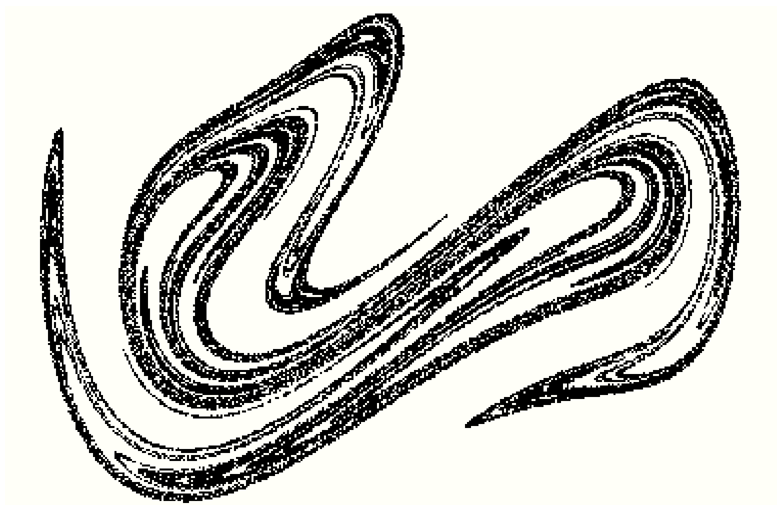
CONTEMPORARY
SIGNAL PROCESSING



RANDLE, INC.
P.O. Box 1010
Great Falls, VA 22066
(703) 759-5257

SUMMARY

- **Chaotic signals, man made and naturally occurring, appear to be broad-band random noise to traditional signal processors.**

- **Many man-made signals are not inherently chaotic, but have broad-band behavior.**

- **Techniques for signal processing derived from the study of chaos provide new approaches for signal detection, signal classification, noise reduction, and secure communications.**

- **Noise may be a deterministic signal from a higher dimension that corrupts data observed in a lower dimension. Chaotic signal processing techniques often can be used to mitigate noise.**

- **Secure communications based on chaotic signals is attractive because the carrier signal appears to be noise to traditional signal detection equipment. Even if an adversary knows a signal is present and knows the signal-generating function, interception is still difficult.**

- **Our unique processor incorporates several novel techniques including:**

  - **An easy and fast method of finding a signal's dimension (degrees of freedom),**

  - **An information theoretic approach to detect signals of unknown and arbitrary form**

  - **A technique for noise reduction and secure communications called Probabilistic Scaled Cleaning.**

TABLE OF CONTENTS

January 1994

STATEMENT OF PROPRIETARY INTEREST

The techniques discussed in this paper are in the scientific literature, as are many other methods. We ask that the use of these specific techniques by Randle, Inc. and the data we have processed be treated as proprietary.

Commercial

The information contained herein constitutes valuable trade secrets. This document shall not be duplicated, used, or disclosed to any person other than the recipient without the written permission of Randle, Inc.

Government

All government personnel must exercise extreme care to insure that the information in this proposal is not disclosed to an individual who has not been authorized access to such data in accordance with 3.104, and is not duplicated, used, or disclosed in whole or part for any purpose other than evaluation, without the written permission of the offeror. If a contract is awarded on the basis of this document, the terms of the contract shall control disclosure and use.

This notice does not limit the government's right to use information contained in the proposal if it is obtainable from another source without restriction.

This is a government notice, and shall not by itself be construed to impose any liability upon the government or government personnel for disclosure or use of data contained in this proposal.

FAR 15.509

INTRODUCTION

One goal of signal processing is to identify signals by invariant parameters. Identification of signals traditionally has relied upon detection of discrete frequencies that can be categorized by Fourier analysis of sensor data. Some parameters that classify the emitter are the Fourier coefficients of the signal.

Fourier analysis, a linear signal-processing technique, relies on assumptions that simplify models of signals, clutter and noise. These idealizations state that the signal and its environment are one degree-of-freedom systems, they are perfectly known, the signal is a pure sinusoid, and the noise is perfectly random. For example, when processing manmade radiated energy, the received waveform is assumed to be a replica of an ideal transmitted linear waveform. A matched filter detector (and the "FFT") recovers the signal by correlating it with a linear model of the signal. Additionally, the noise is assumed to be Gaussian with a known spectrum. The signals of interest either have a narrow-band Fourier spectrum or have a Fourier spectrum that otherwise distinguishes the signal from the noise. The linear approach to noise rejection is to average out the seemingly random components of the data. Linear signals are well understood, well exploited, easily countered, and are not the subject of this paper. Still, we are interested in nonlinear effects that corrupt otherwise linear signals. Fourier analysis is useless for analyzing most nonlinear systems.

Nonlinear science, often called chaos theory, has matured to the point where the signals-processing community can benefit from this new understanding of complex phenomenology. New signal analysis tools derived from the study of chaotic systems and their signals allow the characterization of complex signals, clutter, and noise. **The nonlinear approach is to treat both the signal and the noise as deterministic processes in a multidimension (N-degree-of-freedom) geometric space. No assumptions are made about a signal.** Signals are geometric objects called attractors. They are categorized by invariant geometric and information conveying properties of the attractor. Chaotic signal processing can detect signals of arbitrary (nonsinusoidal) waveform and reduce noise for linear signal detection. It also may be possible to identify low signal-to-noise ratio chaotic signals that are a by product of other emissions.

Randle, Inc., working with the Institute for Nonlinear Science (University of California - San Diego) has developed an integrated software system for detecting and analyzing chaotic signals. This system includes a module that performs Probabilistic Scaled Cleaning (PSC). PSC is a recurrsive cleaning method that relies on the difference in characteristics between a signal (even if it is chaotic) and noise. A signal occupies well-defined regions of state space (i.e., it resides on an attractor), while noise fills the state space according to its distribution.

The novel aspect of chaotic signal processing is that signals are analyzed in a time domain state space, rather than in a frequency domain. **The new and remarkable idea is that systems with only a few degrees of freedom can be chaotic and that it is possible to reconstruct and analyze chaotic behavior with scalar data collected in only one dimension.**

Manmade chaotic signals exist. Chaotic acoustic signals may occur from normal machinery operation or may be associated with transients. Examples include air conditioning noise and pump operations. Chaotic electromagnetic signals can be intentionally generated.

Other types of man-made chaotic signals include ship acoustics, fluid turbulence, flow noise, plasma fusion, electronic circuits, simple neural networks, structural vibrations, vehicle dynamics, and dripping faucets. Other types of manmade signals have broad-band spectrums and can be exploited with these methods. Many types of digital communications signals have a *sine(x)/x* spectrum and are candidates for exploitation.

