamics of spontaneous motor activity may be reduced to a handful in stable CM. This reduction in degrees of freedom is theoretically interesting in its own right as an example of self-organization (Haken, 1983). If similar results are obtained in more extensive analyses of additional samples of CM, the implication is substantial. Finding evidence for low-dimensional dynamics implies that the complex and irregular nature of CM is the output of an effectively simple system. This insight is certainly not accessible through atheoretical statistical analysis of the movement time series, nor is it accessible through otherwise powerful methods such as spectral analysis. Since we know that the number of active elements in the neuromotor system generating CM is very large, the reduction in degrees of freedom implies there is substantial linking or coupling of those elements in stable CM. Identifying the small number of state variables in stable CM is likely to depend on experiments with animal models. In any case, it will be especially important

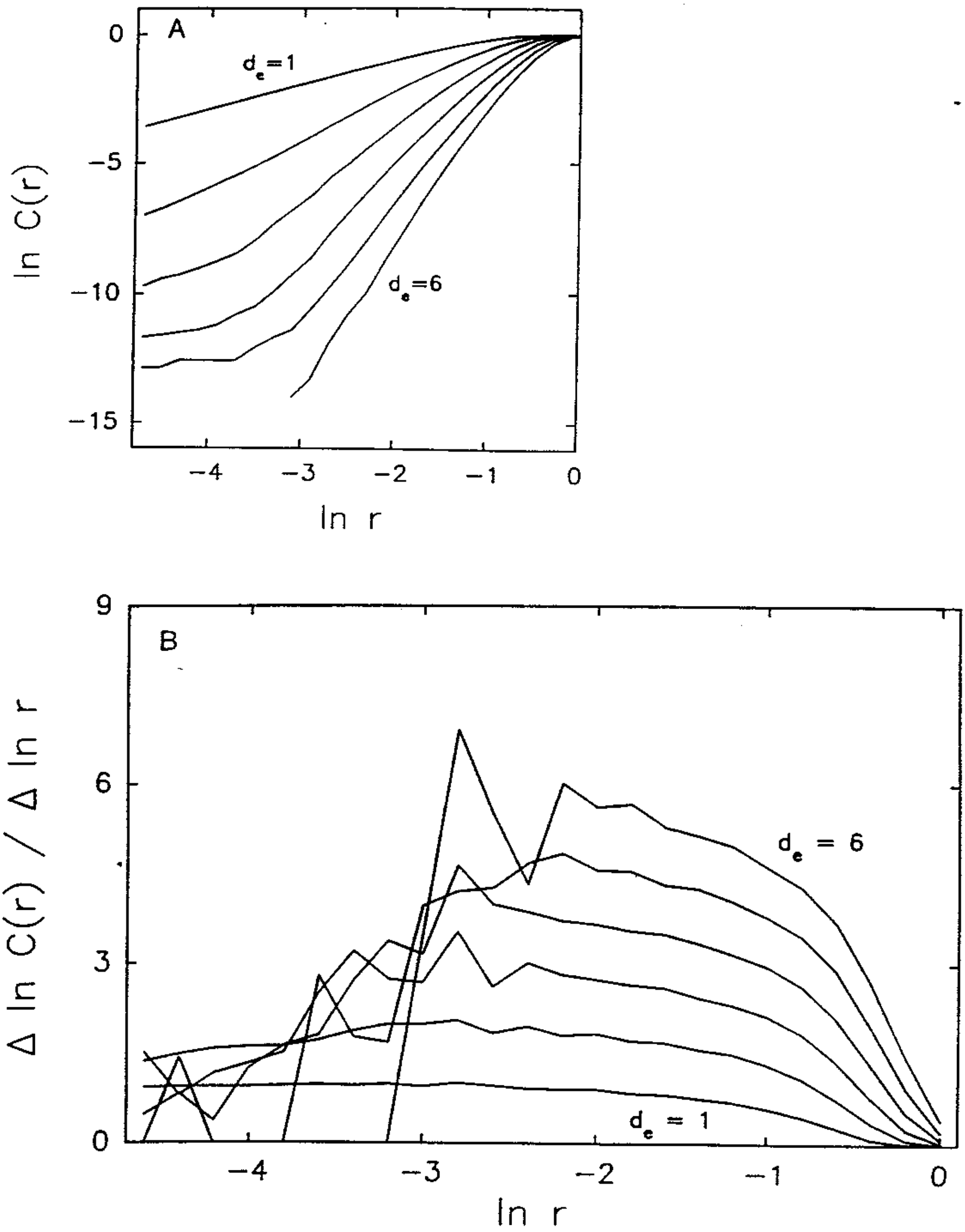

Figure 5.4
Dimension calculations for data obtained during the same 1630 seconds as in figure 5.1, but with narrower thresholds and hence greater noise. *A* and *B* are the same as in figure 5.3, except there is no scaling region (and hence no panels *C* and *D*).