CHAOTIC BEHAVIOR IS:

Apparently random
Deterministic
Broadband
Rapid divergence from near identical points
(Sensitive to initial conditions)
Everywhere

The techniques for processing chaotic signals are now well defined. Much work has been done using these methods on the traditional aspects of signal processing -- detection, characterization, and classification of the invariant properties of signals. A remarkable result of this experience is that a new view of "noise" is emerging. The deterministic signal from a chaotic system can appear to be broad-band noise when processed with traditional schemes such as Fourier analysis or matched filters. Thus, when we observe or hear "noise" we wonder if the "no signal" signal is complex and intractable. Or, are there components that are only a dimension or two higher than the signal processor's domain? So, what is noise? How can a signal that masquerades as noise be separated from other "noise?" And, can messages be encoded onto and separated from "noise?"

Chaotic signal processing is only one tool for analyzing signals. Fourier analysis, high-order spectra, wavelets, and chaos theory are complementary. Chaotic techniques can characterize all classes of signals, but may provide little additional information about signals that are well behaved in other processors:

Signal Processing Paradigms			
Type of signal	Linear	Nonlinear	Chaotic
Definition	Superposition applies, homogeneous	No superposition, inhomogeneous	No superposition, inhomogeneous
Characteristics	Discrete tonals	Multiple, coupled wavelengths	Broad-band
Detection method	Fourier analysis	Higher order spectra	Time delay state space reconstruction
Signal descriptors	Wavelength of tonals	Wavelength pairings	Attractor characteristics, Lyapunov exponents

Chaotic signal processing often works when linear techniques, including polyspectra, break down. A current belief is that new submarines, for example, have become so quiet that detection of tonals by passive sonar may be of little use. Passive detection and identification of submarines has traditionally relied on detection of discrete tonals where target-radiated spikes of acoustic energy in the Fourier spectrum were well above ocean ambient noise. Modern quieting techniques have reduced the energy in these discrete tonals to well below the level of ambient noise. Research has shown that there are radiated signals that can only be detected by chaotic techniques and not by other means.

Besides the detection and classification of signals, these techniques may also be used to reduce noise. The invariant properties of a signal are determined and iterative techniques are used to generate a clean signal that has the same properties as the reference signal. This scheme allows removal of chaotic noise from other signals or the removal of high-dimension noise from chaotic signals. Noise reduction techniques based on chaotic signal processing may provide higher signal-to-noise ratios at much less expense than methods that operate directly on the signal.

Noise reduction techniques can also be used as the basis for secure communications. A message, in the clear or encoded, is added to a chaotic carrier. The transmitted signal is broad-band noise to traditional signal detection equipment. The authorized receiver uses the noise reduction system to strip away the chaotic carrier, leaving the message. The unauthorized receiver is faced with the problem of determining whether a signal is even present, what model is used for the chaotic carrier, and what system parameters were used to generate the signal. Development of new methods for covert communications is important because spread-spectrum techniques are, ironically, easily detected with chaotic signal processing.

THE FUNDAMENTAL THEOREM

The fundamental operation that makes chaotic signal processing possible is the reconstruction of an N-dimension geometric object, called an attractor, from a single time series.

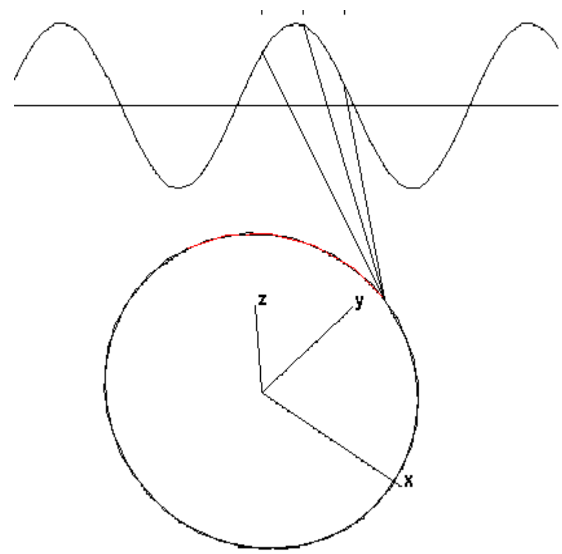
The scalar time series is represented as:

$$x(n), x(n+1), x(n+2), x(n+3), x(n+4), x(n+5), x(n+6) \ldots x(n+N-1)$$

where n is the sample number (usually treated as 0), N is the total number of samples, and x(k) is the scalar value of the sampled signal at k.

To form the reconstructed attractor, this scalar time series is mapped to a vector time series of the form:

$$y(n) = [x(n), x(n+T), x(n+2T), \ldots x(n+(d-1)T)].$$



For any geometric dimension, *d*, the vector *y(n)* will have d components that are comprised of d samples in the original series, separated by a "time delay", *T*. The time delay is sometimes called the time lag. The method for selecting *d* and *T* will be discussed later, but, for the purposes of the fundamental theorem, any *d* and *T* will suffice.

The time lagged vectors are plotted in the geometric ("phase" or "state") space. **The principle that makes chaotic signal processing possible is that there is a unique mapping between *d* degree of freedom vector sets and vectors reconstructed from the time history of scalar data collected in only one dimension.** Counter intuitive, but true. The original signal vectors are not reconstructed. The reconstruction yields a unique new vector set that captures the behavior of the dynamic system. Because the process is invertible, it is an example of a diffeomorphic transform. Figure 1 shows time lagged vectors for a nonchaotic process embedded in a three dimension state space. The attractor for this signal is simple and, to the trained eye, is a sine wave. It will never leave the trajectory shown.

Signals represented in this manner evolve in both time and space. The signal is studied and exploited by analyzing the geometry of the reconstructed attractor and the rate at which information about the signal state is lost.

The techniques for processing chaotic signals are rapidly evolving. A short overview of the current state-of-the-art follows. Each step of the process is identified by the name of the scientist who is credited with the development of the technique and the publication date. These techniques are discussed at length in the following sections.

PROCESSING OVERVIEW

A. *Acquire a string of scalar numbers.*

Despite the degrees-of-freedom of the underlying dynamic system, usually one can collect data from only a single sensor. That is, information from one dimension is usually all that is available. The novel aspect of chaotic signal processing is to use information embedded in this one-dimension time series to capture multidimension chaotic behavior.

Acquisition issues that are important to chaotic signal processing include the sampling rate, the band width, the time over which collection is continuous, the characteristics of the collection system, and the response curve of the sensors.

B. *Perform a Fourier Analysis,* (Fourier: 1807)

An FFT may provide some clues about dominate periods and the bandwidth of the signal. Additional information about nonlinear but nonchaotic dynamics may be obtained from a polyspectral analysis. Then again, these techniques may provide little or no information.

C. *Find the embedding time delay, T*, (Fraser and Swinney: 1986)

The Average Mutual Information, $I(T)$, is a prescription for selecting an appropriate time delay interval for construction of the time lagged vectors that will be embedded in d_E. Mutual information determines how much information about subsequent data is embedded in a given datum. Use of average mutual information to determine the time delay provides better estimates of other signal characteristics later in processing. Average mutual information also is a method for detecting a signal.