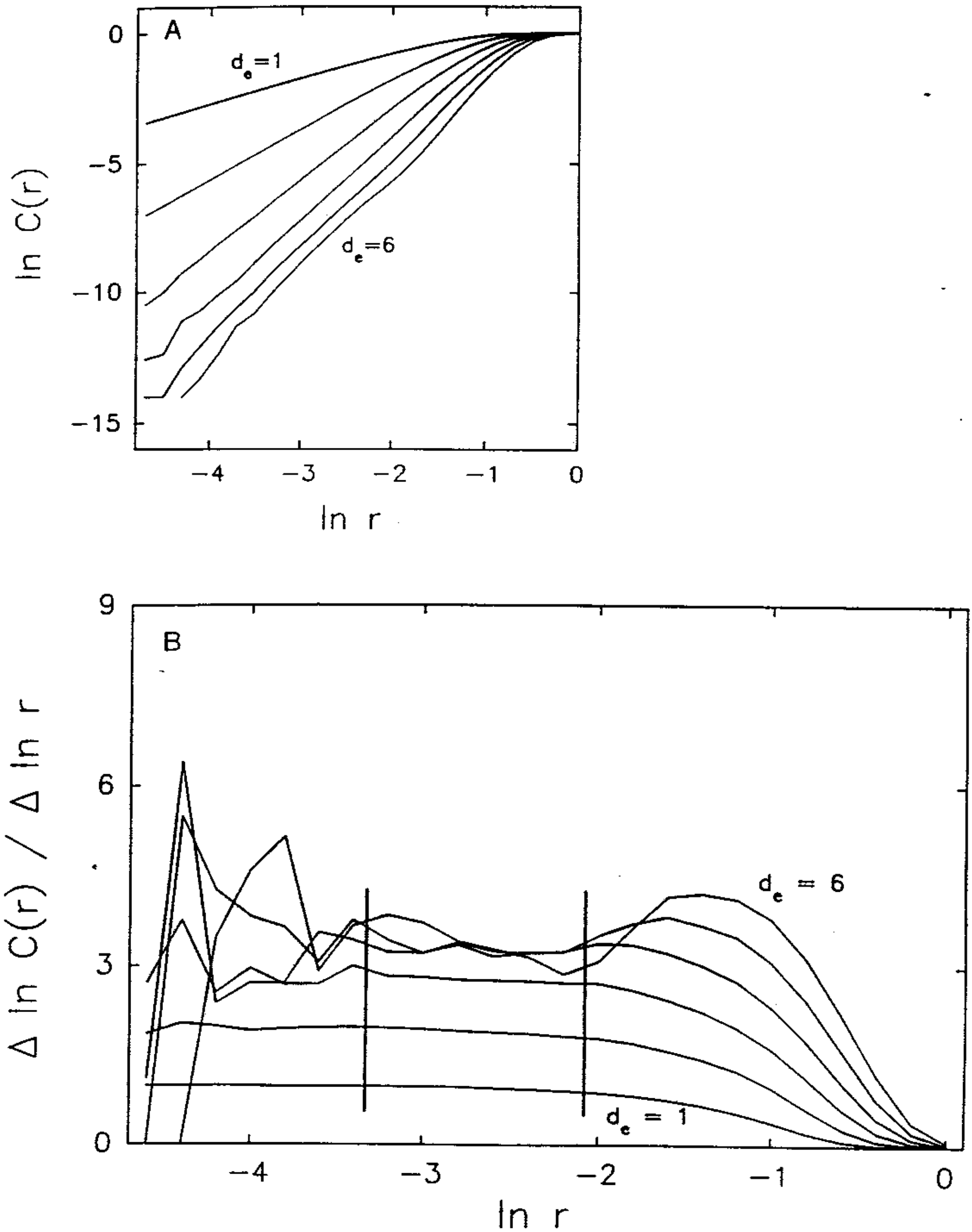

Figure 5.5
Dimension calculations for the data used in figure 5.4, but after smoothing with a 15-second moving average. *A–D* are the same as in figure 5.3.

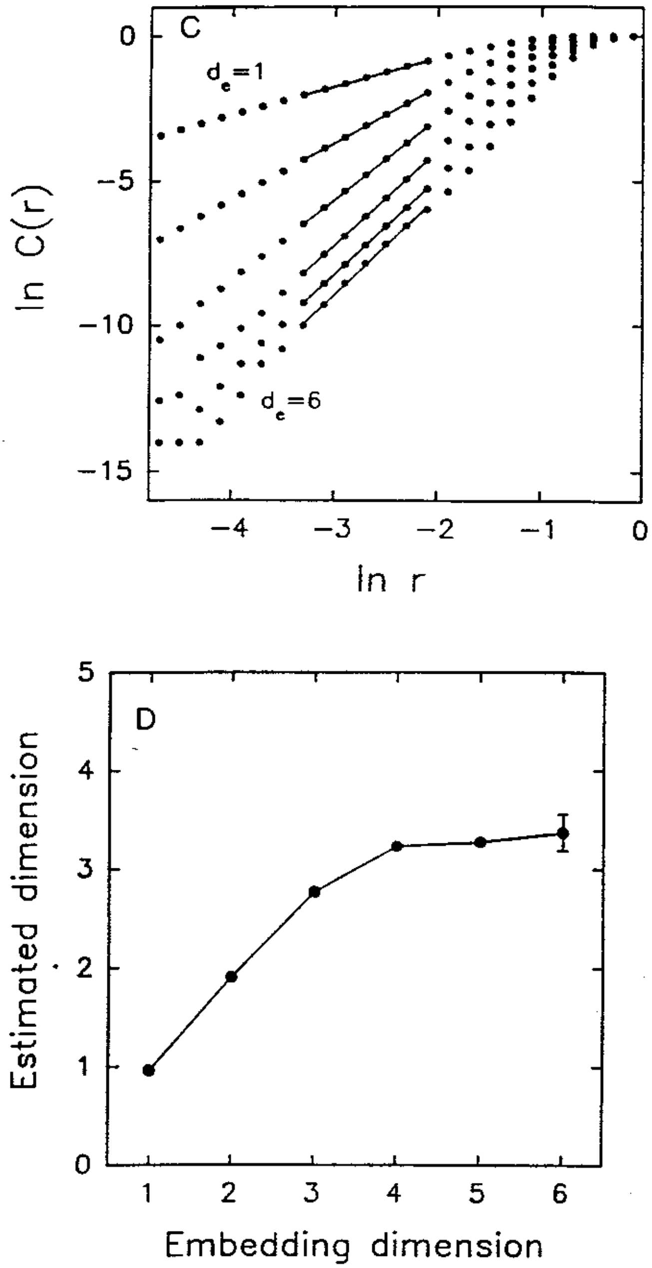

Figure 5.5 (continued)

to determine whether those variables have psychobiological meaning. If they do, it is much more likely that a good dynamic model of CM will lead to insights into the functional significance of CM as an action system through which the infant interacts with the rest of the world.

Finally, it is important to emphasize again the practical difficulties in estimating the dimension of an attractor from the analysis of a phase portrait reconstructed from a time series. We may not be justified in making strong quantitative inferences about dimensionality if the experimental data are limited in amount and quality. Although moderate limitations might not seriously damage conclusions about nonrandomness, low degrees of freedom, or changes in dimension in response to the manipulation of internal and external conditions, serious data limitations will obviously be a major obstacle in building and testing dynamic models of CM. This is significant, because really going beyond the metaphoric redescription of CM as chaos will require building a dynamic model with strong psychobiological foundations.

Divergence of Nearby Trajectories

The stable presence of CM motivated us to invoke the concepts of dynamic systems theory and talk about a possible CM attractor. In an effort to go beyond the metaphoric description, we constructed a phase portrait from the movement time series and found evidence that it might contain the image of a CM attractor with a correlation dimension between 3 and 4.

Similarly, it is the characteristic irregularity of CM which made us speculate that the fluctuations in spontaneous motor activity might be governed by a chaotic dynamic system. Since the defining (Farmer et al., 1983) or diagnostic (Wolf et al., 1985) property of a chaotic attractor is the exponential divergence of nearby trajectories, the next step in pushing beyond the metaphor should be to see if there is any evidence for the divergence of trajectories in CM.

The divergence (and convergence) of nearby trajectories is quantified by Lyapunov exponents, for which there are a number of equivalent definitions (Gershenfeld, 1988). One of the more intuitive definitions, which relates directly to attempts to estimate Lyapunov exponents from experimental data (e.g., Eckmann and Ruelle, 1985; Wolf et al., 1985), is the following: In the k-dimensional state space of a dynamic system, take an infinitesimal volume (a k-sphere) of states and measure the long-term expansion and contraction of that sphere as it becomes ellipsoid due to the divergence (and convergence) of the trajectories originating in the k-sphere. If the principal axes of the ellipsoid after a time t are $a_i(t)$, then the ith Lyapunov exponent is defined as:

$$(6) \quad \lambda_i = \lim_{t \to \infty} \frac{1}{t} \log \frac{a_i(0)}{a_i(t)}$$

where $a_i(0)$ are the lengths of the axes at the start (i.e., the k-sphere). That is, λ_i is the average long-term expansion or contraction (in bits per time unit) of the corresponding principal axis of the original infinitesimal sphere of initial states. The λ_i are conventionally labeled in order of decreasing magnitude, so λ_1 is the largest Lyapunov exponent and corresponds to the direction of greatest divergence if it is positive. It is important to note that the orientation of the ellipsoid changes as it evolves in time, so the direction of greatest divergence does not correspond to a fixed direction in state space. In view of the importance of knowing the dimension of an attractor, it is also interesting to note that for many attractors the information dimension (for which the correlation dimension is a lower bound) appears to be related in a rather simple way to the spectrum of Lyapunov exponents (Farmer et al., 1983; Kaplan and Yorke, 1979; Mayer-Kress and Kaneko, 1989).