D. *Find the minimum embedding dimension,* d_E, (Kennel, Brown, and Abarbanel: 1991)

The minimum dimension of the state space is found by counting the number of close points (near neighbors) in each dimension that move to more distant regions of the attractor as dimensions are added. The minimum embedding dimension is the dimension where the number of false neighbors drops to some small value.

E. *Embed the time delay vectors in a* d_E *state space* (Packard, et. al., Mane, Takens: 1981).

This remarkable operation, while counter intuitive, is easily derived. Phase space representations of systems have been used for 100 years. A phase space representation of a system plots various system values (the system state vectors) on different axis. A typical phase space plot might show displacement, velocity, and acceleration. One needs one axis for each independent variable to fully capture the behavior of a system. But, the problem for most real world data collection is that often only one variable can be collected. The challenge is to reconstruct the phase portrait, or a proxy for it, from this one data series.

The fundamental theorem demonstrates that time delay vectors built from a single series is, in fact, a proxy for the vectors for a perfectly known system.

F. *Find the local embedding dimension,* d_l, (Kennel, and Abarbanel: 1993)

The global embedding dimension is necessary to completely unfold the attractor. However, local evolutions may be adequately described in fewer dimensions. This local dimension, $d_L \leq d_E$, quantifies the degrees-of-freedom that captures motion on the attractor. Thus, models for prediction or control may also be constructed in d_L dimensions. The test for local false nearest neighbors finds, in each $d < d_E$, the percentage of nearest neighbor pairs that move far apart over some prediction horizon. The dimension where the percentage of bad predictions becomes independent of the global dimension is d_L.

G. *Compute fractal dimension of the attractor, d_A* (Ruelle: 1983).

The fractal dimension, d_A, of the attractor is one method of classifying a process and provides information on how much of the state space is filled by the system. For the sine wave of Figure 1, the fractal dimension is precisely 1.0. Because this signal has an integer fractal dimension, it is not chaotic. A chaotic signal has, by definition, a noninteger fractal dimension and the attractor is called a "strange" attractor (Ruelle, 1976).

H. *Develop the map.*

The map is a function that moves a vector in state space to the next vector as a function of time by using a local polynomial. The Taylor series expansion of the polynomial is the Jacobian of the underlying dynamics. Recent approaches retain the higher order terms to better fit local curvatures in the attractor (Abarbanel: 1989).

Noise corrupts the local Jacobian and can affect the accuracy of the calculation of the Lyapunov exponents (Abarbanel: 1990). The effect of noise is to blur the lower dimension attractor. Iterative techniques may be used to refine estimates of the local mapping function and then pull the measured data back onto the attractor (Hammel: 1990). The recovered attractor can then be exploited.

I. *Compute Global Lyapunov exponents* (Eckmann, Kamphorst, Ruelle and Ciliberto: 1986, Abarbanel: 1989).

The Lyapunov exponents describe the rate at which close points diverge and are a measure of predictability. If one or more Lyapunov exponents are positive, the system is chaotic. A QR matrix decomposition of the Jacobian gives Lyapunov exponents for the system. The Lyapunov exponents are invariant with respect to initial conditions. Therefore, they are another way of classifying a chaotic system. Global Lyapunov exponents are an invariant because they describe the effect of infinitesimal perturbations over infinite time.

J. *Compute Local Lyapunov exponents* (Abarbanel, and Sushchik: 1993).

Local Lyapunov exponents govern the growth of small deviations for a finite time and describe variations in predictability. As noted in the discussion of local false neighbor testing, the minimum embedding dimension, d_E, may be larger than the actual local dimension of the underlying dynamics, d_L.

Another method for determining d_L is to use the local Lyapunov exponents (Abarbanel and Sushchik: 1992) which is based on earlier observations that when time is reversed the true Lyapunov exponents reverse sign while false exponents do not change sign Eckmann and Ruelle, and Parlitz). The false exponents result from over embedding.

Spurious exponents and their elimination

Lyapunov exponents for dimensions above d_E are spurious. In fact, exponents for dimensions greater than the degrees-of-freedom that are active locally are also spurious. Identification of these unwanted, misleading, inaccurate, and meaningless parameters is accomplished by reversing the time series and computing the Lyapunov exponents for both the forward and backward time evolutions.

The signs of the exponents for the reverse time series are changed and the values are compared to the forward pass. The true exponents for each dimension will now be identical (or very close) for both the forward and reverse pass. The spurious exponents will have different values.

Chaotic System Invariants

Dimension of embedding space, d_E
Local embedding dimension, d_L
Fractal dimension of the attractor, d_a
Time delay for state space reconstruction, T
Global and local Lyapunov Exponents, λ_i

K. *Noise Reduction*

In signals that have a chaotic (or noiselike) components, we rely on the low dimensional, deterministic, nature of the signal to distinguish it from other signals that contaminate it. There are three methods for separating a chaotic or noiselike signal *C(n)* that has additive contamination *z(n)*:

The dynamics are known,
A "clean" signal has been measured, or
Only a contaminated signal *C(n)* + *z(n)* is measured.

If the dynamics are known and we have knowledge of the actual mapping that evolves the system in d-dimensional space: x - F(x), we can then use the properties of the vector field F(x) to identify stable, unstable, and neutral manifolds. The method of Hammel or its generalizations can then be used to separate the signal at each time step from the contamination *z(n)* when the signal-to-noise ratio is a few to ten percent.

In the second case, if a "clean" signal from the system has been observed, or can be derived, we can use this "reference" orbit $C_r(n)$ to establish the statistics of the system on its attractor. From the attractor in d-dimensional embedding space, we can find the invariant density $p(x)$ and the Markov transition probabilities $p_t(x,y)$ for the system to be in state space volume dx around x at time n and in dy around y at time $t+n$. Using these two quantities there are at least two algorithms that allow one to separate the signal from the contamination when another observation is made. The signal $C(n)$ and the reference $C_r(n)$ come from the same dynamical system, but have different initial conditions. Thus, they have entirely different orbits on the system attractor. Their statistics are the same, however. So, knowledge of $C_r(n)$ can be used to identify C(n). The Hidden Markov Model method and the Probabilistic Cleaning method are cruder than the first technique discussed but are more robust. They have been demonstrated in cases where the signal-to-noise ratio is as low as -20 dB, and there is no reason to believe they will not work at lower SNRs within the limits of numerical accuracy.

The final case is where only the contaminated signal is available. A reference orbit is not available and the system dynamics are not known. This a risky proposition, but there is some evidence, based on the work of Abarbanel, et. al., that even if the chaotic contamination-to-signal ratio is high (i.e., the SNR is very low), the Probabilistic Scaling method can isolate and remove the chaotic component, leaving the signal z(n). At present, this is unexplored territory. This will be one issue that we are exploring with both Probabilistic Scaling and Hidden Markov modeling to find the limits on "self cleaning." The self cleaning can easily be spoofed, and we are exploring methods for that as well.

APPLICATIONS

Detection and Classification of Signals and Noise Reduction

The most common signal-processing function is the detection and classification of signals. Often, just knowing that a signal is present is a major objective. Chaotic techniques can detect signals of arbitrary waveform and signals that are noiselike. These techniques also can detect signals that are intended to be covert, such as spread-spectrum signals. Foreign literature[1] discusses techniques that generate signals that cannot be detected by Fourier or higher order spectra. These signals can be detected by state space embedding methods.

The primary signal-detection tools are the nearest false neighbor embedding method, average mutual information, and naive state space embedding. Often, a low dimension signal can be detected simply by selecting an arbitrary time lag and reconstructing a three-dimension attractor. This can be done very quickly because the operations are all arithmetic.

As discussed earlier, after a signal is detected, it is categorized by its invariants.

Modeling

These methods are useful in the construction, verification, and analysis of other models. Examination of the output of highly complex models may verify the complexity of the model or

[1] V. P. Ponomarenko, "Capture Region in a Nonlinear System for Filtering of a Pseudorandom Signal with an Arbitrary Manipulation Angle," *Radio Technika i Electronica*, **20** 11 (1975) pp. 53-58 (USSR publication).

aid in its simplification. The output of one high degree-of-freedom model has been shown to be a much lower degree-of-freedom system, suggesting the original model can be greatly simplified.

Covert Communications

The PSC method works remarkably well when the signal of interest is a binary sequence (0's and 1's or the equivalent ±1). If a message is encoded or encrypted for IFF (Identification Friend or Foe), for example, as a binary sequence then this message can be added to a chaotic signal with a known reference orbit where the CSR (defined earlier) is large, then the IFF signal would be masked by the broadband carrier.

The intended recipient can easily strip off the chaotic carrier and read the binary IFF sequence. This process can be done in near-real time on existing general purpose computers. Thus, it lends itself well to simple stand alone implementation on a chip. The interrogator would send a brief signal that essentially asks for broadcast of the IFF signal. The respondent could then transmit the IFF signal (binary sequence plus the chaotic carrier) while remaining covert.

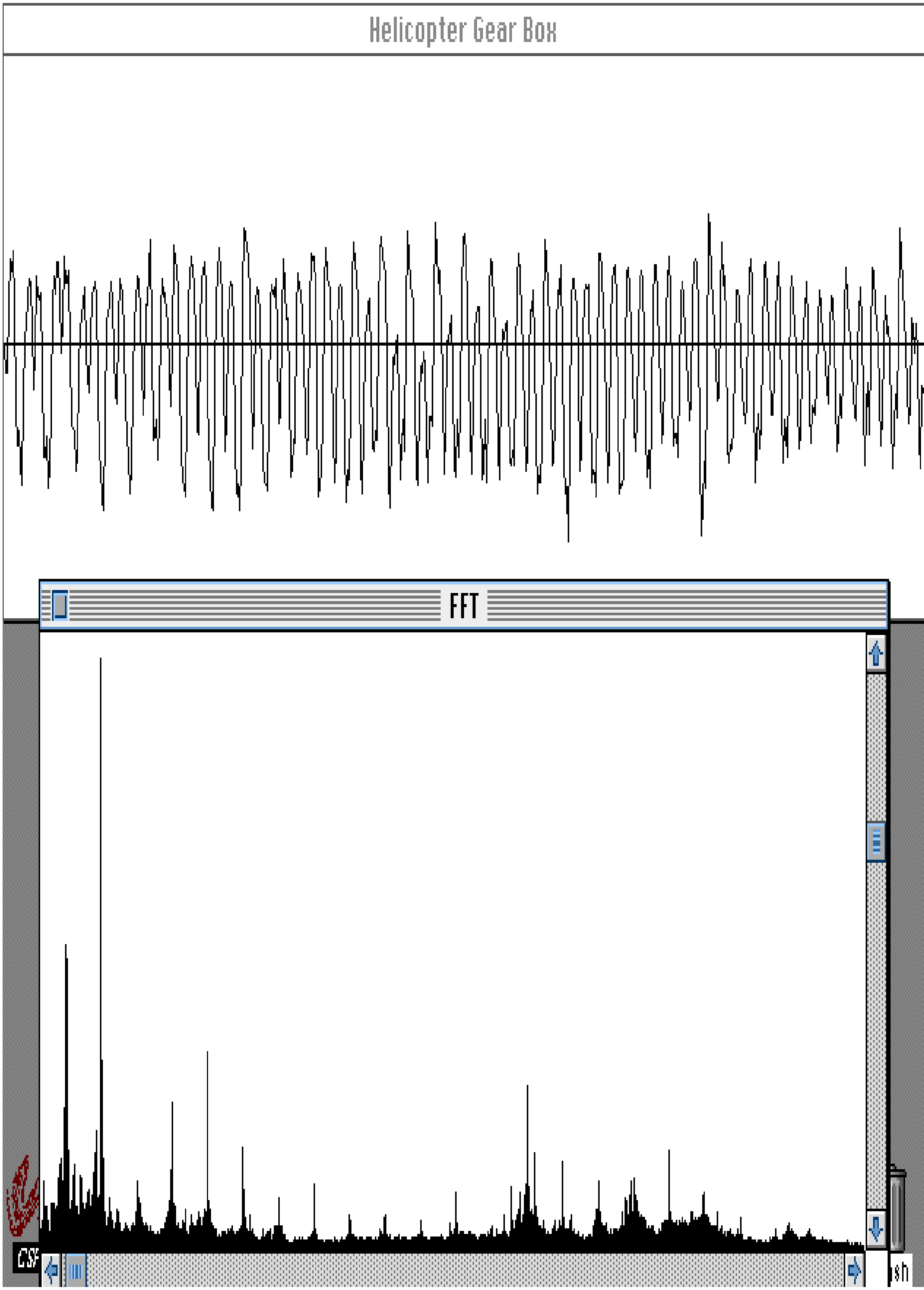
A chaotic carrier from any model system should work. Conceptually, a long sequence of a naturally occurring chaotic signal also can be used if its chaotic characteristics are known. Some issues that we are exploring are the desirable SNR (or, to use our terminology, the SCR) of the encoding on the carrier, and the effect of the encoding on the FFT of the signal.