It is common to take $\lambda_1 > 0$ as the definition of chaos (e.g., Farmer et al., 1983). The magnitude of λ_1 reflects how rapidly the system's behavior becomes unpredictable, which results from the inability to resolve infinitesimal differences in initial conditions. In contrast, a fixed point attractor in a three-dimensional state space has all $\lambda_i < 0$, reflecting the convergence of infinitesimal perturbations in all directions. A limit cycle in three-dimensional state space has $\lambda_1 = 0$ and $\lambda_2, \lambda_3 < 0$, with a similar interpretation.

The calculation of all λ_i is relatively easy when the differential equations defining the dynamic system are known. The problem experimentalists face is how to estimate λ_i from data. The Lyapunov exponents are preserved in the state space reconstructed from a suitable time series, but there are two aspects of the theoretical definition of λ_i that must be addressed by any computational algorithm. First, the definition of λ_i is based on infinitesimal distances (the initial separation of nearby trajectories). The limited resolution in the measurements that make up the experimental time series makes very small length scales inaccessible, and noise (although it may play an important dynamic role) is space-filling at small length scales, so intermediate distances must be used. Second, the definition of λ_i is based on infinite time, but data sets are finite. Wolf and Vastano (1986) present some limited evidence that one of the most widely used algorithms (Wolf et al., 1985) yields reasonable results in spite of these inherent problems in applying the theoretical definition. They also argue that if an estimate of λ_1 is to be used as a diagnostic test for the presence of chaotic dynamics, then even the minimal information that $\lambda_1 > 0$ is useful. The algorithm of Wolf et al. (1985) for estimating the largest Lyapunov exponent is intuitively straightforward. First, the nearest neighbor to the initial point in the reconstructed state space is found, and the separation between them, D, is calculated. The subsequent trajectories of these two points are followed for a fixed amount of time, and the new separation between the trajectories, D', is calculated. A new test point is then found that is as close as possible to the original trajectory but which preserves as much as possible the relative orientation of the last-used point on the old test trajectory. This process is an attempt to monitor the long-term behavior of a single axis in the theoretical ellipsoid of evolving states. The procedure is repeated until the end of the data set is reached. The overall time average of the values of $\log_2 D'/D$ obtained at each step is used as the estimate of λ_1.

There are a number of important decisions which must be made in attempting to estimate the largest Lyapunov exponent from experimental data (Wolf et al., 1985). These are not merely technical details, but represent fundamental issues that must be addressed if we are going to bring dynamic theory into contact with data. In dealing with these issues, we are forced to respect the constraints imposed on us by our experimental systems.

The first set of decisions concerns how to reconstruct the phase portrait from the time series. The objective is to facilitate accurate measurements of the separation of trajectories in the state space. Thus a time delay that minimizes the compression of the reconstructed trajectories will be optimal. This is the same criterion used to select the time delay for estimating the attractor's dimension. Similarly, the embedding space should be large enough to resolve the attractor's geometry (i.e., the embedding dimension should be larger than the attractor's dimension), but no larger, since noise plays an increasingly important role in the computations as the embedding dimension increases.

The second decision is when to replace test trajectories with new ones in monitoring their divergence. If replacements are done too often, the errors involved in finding a new trajectory that is nearby and in the same orientation as the old one will accumulate and thereby reduce the precision of the estimate. In addition, if test trajectories are replaced so often that the interval is on the order of noise-induced variations in the data, then λ_1 is likely to be overestimated because replacement is always done with the closest point (subject to orientation constraints). On the other hand, replacement should be done before the separation between the monitored trajectories grows to the size of the attractor, after which λ_1 is likely to be underestimated.

Finally, there are significant issues concerning the amount and quality of data required. Wolf et al. (1985) analyze the amount of data needed from a number of perspectives, and suggest at least 10^d, where d is the attractor's dimension. This is similar to the requirement others have recommended in estimating the attractor's dimension, and emphasizes the difficulty in measuring the divergence of trajectories on a high-dimensional attractor. Data quality is degraded by the presence of noise, which makes small distances in the reconstructed state space unusable. One solution is to avoid small distances in choosing nearby trajectories at each replacement step. Another solution is to smooth the time series to reduce the effects of noise. As Wolf et al. (1985) point out, smoothing may distort the shape of the reconstructed attractor, but they find that modest low-pass filtering often has little effect on the long-term rate at which nearby trajectories diverge.

Estimating Divergence on a Cylic Motor Activity Attractor

If we are to have any justification for thinking that CM might be the readout of a chaotic dynamic system, we must look for evidence of "the dynamic diagnostic for chaotic systems" (Wolf et al., 1985), a positive Lyapunov exponent. The time series in figure 5.1, which yielded some evidence for a finite-dimensional CM attractor, was therefore analyzed in an attempt to estimate the largest Lyapunov exponent.

The time series was embedded in a five-dimensional embedding space using Takens's (1981) method of delays. The embedding dimension was selected to be larger than the estimated dimension of the attractor (between 3 and 4) to avoid mixing of different regions of the reconstructed attractor, which would be catastrophic for the estimation of the rate at which trajectories diverge. The time delay was 15 seconds, which was found previously to provide an optimal separation of the trajectories in the reconstructed state space, and would therefore facilitate an accurate measurement of their rate of divergence.

The fixed evolution time algorithm of Wolf et al. (1985) was used, with computations done by FORTRAN routines written by Schaffer et al. (1989). The fixed evolution time refers to the fact that test trajectories are replaced after their divergence from the original (fiducial) trajectory has been followed for a fixed interval of time. The computations were repeated for a range of evolution times in order to obtain some information about the stability of the estimates of λ_1, and to identify at what point overestimation (with short evolution times) became evident. Spectral analysis of the

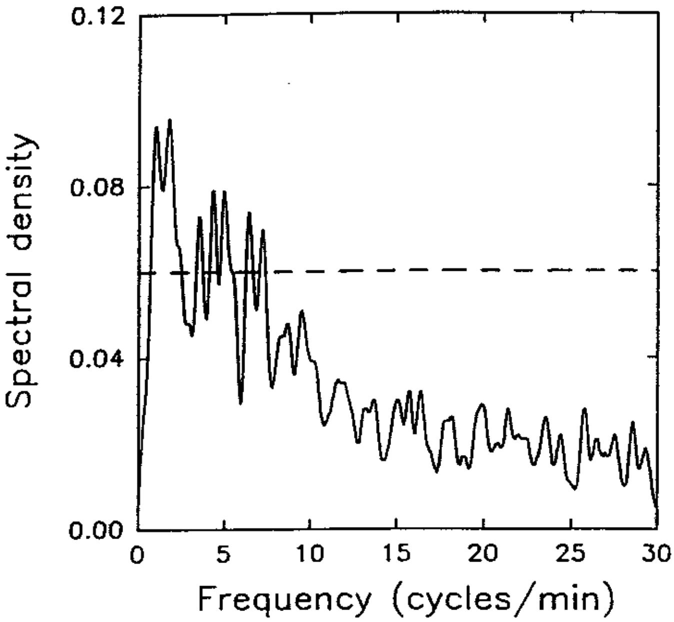

Figure 5.6
Power spectrum of the data in figure 5.1 (unsmoothed) using a Tukey lag window (Jenkins and Watts, 1968) covering 150 points, which has a bandwidth of .53 cycle/min. *Broken lines* are the 99% confidence limits for the spectral estimates of a white noise process.

time series (figure 5.6) indicated that the dominant period in the data was near 40 seconds (1.5 cycles/min), which was taken as a measure of the average orbital time for trajectories on the CM attractor. The largest Lyapunov exponent was estimated for evolution times ranging from 25% to 300% of the average orbital time.

Acknowledging the limited resolution in the original measurements, the initial separation between the fiducial and any test trajectory was not permitted to be less than 10% of the measurement range. In order to provide for an adequate number of replacement points to select from, the maximum allowable separation between the fiducial and any test trajectory was set at 30% of the measurement range, which is rather large.

The results of these computations are shown in figure 5.7. It is clear that over quite a large range of evolution times, the estimates of the largest Lyapunov exponent are uniformly greater than 0. The estimate of λ_1 is greatest using an evolution time of 10 seconds, with a rather precipitous drop with evolution times of 20 or 30 seconds. Evolution times of 30 to 60 seconds yield more stable estimates between 1.1 and 1.5×10^{-2} bits/sec. For evolution times longer than 60 seconds, the estimates of λ_1 are somewhat lower and range from 0.7 to 1.0×10^{-2} bits/sec.

Since the dynamics generating CM are unknown, it is difficult to know which estimates of λ_1 to use. Wolf et al. (1985) note that for systems that produce substantial divergence of nearby trajectories during each orbit, evolution times between 50% and 150% of the average orbital time yield the most reliable estimates of λ_1, with longer evolution times tending to underestimate λ_1. If CM were governed by such a mechanism, evolution times between 20 and 60 seconds might be expected to yield the most reliable estimates in this case. However, the estimates do not appear to stabilize until evolution times of 30 seconds or longer are used. Given the expected tendency toward underestimation as evolution times get longer, perhaps the estimates based on 30 to

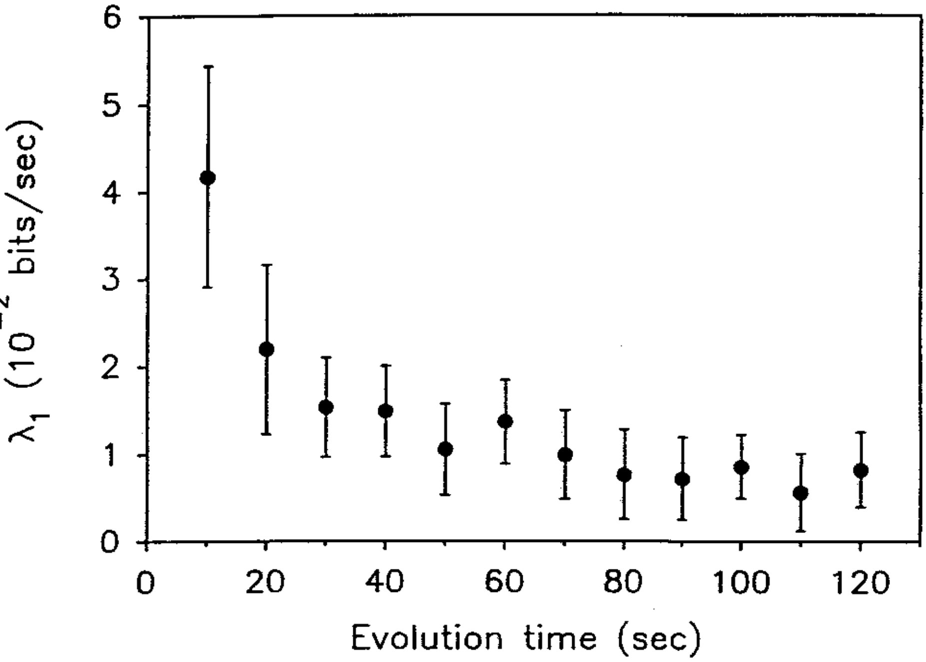

Figure 5.7
Estimates of the largest Lyapunov exponent, and 95% confidence intervals, for evolution times between 10 and 120 seconds.

60 or 70 seconds, which represent the first relative plateau or stable region after the rapid decline, should be used. This would put λ_1 in the range of 1.0 to 1.5 $\times$ 10^{-2} bits/sec, with 95% confidence intervals of $\pm 0.5 \times 10^{-2}$ bits/sec.

Our ability to accurately estimate the magnitude of λ_1 is clearly limited, which should inhibit us from making strong statements about the precise rate at which nearby trajectories diverge in this system. On the other hand, there appears to be relatively good evidence in this case to support the conclusion that $\lambda_1 > 0$. Numerous other calculations in which λ_1 was estimated using the algorithm of Wolf et al. with less defensible input parameters yielded similar results, always with $\lambda_1 > 0$. Thus it appears that the defining or diagnostic criterion for chaotic dynamics has been satisfied.

Implications of Diverging Trajectories
There are at least two important implications of finding evidence that the irregular fluctuations in spontaneous movement might be due to chaotic dynamics in the underlying mechanism. First, it provides some justification for considering the sustained oscillations and their irregularity as equally fundamental properties of the same mechanism. That is, it may not be necessary to postulate stochastic events or noisy environments as sources of the irregularity distinct from the regularity.

The second important implication of finding evidence for chaotic dynamics governing CM is the constraints and guidance it provides for thinking about the functional significance of this type of behavioral organization. If CM is fundamentally unpredictable, then we cannot look for possible consequences of CM that would depend on predictability. For example, it would make no sense to view CM as a timekeeping device. Instead, we should think of ways in which bounded unpredictability might be particularly useful. Others have suggested that chaotic dynamics in the central nervous system is highly adaptive, although the proof of this contention has been difficult to establish (Mpitsos et al., 1988; Skarda and Freeman, 1987).

In the newborn and young infant, spontaneous motor activity constitutes an important link to the physical and social environment. Increased activity is likely to elicit responses from the environment, even change it, and is also likely to reorient the infant's focus of attention. Decreased activity, on the other hand, permits sustained attention to a portion of the environment and the opportunity to process information extracted from the environment. Thus it has been suggested that fluctuations in motor activity may be especially useful because they combine the benefits of action and inaction over periods of minutes (Robertson, 1989).