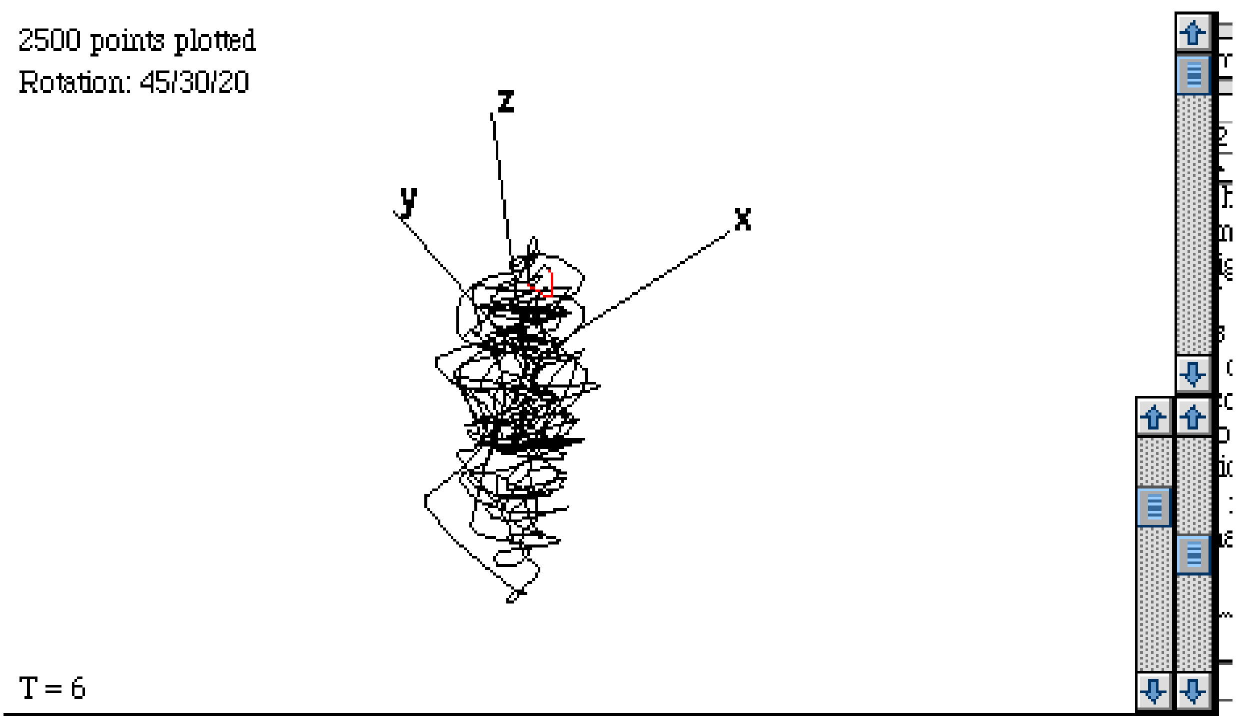
Machinery Maintenance

These signal-processing techniques may be able to discriminate small changes in the vibration of machinery due to incipient faults. An ability to detect impending machinery failure will reduce maintenance costs and will probably reduce fatality and injury rates.

If the signals from these devices can be characterized by an attractor, then the issue becomes one of determining if the attractor changes when a fault develops and how those fault-induced changes affect the invariants that characterize the attractor. If so, then incipient faults may be detected by a change in the geometric description of the attractor. Figure 2 shows the signal and FTT from an accelerometer attached to a high speed gear box. These signals are not sinusoidal, although they do exhibit a strong periodicity that is driven by cyclical processes. While there is a predominate period, for the most part the signal appears to be broad band noise. But, we can show that this signal is deterministic and can be characterized by a "strange" attractor (figure 3). Another study[2] has demonstrated that helicopter airframe vibrations are chaotic and can be characterized by an attractor.

[2]Sarigul-Klijn, M. "Application of Chaos Methods to Helicopter Vibration Reduction Using Higher Harmonic Control," Ph.D. Dissertation, Naval Postgraduate School (March 1990).


Helicopter Gear Box
FFT


2500 points plotted
Rotation: 45/30/20
z
y
x
T = 6

Multipath Correction

Multipath essentially increases the dimension of the signal. If the objective is to eliminate multipath effects, then the problem is to simply pull the signal back to a lower dimension by eliminating the highest dimension. The operation is conceptually simple and is done by resetting the coordinate for the highest dimension to 0 for all reconstructed vectors.

A variation on this technique can be used to determine the point of origin for a signal -- a multipath direction finding problem. The attractor is formed in real time. The receiver is moved in the direction that lowers the dimension of the attractor. The point where the signal is of the lowest dimension should be the closest point to the transmitting antenna.

Medical

The evidence to date suggests that all signals from biological organisms are chaotic. This makes sense because chaotic systems are, in fact, more stable in a global sense than linear systems. A linear system is easily destroyed by perturbations while a chaotic system simply returns to an attractor. The inherent stability in chaos is probably the mechanism that allows life to survive the near continuous perturbations in nature. Studies (see Glass and Mackey [1988] in the reading list) have shown that normal cardiac signals are chaotic, as are brain activities, micturition, ion levels in the blood, and breathing, to name a few. From a diagnostic point of view, the techniques for the application of chaotic signal processing are the same as for machinery maintenance. A normal attractor is characterized. Deviations from the attractor may be manifestations of problems. For example, heartbeats may become more regular prior to a heart attack. Thus, a lowering dimension of cardiographic traces may indicate preventive treatment.

The pattern of many disease outbreaks is also chaotic. A classical example is influenza in New York City. Records on instances of influenza have been kept since 1900. These data have a well-defined attractor. Thus, chaotic signal processing may be used to predict outbreaks of diseases and their severity.

Economic

Economic systems are good candidates for study by chaotic techniques. For the most part, economic indicators appear to be low dimension noninteger fractal dimension systems. Much work remains to be done, but a primary interest is prediction of market movements and detection of leading indicators. If these systems are chaotic, then the limits on predictability are the major issue. Chaotic techniques (average mutual information, for example) can determine if economic indicators are related and if there is predictive value in a series.

A major issue, however, is that there may not be sufficient data to adequately characterize an attractor. For example, if an indicator is sampled one a day for ten years, there are approximately 2,500 samples. This may be sufficient to characterize a three- or four- dimension system, but is definitely not adequate for six or seven dimensions. Processing these data may result in "saturation" at three to five dimensions that may be mistaken for a low- dimension system.

Music

A pure sine tone is very aggravating. It has no aesthetic appeal what so ever. But, the holding of a note by a trained musician is a melodic accomplishment. Work to date is slim, but multiphonic tones from saxophones and clarinets are chaotic (Keefe and Laden, 1991). Another experiment used a chaotic map to produce a rather pleasing tone. Thus, we have evidence that music has chaotic qualities working from two directions. In one case, music is shown to be chaotic, in the other case one type of chaos makes music. Randle, Inc. is processing larger segments of symphonic music looking for these attributes. Early results are encouraging, but music is a very complex process.

Aside from the scientific interest, the application of chaos theory may lead to better music theory and insights about the perception of sound.

Environmental

The chaotic nature of weather is evident to anyone who lives in the Washington, D.C. area. Local weather forecasts are often good for only a few hours. The primary problem in the application of chaotic signal processing to environmental data is the amount of data that can be collected. One hundred years of daily weather readings is only about 36,500 data. Intuition suggests that weather is a very high degree-of-freedom system. Thus, even several temperature readings a day for a hundred years may not adequately define a long-term behavior. On the other hand, several reading a day for a year may define an annual attractor that is representative of yearly fluctuations.

Nonetheless, chaos theory and chaotic signal processing are an important tool in the analysis of the global environment.