Is there any reason to think that chaotic fluctuations might have different or greater benefits than regular oscillation? The exponential divergence of nearby trajectories on a CM attractor means that in the long run the fluctuations are unpredictable. In the short run predictability is greater, in contrast to random fluctuations (Guckenheimer, 1982). Perhaps there are benefits to somewhat predictable actions in the short run and unpredictable actions in the long run. Over short time scales, repetitive action patterns might support stable interactions with the environment, especially the social environment (as in behavioral dialogues with an adult). But over longer time scales, irregularity in the infant's action might reduce the chance of becoming locked into repetitive patterns of interaction with the environment, facilitate exploration, and maximize the information extracted from the environment.

We should remember, of course, that the existence of fluctuations in activity (chaotic or otherwise) is not, in itself, evidence of their utility (Gould and Lewontin, 1979). It is also possible that chaotic dynamics in CM has no benefits, or may instead be disruptive. For example, it may significantly reduce the success of early social interactions while adding little or nothing to the efficient extraction of information from the environment. These are empirical questions which developmental studies should go a long way in answering.

The theoretical significance of $\lambda_1 > 0$ in CM is that irregularity is being produced by the system, not imposed on it from the outside. Thus we are forced to consider the functional implications for an action system driven by such a mechanism. That is, we must examine the possible benefits (as well as the costs) of built-in unpredictability, something we are not accustomed to doing in the functional analysis of behavioral organization.

Next Steps

Our main goal has been to illustrate how the application of dynamic systems theory might be taken beyond the metaphoric level of redescription by asking some specific questions about a particular behavioral phenomenon. We have focused very closely on the two basic characteristics of CM which motivated the use of concepts from dynamic systems theory. The stability of CM suggested, at a metaphoric level, the concept of an attractor. The irregularity of CM suggested that the underlying dynamics might be chaotic. In both cases, we illustrated what was involved in going beyond the metaphor in order to (1) obtain the first level of knowledge necessary to characterize an attractor (its dimension) and (2) look for diagnostic symptoms of chaos (exponential divergence of trajectories).

What else might be done to bring dynamic theory and experimental data into contact for this particular phenomenon? There are a number of reasonable next steps: (1) Look for more evidence for the existence of a CM attractor; (2) characterize the properties of the attractor and the underlying dynamics in more detail; (3) search for evidence of structural instabilities in the CM attractor (bifurcations) and the relevant control parameters. In the next few sections we briefly explore these possibilities.

More Evidence for an Attractor

The central dilemma faced by any experimentalist hoping to apply dynamic systems theory is ignorance, in particular, ignorance of the state variables and therefore an inability to search directly in the system's state space for an attractor. Instead, we must conduct an indirect search in the state space reconstructed from experimental time series. Nevertheless, there are at least two kinds of experiments that would provide converging evidence for the existence of a CM attractor, both of which use a perturbation paradigm.

First, if typical CM does indeed reflect the existence of an attractor, then it might be possible to indirectly monitor the return to the attractor following small perturbations. For example, projections of the reconstructed phase portrait could be studied graphically for evidence that perturbed trajectories return to the original region of the reconstructed state space. It might also be possible to make rough estimates of the time taken to return, which would provide some information about the magnitude of negative Lyapunov exponents in the near vicinity of the attractor. Evidence that the correlation dimension and λ_1 had not changed (with respect to the preperturbation values) after the transient deflection and return had died out would increase our confidence that trajectories had returned to the same attractor.

A second set of experiments inducing larger perturbations in CM might reveal the boundary of the basin of attraction, and perhaps the existence of additional CM attractors representing qualitatively different modes of organization in spontaneous motor activity. For example, experiments that vary the strength of the perturbation pulse might yield evidence of a discontinuous jump to a new region of state space above a critical pulse amplitude. If it is stable, the postperturbation phase portrait could be analyzed in the same ways as the preperturbation phase portrait, by estimating the correlation dimension, the largest Lyapunov exponent, and studying the response to small perturbations. Furthermore, if a second attractor exists, one would predict return times to the preperturbation attractor to increase as the critical pulse amplitude was approached.

In principle, much detailed information about CM state space, attractors, and their basin boundaries could be obtained from systematic perturbation experiments. In reality, the number of experiments that would have to be performed in order to reveal that level of detail is likely to be prohibitive. It is more reasonable to expect that only critical (or nearly critical) experiments will actually be done. That is, experiments to test the stability of the attractor and to test for the existence of a second attractor are far more likely to be done than experiments to trace out the basin boundaries between attractors. A good dynamic model of CM, however, would guide the experimentation, and through numerical simulations might also provide useful information about regions of state space that will not be explored in real experiments.

Information about the extent and strength of CM attractors (if they exist), and their relative sensitivity to different types of perturbations (e.g., stimulation in different sensory modalities), will also be relevant to a functional analysis of CM. Detailed knowledge of differential sensitivities might suggest conditions under which the dynamic system underlying spontaneous activity could be bumped out of the basin of a high-dimensional attractor (which might facilitate extracting information from a relatively static environment) to a lower-dimensional attractor (which might permit focused attention or even entrainment by periodic environmental stimulation). The possibility that qualitatively different modes of behavioral organization, with corresponding differences in functional significance, could be accounted for in specific detail by the same underlying mechanism would be a significant contribution of dynamic theory. An important theoretical (if not practical) challenge, however, would be to distinguish such events from bifurcations induced by changing control parameters, as discussed below.

More Detailed Characterization of the Attractor

As Farmer et al. (1983) argue, an attractor's dimension is the first level of knowledge needed to characterize its properties. But as we alluded to earlier, the calculation of the correlation dimension using the method proposed by Grassberger and Procaccia (1983) results in the loss of other information. In particular, information about nonuniformity in the dimension of an attractor is lost when $N(r)$ is averaged over the attractor.

Mayer-Kress (e.g., 1986) has argued that in some cases it is better to select a set of reference points on the reconstructed attractor and estimate the local dimension at each of those reference points. The variation in local dimension among the reference points would indicate the dimensional nonuniformity of the attractor. For example, an extreme flattening and compression of trajectories in one region of the attractor would be reflected by a drop in the estimated local dimension for reference points in that region. Furthermore, the evolution of dimensional complexity can be studied by estimating local dimensions for a sequence of reference points ordered in time. Either way, information about dimensional nonuniformity may be critical in detecting structural changes in reconstructed phase portraits as a potential control parameter is varied.

The defining property of a chaotic attractor is the exponential rate at which nearby trajectories on the attractor diverge as they evolve in time (Farmer et al., 1983). This leads to the well-known sensitive dependence on initial conditions, since the inability to specify the state of the system with more than finite precision means that initially indistinguishable states are later found to be distinct. The sensitive dependence on initial conditions, therefore, is the basis for the unpredictability of a chaotic dynamic system in practice. The focus on unpredictability, although perfectly legitimate and a fundamental property of chaotic dynamics, reveals a conceptual bias. The bias is that predicting the future is what really matters.

An equivalent interpretation is that a system which exhibits a sensitive dependence on initial conditions is one that produces information (Eckmann and Ruelle, 1985). That is, states of the system that are not distinguishable at one moment evolve into states that are. The second measurement therefore yields additional information about the system. Entropy is the average rate at which information is produced by a chaotic system, and has a number of technical definitions. Fraser (1986) describes a method of calculating metric entropy from a time series using mutual information. Entropy is also related to the Lyapunov exponents (Eckmann and Ruelle, 1985), which is not surprising since they quantify the rate at which the separation of nearby trajectories changes.

If for no other reason than to correct an interpretive bias, it might be useful to calculate entropy when characterizing a CM attractor. However, in addition to quantifying the average rate of information production by the mechanism generating CM, a consideration of entropy might also suggest novel ways to view the functional consequences of CM in the real world. As discussed below, to the extent that information about the environment is acquired as the result of actions in it, a behavioral system with positive entropy may increase the information yield from exploration of the surroundings.

Finally, it may be worth considering an approach to characterizing the reconstructed phase portrait which makes fewer assumptions about the underlying dynamics and the relationship between the experimental time series and the dynamic system. Eckmann, Kamphorst, and Ruelle (1987) have proposed a graphic tool, called a recurrence plot, which yields relatively low-level (i.e., unprocessed) information about temporal patterns in the trajectories of the reconstructed phase portrait.

A recurrence plot (figure 5.8) is constructed by taking each point $\mathbf{y}_i$ in the reconstructed state space, following the trajectory from that point on, and measuring the distance (in state space) from the starting point. As the trajectory evolves, it will return to regions of the state space that are close to its starting point. The closeness of the recurrences is plotted (as shades of gray in figure 5.8) as a function of the starting point and elapsed time with respect to the starting point.

The pattern of close encounters displayed in the recurrence plot represents space-time information extracted from the reconstructed state space. This gives direct and independent information about dynamic quantities (local divergence rates) which are used in the calculation of both Lyapunov exponents and Kolmogorov-Sinai entropies. The lengths of regions representing close recurrence give direct measures of the separation of high-dimensional state-space trajectories. As a diagnostic tool, recurrence plots can also provide information about the sensitivity of the system to external perturbations. The information can be quite rich (see figure 5.8), and can reveal complicated patterns which might be useful in visualizing the effects of perturbations or parameter changes in experiments or simulations. For the present, however, the richness of recurrence plots exceeds our ability to interpret the patterns.

Structural Instability of the Attractor

Up to this point, we have focused on how to identify and characterize the image of a CM attractor (or multiple attractors) in the state space reconstructed from an experimental time series. The assumption has been that the attractor was structurally stable under normal conditions. That is, at a metaphoric level the apparently strong stationarity of CM suggested the existence of an attractor. Estimating its dimension and determining if there was evidence for the exponential divergence of trajectories took us two important steps beyond the metaphor. We have also described, in far less detail, some further steps that might be taken. Although some (e.g., perturbation experiments) are designed to probe the local stability (in state space) of the mechanism generating CM, we have not considered the structural stability of the attractor when important parameters are varied. It is to this subject of bifurcations that we now briefly turn.

Bifurcation theory (e.g., Guckenheimer and Holmes, 1983) deals with the qualitative changes that occur in the dynamics of a system as a consequence of varying parameter values. The descriptive study of bifurcations is relatively straightforward when the state variables, parameters, and differential equations defining the dynamic system are known. It is then possible to simulate the evolution of the dynamic system numerically for different parameter values and identify bifurcation points where attractors undergo structural change. For example, at a particular value of some parameter, a previously stable fixed-point attractor might become unstable and be replaced by a limit cycle (a Hopf bifurcation). That is, a system which in one parameter range has a constant output may oscillate in an adjacent range of the same parameter.

Structural stability and bifurcation are seductive metaphors to those of us in the behavioral and developmental sciences. Because we are interested in patterns and organization, and especially in how they emerge, transform, and disappear in both real and developmental time, the concept of bifurcation seems particularly well suited to our needs. Furthermore, the possibility of accounting for qualitatively different types of organization (homeostasis, oscillations, and chaotic fluctuations) and the transitions between them with the same theory is especially attractive. Finally, the notion that potentially simple, scalar quantities that change in a slow, gradual fashion can trigger massive reorganization of the whole system is almost too good to be true. It is the way in which parameter space and state space are linked by the concept of bifurcation that permits us to think of control, regulation, and the influence of measurable factors in the internal and external environments of the organism.

How can we go beyond the bifurcation metaphor? For the particular phenomenon of CM, we would first need evidence of the structural stability of an attractor. This means evidence that a CM attractor exists (as discussed in preceding sections), plus evidence that it is structurally unchanged for a range of values of some parameter. Then we would need evidence of structural instability as the parameter was varied through a critical interval, with the appearance of a qualitatively different attractor or attractors with further change in the parameter.

But now we must deal with another category of ignorance. Not only do we not know the state space of CM, we also do not know the parameter space. We could do experiments and manipulate factors that might be important based on other knowledge of the behavioral system, such as communication along the spinal cord in the fetal rat (Robertson and Smotherman, 1990), or the rate of rhythmic driving by an infant's mother during social interaction (Robertson, 1989). Such experiments could yield useful information about the structural stability and instability of CM, and might reveal different routes into and out of the (typical) chaotic dynamics for CM. Of course, any such experiments would have to be done in a state space reconstructed from movement time series without the benefit of any specific knowledge about the dynamics underlying CM. Once again, the need for a dynamic model for CM, in which parameters with behavioral or biological underpinnings are explicitly represented and can be studied directly, becomes apparent.

Mechanisms and Models

In the attempt to understand the behavioral phenomenon of CM, an important objective is to build a successful model of the underlying mechanism. What, then, is the consequence of having begun to push dynamic systems theory beyond the level of metaphoric redescription when it comes to the model-building process? The consequence is that, if the basic results are confirmed in more samples of CM, there is justification for considering models based on chaotic dynamic systems. That is, the dynamic model and its biological substrate should be low-dimensional and capable of generating complex fluctuations that are sustained and irregular; both core characteristics of CM should be inherent properties of the same, effectively simple mechanism.

One possible biological substrate is a single oscillatory network in the central nervous system which generates the chaotic fluctuations in motor output. An example of a single chaotic neural oscillator has been described by Mpitsos et al. (1988) in the buccal ganglion of the sea slug *Pleurobranchaea californica*. Although a single source would be parsimonious, it is not supported by the available data. It was recently shown that the fetal rat near the end of gestation exhibits fluctuations in spontaneous motor activity which are quantitatively similar to human fetal CM in terms of their spectral properties (Smotherman, Robinson, and Robertson, 1988). When the spinal cord is transected in a midthoracic location, movements of the rear legs are uncoupled from other body and limb movements, but both rear leg and non-rear leg activity continue to exhibit the temporal organization characteristic of CM (Robertson and Smotherman, 1990). These results rule out a single source of oscillations in the rat fetus. They are in other respects also consistent with the results of a noninvasive perturbation experiment with human neonates (Robertson, in press) and suggest that rostral sources may be slower than caudal sources, with the latter normally dominant.