DYNAMIC SYSTEMS

A dynamic system is a system that changes over time. Systems analysis defines attributes of a system that are invariant and encompasses characterization, prediction, and control. Useful tools for the issues discussed in this article include the Fourier spectrum, state portraits (also called flow charts), and difference maps.

One method of characterizing a periodic linear system is to describe it in terms of the Fourier spectrum that is a graph of how much energy is contained in a given frequency band. Because a nonlinear system can display multiperiod or aperiodic behavior, the utility of Fourier analysis diminishes as chaos is approached and other methods of characterizing a system are needed. Fourier analysis is inherently a one-dimensional tool and the dimensionality of many processes of interest is much higher.

The "trajectory" of a system is the change in the state variables over time. Various methods of plotting trajectories are one means of visualizing the characteristics of a dynamic system. A plot of the displacement-versus-velocity of a forced pendulum is one example of a trajectory. A "flow" is a group of trajectories that originate from adjacent initial conditions. Finally, an "action" is a stroboscopic view of a variable at some defined periodic time (such as a Poincaré map) or circumstances in state space (x = 10, y = 15, for example).

An "attractor" is the point or set of points toward which the trajectory of a system tends in the long term (Ruelle and Takens: 1971). An attractor may be a single point, such as a local minimum, or a closed curve for periodic behavior. The attractor may be such a complicated set of points that trajectory appears to wander randomly about the state space. The system dynamics are chaotic and is said to have a "strange attractor."

While this paper only addresses one restricted aspect of signal processing, the terminology and ideas apply to many classes of problems. A panoramic view is that signal processing is just one type of information extraction from data. In this sense, signal processing is no different from characterization of invariant information in a text-type data base.

Duffing's oscillator will be used as a model system in the discussions that follow. It is an example of a deterministic system that can display low-dimension nonlinear and chaotic behavior. This is a simple system that can be implemented electronically (Georg Duffing: 1904) or easily simulated on a computer. It has the form:

$$\ddot{x} + \delta\dot{\xi} + \alpha\xi + \xi^3 = \phi\sigma\iota\nu(\omega\tau)$$

where *f* is the magnitude of the forcing, *d* is the damping, and *a* is a linear stiffness term. Because it is a low dimension system, its behavior can easily be visualized.

Linear Systems

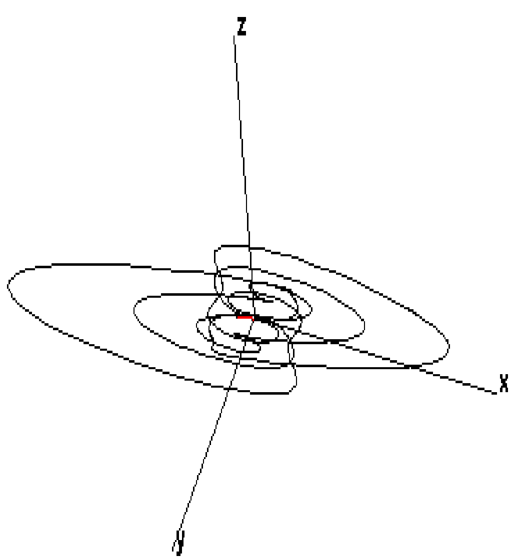
A linear system is one for which superposition and homogeneity are satisfied. That is, additive excitations result in simple additive responses where the relative magnitudes are preserved. The assumption is that there is no interaction between components of the input signal within the system. A small change in one parameter results in a small and proportionate change in the output. In reality, these assumptions are only met for a narrow range of system behaviors. A challenge for engineers was to insure that systems to be modeled remained, or could be assumed to remain, near a state where these assumptions were almost true[3]. For example, a pendulum can be treated as a linear system only if the displacements from the vertical are small.

Many systems that are treated and controlled as linear systems are, in fact, nonlinear systems whose operating regime is restricted to a region where it appears to behave linearly. The caveat that is traditionally applied to these systems is "assuming the system can be kept close to equilibrium . . . .." The implication is that if the system transcends these bounds, the control problem becomes intractable because the linear control model fails.

Nonlinear Systems

A nonlinear system may exhibit more complex behavior than a linear system. A small change in the input may result in disproportionate or counter-intuitive changes in the system response. A pendulum with forcing and damping is nonlinear, especially when far from the vertical. Friction is another example of a nonlinear influence because the amount of friction depends on the initial difference in energy between two surfaces. Its effect varies with the system conditions.

Figure 5 shows the trajectory of Duffing's oscillator in a nonlinear regime. For the parameters shown, the system exhibits a three-period behavior -- the orbit repeats itself every three periods. A Poincaré map of this system is just three points. The Fourier power spectrum for the acceleration term, Figure 7, shows that even a simple nonlinear system may have a broad-band power spectra.



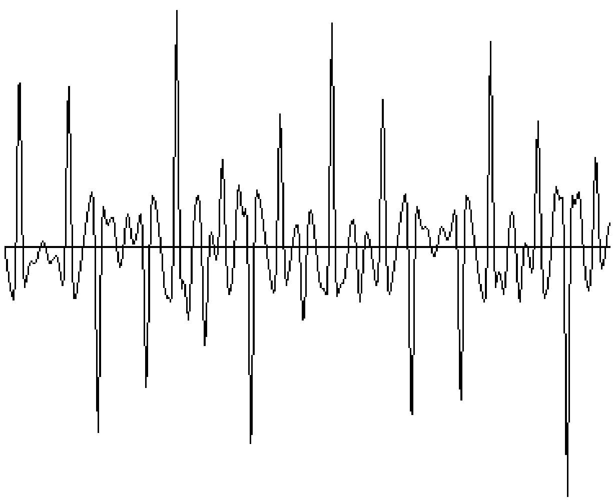
Neither superposition nor homogeneity apply to nonlinear systems. From a modeling standpoint, the nonlinear terms in the equations describe how energy flows between the various components of the system. Because all state variables in a nonlinear system are related, even if weakly, measurement of one variable is sufficient to capture the behavior of all variables.

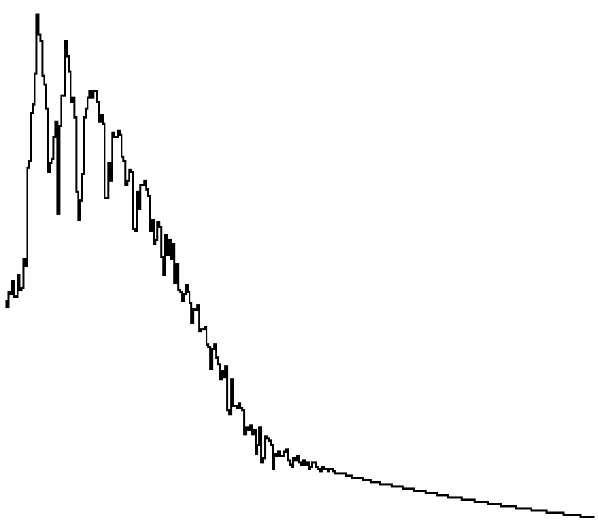
[3]Richard C. Dorf (1967), Modern Control Systems. Addison-Wesley: Reading, MA. pp. 21-22.