Another possible biological substrate is two or more oscillatory networks which alone are quite stable but which are coupled together in such a way that their combined interactions produce a chaotic output. Neural oscillators are well known to produce stable rhythmic movements in vertebrate nervous systems (Cohen, Rossignol, and Grillner, 1988). An example of a group of stable oscillators that produce an apparently chaotic motor output by virtue of their coupling has been described by Cohen, Baker, Guan, and Kiemel (1990) in lampreys subjected to partial spinal cord transections. Following regeneration of the transected fiber tracts, the swimming motor pattern produced by the spinal cord becomes highly unstable. The pattern can be easily reproduced by assuming that the intersegmental coordinating system has only been partially restored, an assumption well justified by direct experimental evidence (Cohen, Baker, and Dobrov, 1989). The modeling demonstrated quite clearly how neural oscillators could be coupled together to produce chaotic motor patterns.

The data obtained from the fetal rats with spinal cord transections (Robertson and Smotherman, 1990) were not subjected to the kinds of analyses illustrated in this chapter. However, spectral analysis indicated that the fluctuations generated above and below the transection continued to be irregular. Thus it is unlikely that CM is the consequence of only two oscillators that are independently stable but yield chaotic output when coupled. The possibility remains, however, that there are multiple sources distributed throughout the motor system which interact to produce CM. It is also possible that there are no localized sources, but that the fluctuations in motor activity emerge only at the macroscopic level.

We clearly know very little about the biological substrate of CM. What we do know puts some constraints, such as the requirement for more than one source, on any specific dynamic model that might eventually be developed. But the guidance provided by our minimal knowledge is not great. We are therefore a long way from the goal of building a dynamic model of CM in which the state variables and parameters have a clear correspondence with psychobiological and environmental factors. Thus the obvious benefits of such a model described throughout this chapter in (1) searching for multiple attractors and the boundaries between their basins of attraction, (2) predicting how various intrinsic or extrinsic perturbations or parameter changes might qualitatively change the organization of CM, and (3) suggesting ways in which a specific chaotic dynamic system might have functional consequences for how infants act on the world and extract information from it, are all yet to be realized.

Conclusion

The point of this chapter has been to illustrate what might be involved in taking our seduction by dynamic systems theory a few steps beyond the mere redescription of what we already know. We have argued that it is necessary to examine specific phenomena as closely as possible, to bring the theory into contact with the data, and to look for concrete evidence that the fundamental explanatory concepts provided by dynamic systems theory are appropriate.

These are, however, just the preliminaries. It is encouraging to think that a theory with successful applications across diverse scientific disciplines might be brought to bear on questions of behavior and development. The real test, however, is not whether dynamic systems theory can account for the phenomena we study, but whether it yields any new insights, integrates previously unrelated empirical facts, or in some other way leads to a deeper understanding of those phenomena.

We illustrated in some detail how the application of dynamic systems theory could be taken beyond the metaphor for the specific phenomenon of human CM. More extensive work is of course needed to make this more than an illustration. What might be the fundamental insight gained in this process for the specific case of CM? The fundamental insight, with evidence to support it, would seem to be that the complex fluctuations in spontaneous motor activity can be explained by a dynamic system with relatively few degrees of freedom, and that both the sustained presence of oscillations and their irregularity might be inherent properties of the same, effectively simple mechanism. This is in contrast to our usual tendency to think of regularity as the manifestation of a lawful underlying mechanism and irregularity as the consequence of stochastic inputs to the mechanism.

A significant consequence of this insight into the mechanism of CM is that it reorients our thinking about the possible functional benefits (or costs) of CM. We are led to consider the adaptive consequences of unpredictability built in to an action system that regulates the young infant's interaction with the physical and social environment. A common assumption, of course, is that the adaptive value or utility of behavioral organization lies in its predictable components.

There is no question that dynamic systems theory can provide us with rich metaphors in our attempt to understand behavior and development. Although we face severe limitations imposed by our ignorance of the systems we study and our inability to measure them accurately over long periods of time, we should nevertheless try, whenever possible, to go beyond metaphor in our analysis of specific phenomena in order to put any new insights on firm ground.

References

Abraham, R. H., and Shaw, C. D. (1982). *Dynamics—The geometry of behavior. Part 1. Periodic behavior.* Santa Cruz, CA: Aerial Press.

Abraham, R. H., and Shaw, C. D. (1983). *Dynamics—The geometry of behavior. Part 2. Chaotic behavior.* Santa Cruz, CA: Aerial Press.

Abraham, R. H., and Shaw, C. D. (1985). *Dynamics—The geometry of behavior. Part 3. Global behavior.* Santa Cruz, CA: Aerial Press.

Abraham, R. H., and Shaw, C. D. (1987). Dynamics a visual introduction. In F. E. Yates (Ed.), *Self-organizing systems* (pp. 543–597). New York: Plenum.

Abraham, R. H., and Shaw, C. D. (1988). *Dynamics—The geometry of behavior. Part 4. Bifurcation behavior.* Santa Cruz, CA: Aerial Press.

Albano, A. M., Mees, A. I., DeGuzman, G. C., and Rapp, P. E. (1987). Data requirements for reliable estimation of correlation dimensions. In H. Degn, A. V. Holden, and L. F. Olsen (Eds.), *Chaos in biological systems* (pp . 207–220). New York: Plenum.

Cohen, A. H., Baker, M. T., and Dobrov, T. A. (1989). Evidence for functional regeneration in the adult lamprey spinal cord following transection. *Brain Research, 496,* 368–372.

Cohen, A. H., Baker, M. T., Guan, L., and Kiemel, T. (1990). Neural regeneration leads to dysfunction in lamprey spinal cord. Society of Neuroscience Abstract No. 211.1.

Cohen, A. H., Rossignol, S., and Grillner, S. (Eds.) (1988). *Neural control of rhythmic movements in vertebrates.* New York: Wiley.

Corner, M. A. (1977). Sleep, and the beginnings of behavior in the animal kingdom—studies of ultradian motility cycles in early life. *Progress in Neurobiology, 8,* 279–295.

de Vries, J. I. P., Visser, G. H. A., and Prechtl, H. F. R. (1982). The emergence of fetal behaviour. I. Qualitative aspects. *Early Human Development, 7,* 301–322.

Drachman, D. B., and Sokoloff, L. (1966). The role of movement in embryonic joint development. *Developmental Biology, 14,* 401–420.

Eckmann, J. P., Kamphorst, S. O., and Ruelle, D. (1987). Recurrence plots of dynamic systems. *Europhysics Letters, 4,* 973–977.

Eckmann, J. P., and Ruelle, D. (1985). Ergodic theory of chaos, and strange attractors. *Review of Modern Physics, 57,* 617–656.

Farmer, J. D., Ott, E., and Yorke, J. A. (1983). The dimension of chaotic attractors. *Physica, 7D,* 153–180.

Fraser, A. M. (1986). Using mutual information to estimate metric entropy. In G. Mayer-Kress (Ed.), *Dimensions, and entropies in chaotic systems* (pp. 82–91). Berlin: Springer-Verlag.

Fraser, A. M., and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A, 33,* 1134–1140.

Gershenfeld, N. (1988). An experimentalist's introduction to the observation of dynamic systems. In Hao Bail-lin (Ed.), *Directions in chaos* (pp. 310–381). Singapore: World Scientific.

Glass, L., and Mackey, M. C. (1988). *From clocks to chaos: The rhythms of life.* Princeton, NJ: Princeton University Press.

Gould, S. J., and Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London, B205,* 581–598.

Grassberger, P., and Procaccia, I. (1983). Characterization of strange attractors. *Physical Review Letters, 50,* 346–349.

Grebogi, C., Ott, E., Pelikan, S., and Yorke, J. A. (1984). Strange attractors that are not chaotic. *Physica, 13D,* 261–268.

Guckenheimer, J. (1982). Noise in chaotic systems. *Nature, 298,* 358–361.

Guckenheimer, J. (1984). Dimension estimates for attractors. *Contemporary Mathematics, 28,* 357–367.

Guckenheimer, J., and Holmes, P. (1983). *Nonlinear oscillations, dynamic systems, and bifurcations of vector fields.* New York: Springer-Verlag.

Haken, H. (1983). *Synergetics: An introduction,* 3rd ed. Berlin: Springer-Verlag.

Hamburger, V., Balaban, M., Oppenheim, R., and Wenger, E. (1965). Periodic motility of normal and spinal chick embryos between 8 and 17 days of incubation. *Journal of Experimental Zoology, 159,* 1–14.

Harris, W. A. (1981). Neural activity and development. *Annual Review of Physiology, 43,* 689–710.

Jenkins, G. M., and Watts, D. G. (1968). *Spectral analysis and its applications.* San Francisco: Holden-Day.

Kaplan, J. L., and Yorke, J. A. (1979). Preturbulence: a regime observed in a fluid flow model of Lorenz. *Communications in Mathematical Physics, 67,* 93–108.

Koslow, S. H., Mandell, A. J., and Shlesinger, M. F. (1987). Perspectives in biological dynamics and theoretical medicine. *Annals of the New York Academy of Sciences, 504.*

Kurths, J., and Herzel, H. (1987). An attractor in a solar time series. *Physica, 25D,* 165–172.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Science, 20,* 282–291.

Mandelbrot, B. (1982). *The fractal geometry of nature.* San Francisco: Freeman.

Mayer-Kress, G. (Ed.) (1986). *Dimensions and entropies in chaotic systems.* Berlin: Springer-Verlag.

Mayer-Kress, G. (1987). Application of dimension algorithms to experimental chaos. In Hao Bai-lin (Ed.), *Directions in chaos* (pp. 122–147). Singapore: World Scientific.

Mayer-Kress, G., and Kaneko, K. (1989). Spatiotemporal chaos and noise. *Journal of Statistical Physics, 54,* 1489–1508.

Mayer-Kress, G., Yates, F. E., Benton, L., Keidel, M., Tirsch, W., Poppl, S. J., and Geist, K. (1988). Dimensional analysis of nonlinear oscillations in brain, heart, and muscle. *Mathematical Biosciences, 90,* 155–182.

Mpitsos, G. J., Creech, H. C., Cohan, C. S., and Mendelson, M. (1988). Variability and chaos: Neurointegrative principles in self-organization of motor patterns. In J. A. S. Kelso, A. J. Mandell, and M. F. Shlesinger (Eds.), *Dynamic patterns in complex systems* (pp. 162–190). Singapore: World Scientific.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time series. *Physical Review Letters, 45,* 712–716.

Prechtl, H. F. R. (1984). Continuity and change in early neural development. In H. F. R. Prechtl (Eds.), *Continuity of neural functions from prenatal to postnatal life. Clinics in Developmental Medicine No. 94* (pp. 1–15). Philadelphia: JB Lippincott.

Rapp, P. E., Zimmerman, I. D., Albano, A. M., de Guzman, G. C., Greenbaun, N. N., and Bashore, T. R. (1986). Experimental studies in chaotic neural behavior: Cellular activity, and electroencephalographic signals. In H. G. Othmer (Ed.), *Nonlinear oscillations in biology and chemistry* (pp. 175–205). Berlin: Springer-Verlag.

Robertson, S. S. (1982). Intrinsic temporal patterning in the spontaneous movement of awake neonates. *Child Development, 53,* 1016–1021.

Robertson, S. S. (1985). Cyclic motor activity in the human fetus after midgestation. *Developmental Psychobiology, 18,* 411–419.

Robertson, S. S. (1987). Human cyclic motility: Fetal-newborn continuities and newborn state differences. *Developmental Psychobiology, 20,* 425–442.

Robertson, S. S. (1988). Infants of diabetic mothers: Late normalization of fetal cyclic motility persists after birth. *Developmental Psychobiology, 21,* 477–490.

Robertson, S. S. (1989). Mechanism and function of cyclicity in spontaneous movement. In W. P. Smotherman and S. R. Robinson (Eds.), *Behavior of the fetus* (pp. 77–94). Caldwell, NJ: Telford.

Robertson, S. S. (in press). Probing the mechanism of oscillations in newborn motor activity. *Developmental Psychology.*

Roberston, S. S., and Dierker, L. J. (1986). The development of cyclic motility in fetuses of diabetic mothers. *Developmental Psychobiology, 19,* 223–234.

Robertson, S. S., and Smotherman, W. P. (1990). The neural control of cyclic motor activity in the fetal rat. *Physiology and Behavior, 47,* 121–126.

Schaffer, W. M., Truty, G. L., and Fulmer, S. L. (1989). *Dynamical software,* Vol. I.4.I and II.2.I. Tucson, AZ: Dynamical Systems.

Schuster, H. G. (1988). *Deterministic chaos: An introduction,* 2nd rev. ed. New York: VCH Publishers.

Skarda, C. A., and Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences, 10,* 161–195.

Smotherman, W. P., Robinson, S. R., and Robertson, S. S. (1988). Cyclic motor activity in the rat fetus. *Journal of Comparative Psychology, 102,* 78–82.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand, and L. S. Young (Eds.), *Dynamical systems and turbulence* (pp. 366–381). New York: Springer-Verlag.

Thompson, J. M. T., and Stewart, H. B. (1986). *Nonlinear dynamics and chaos.* New York: Wiley.

Thompson, W. (1983). Synapse elimination in neonatal rat muscle is sensitive to pattern of muscle use. *Nature, 302,* 614–616.

West, B. J., and Goldberger, A. L. (1987). Physiology in fractal dimensions. *American Scientist, 75,* 354–365.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica, 16D,* 285–317.

Wolf, A., and Vastano, J. A. (1986). Intermediate length scale effects of Lyapunov exponent estimation. In G. Mayer-Kress (Ed.), *Dimensions and entropies in chaotic systems* (pp. 94–99). New York: Springer-Verlag.