Chaotic Systems

Chaotic systems display apparently random behavior but are deterministic. Chaos is not random and is structured when viewed with the proper tools.

The signal trace shown in figure 6 is a signal from Duffing's oscillator when it operates in a chaotic regime. The Fourier analysis of this signal is shown in figure 7. Figure 8 shows the chaotic trajectory. The system is operating in a regime that very close to the nonchaotic case shown in figure 6, so, one would expect the FFTs to be similar. Over a longer period, the FFT becomes more broad band as more of the state space is visited.





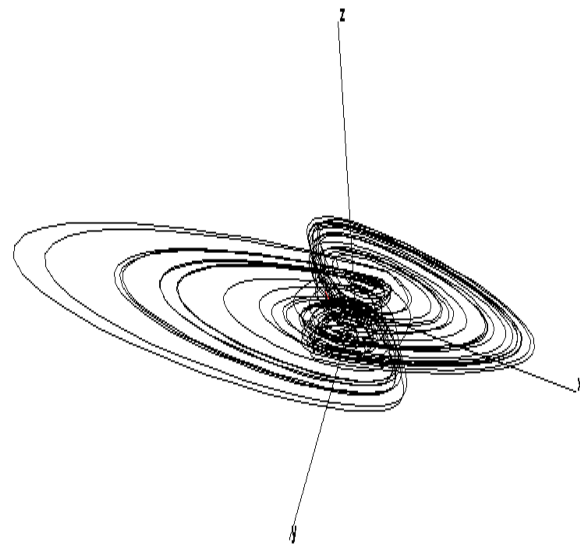


Figure 8 has many places where the trajectories cross. In reality, this cannot be. If the trajectories really crossed, the state vectors would be identical at the intersection. If the state vectors were truly identical, the trajectories would never diverge. But, they do.

There are two explanations. First, we know that this system has three degrees of freedom. Thus, displaying a portrait of the trajectory of any two discards information about the third dimension. The trajectories do not really cross, we have just chosen too low a dimension to display the system adequately.

Second, there are regions where the trajectories appear to be identical, then diverge. Besides the influence of the hidden variable, this illustrates the effects of round off error in displaying chaotic portraits. The trajectories are not really identical, they are just too close to resolve. No matter how much the resolution is improved, seemingly similar trajectories will diverge if the system is chaotic.

This divergence of very close points in state space is a defining attribute of chaos and is measured by the Lyapunov exponents. If at least one Lyapunov exponent is positive, close trajectories will diverge and the system is, by definition, chaotic.

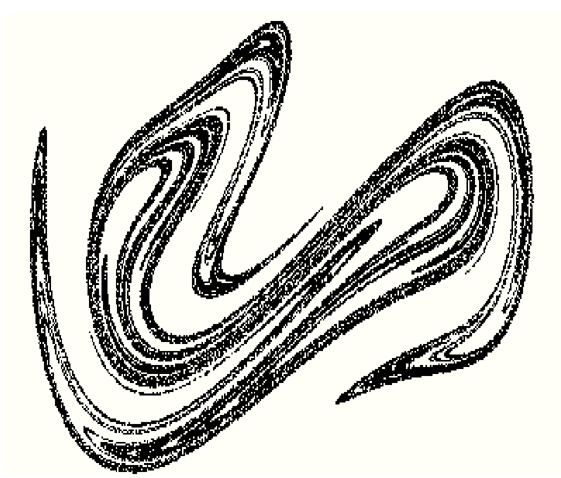


Figure 9 shows a Poincaré section of the portrait of figure 9, taken over 6,000 cycles. The section is a plot of the displacement and velocity terms at the end of each cycle -- one method for forming a Poincaré section. A sense of the fractal nature of the strange attractor for this system is evident.

A defining attribute of a chaotic system is that the future state of the system is critically dependent on its initial conditions. Close but unresolvable initial conditions lead to large divergences in trajectories, so information about prior states is lost as the system evolves. Thus, long term predictions of chaotic systems state are, by definition, doomed to failure. The inability to predict weather for even a few days in advance is a dramatic example of mans' futile grappling with a chaotic system. While long-term predictions of chaotic systems state are not possible, it is sometimes possible to make short-term predictions depending on the accuracy of the model, the precision of the variables, and how fast adjacent trajectories diverge, as measured by the Lyapunov exponents. The Lyapunov exponents are an invariant of a system and are independent of initial conditions. Thus, they also are one means of classifying a system.

A second necessary, but not sufficient condition, for identifying chaos is that a chaotic signal (or more generally, information in data) will masquerade as broad-band noise. Thus, linear methods for analyzing chaotic systems (such as the Fourier power spectrum) do not work. Another problem in dealing with chaotic systems is that the signal can be contaminated by external broad-band noise, which is indistinguishable from the desired signal. A corporate lack of understanding of "noise" prohibits development of simple models for noise rejection.

Finally, the Hausdorff dimension (or more commonly the "fractal dimension") provides a measure of the extent to which the strange attractor (and therefore, the Poincaré mapping of a chaotic system) fills the *m*-dimensional embedding space and is one means of classifying a signal.

**SIGNAL PROCESSING ISSUES
IN LINEAR AND NONLINEAR SYSTEMS**

LINEAR SIGNAL PROCESSING	NONLINEAR SIGNAL PROCESSING
FINDING THE SIGNAL NOISE REDUCTION: DETECTION ------- Separate broad-band noise from narrow-band signal using different spectral characteristics. If system is known, make matched filter in frequency domain.	FINDING THE SIGNAL NOISE REDUCTION: DETECTION ------- Separate broad-band noise from narrow-band signal using deterministic nature of the signal. If system is known, make matched filter in time domain.
FINDING THE SPACE FOURIER TRANSFORMS ------- Use Fourier space methods to turn differential equations or recurrsion relations into algebraic forms: *x(n)* is observed; $x(f) = 3\ x(n)e^{(i2\pi nf)}$ is used	FINDING THE SPACE STATE SPACE RECONSTRUCTION ------- Using time lagged variables, form coordinates for the state space in d dimensions: *y(n) = [x(n), x(n+T),...,x(n+(d-1)T)]* How to best determine d and T? Use false neighbor embedding and average mutual information.
CLASSIFY THE SIGNAL Sharp spectral peaks Resonant frequencies of the system Quantities independent of the initial conditions	CLASSIFY THE SIGNAL Invariants of orbits Lyapunov exponents; Moments of invariant distribution in state space Quantities independent of initial conditions
MAKE MODELS, PREDICT, and CONTROL $x(n+1) = 3\ c_f(n-j)$ Find parameters c_j consistent with invariant classifiers (spectral peaks).	MAKE MODELS, PREDICT and CONTROL $y(n)\ 6\ y(n+1)$ ασ τιμε εϖοl $\psi(\nu+1) = \Phi[\psi(\nu), \alpha_1, \alpha_2, \ldots \alpha_\pi]$ Find parameters a_j consistent with invariant classifiers (Lyapunov exponents, dimensions)

Table by Henry D. I. Abarbanel

PROCESSING CHAOTIC SIGNALS

The methods we use, shown in the table, were developed by Henry D.I. Abarbanel, Reggie Brown, Matt Kennel and others at the Institute for Nonlinear Science, University of California, San Diego. These procedures are based on information theoretic analyses to ensure that the best parameters are derived. Other less rigorous methods yield results that lead to greater errors in subsequent procedures.

The sensor output or signal detector provides a sequence of scalar numbers, *x(1), x(2), x(3) . . . x(N)* at times $t_0 + n\Delta t$. After the signal is detected, time lagged vectors are used to construct a d dimension phase space that serves as the coordinate system for capturing the attractor for the system[4]. The time lagged vectors have the form:

$$y(n) = [x(n), x(n + T), x(n + 2T), \ . \ . \ . \ x(n + (d - 1)T)].$$

Embedding the vectors *y(n)* in the *d* dimension phase space to form a *phase portrait* (figure 1) is a major objective. The use of the word "phase" may lead to some confusion. Our usage bears no relationship to signal phase. For the sake of clarity, we use the term "state and a suitable number of dimensions must be identified. After the state space portrait is formed, the system can be analyzed.

The organization of our processor is shown in figure 12. The procedures, in detail, are:

I. **Acquire a String of Scalar Numbers.**

This may be a nontrivial exercise if the signal-to-noise ratio is low, the bandwidth of the signal is unknown, or the location of the signal in the spectrum is not known. Because a chaotic signal is, by definition, broad band and may masquerade as noise, collection may be difficult. Other factors that influence detection of chaotic signals are identical with those experienced with linear signals. The bandwidth of measuring equipment and operating the equipment at the correct time and place are important considerations.

Despite the degrees-of-freedom of the underlying dynamic system, one can usually collect data from only a single sensor. That is, information from one dimension is usually all that is available. The novel aspect of chaotic signal processing is to use information embedded in this one dimension time series to capture multidimension chaotic behavior.

The dynamical dimension is the number of degrees-of-freedom for the system and is the number of Lyapunov exponents for the system. The time series of a single variable (acquired in one dimension) can be used to construct coordinates for the multivariate state space because in a nonlinear system all variables are coupled. The behavior of any one variable has embedded in it full knowledge of the behavior of all other variables. The bottom signal trace in figure 8 is chaotic and will be used as an example in this discussion.

II. **Find a Suitable Time Lag, *T*.**

---

[4]Eckmann, J.-P. and D. Ruelle, "Ergodic theory of chaotic and strange attractors", *Rev. Mod. Phys.* **57** 3, pp. 617-656 (1985)

Recent approaches are to define *T* based on the idea of average mutual information[5]. The Average Mutual Information, *I(T),* is a prescription for selecting an appropriate time delay interval (T) for construction of the time lagged vectors that will be used to build the attractor.

Within a given dimension, it also is a signal detection tool for unknown signals. *I(T)* defines how much one learns about a datum by having knowledge of another datum. The mutual information of a system is the amount of knowledge (expressed as bits) that one can derive about two datums separated by the time lag, *T*.

Mutual information was first used by Shannon[6] to measure corruption in communications channels. He postulated a transmitter, a channel, and a receiver. Mutual information quantitatively describes the amount of information that is conveyed through the channel. Shannon also defined "information entropy."

This idea was extended to dynamical systems in 1960 by Y. Sinai and to chaotic systems in 1981 by Robert Shaw. The physical communications channel of Shannon can be replaced by a dynamical system. That is, given a knowledge of the system state at some time (or location), the state of the system at other times or locations can be derived. Mutual information tells us how much knowledge is conveyed. For a true linear system with no noise, such as a pure sine wave, knowledge of the system state (the amplitude) at any time is sufficient to define completely the system at any other time. If the signal is corrupted by noise, or if the system is chaotic, then information is lost over time.

In 1986, Andrew Fraser and Harry Swinney published "Using Mutual Information to Find Independent Coordinates for Strange Attractors."[7] This landmark work showed that mutual information does provide significant information about chaotic systems. A subsequent paper and dissertation extended the two-dimension ideas to multiple dimensions and called the technique "minimum redundancy analysis." We retain the term Mutual Information for no particular reason even though we are technically performing minimum redundancy computations.

The mutual information in two dimensions is defined in terms of the joint probabilities of the two datums:

$$I_{AB}(T) = \sum \Pi_{A,B}(\alpha,\beta)\log_2\left[\frac{\Pi_{A,B}(\alpha,\beta)}{\Pi_A(\alpha)\Pi_B(\beta)}\right]$$

Mutual information is a measure of general dependence and is loosely related to the idea of correlation functions and autocorrelation. But, correlation functions only measure linear dependence. The first zero crossing of the autocorrelation is often used to determine the time delay. But, average mutual information provides better estimates of the time lag to use for the state space reconstruction of the attractor.

The determination of the time lag, *T*, is important because an optimum selection of *T* gives best separation of neighboring trajectories within the minimum embedding space. This is important because calculation of the Lyapunov exponents relies on solving a matrix that is comprised of descriptions of how close trajectories diverge. If the trajectories are not separated, then the matrix will be ill conditioned and may not be solvable.

---

[5]A. M. Fraser and H. L. Swinney. "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, 33:1134-1140, Feb. 1986.

[6]C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[7]A. M. Fraser and H. L. Swinney, "Using Mutual Information to Find Independent Coordinates for Strange Attractors", *Phys. Rev. A*. **33**: 1134-1140, Feb 1986.

If *T* is too small, there is little new information contained in each subsequent datum. If *T* is too large, *x(n)* and *x(n+T)* will appear to be random with respect to each other for a chaotic system. In fact, *T* can be somewhat arbitrary for an infinite amount of noise free data. The quality of the state space portrait depends on *T*, so it is desirable to select some value that is reasonable.

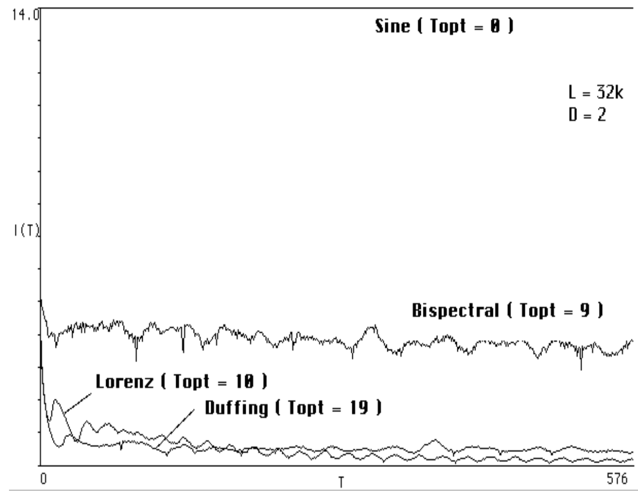
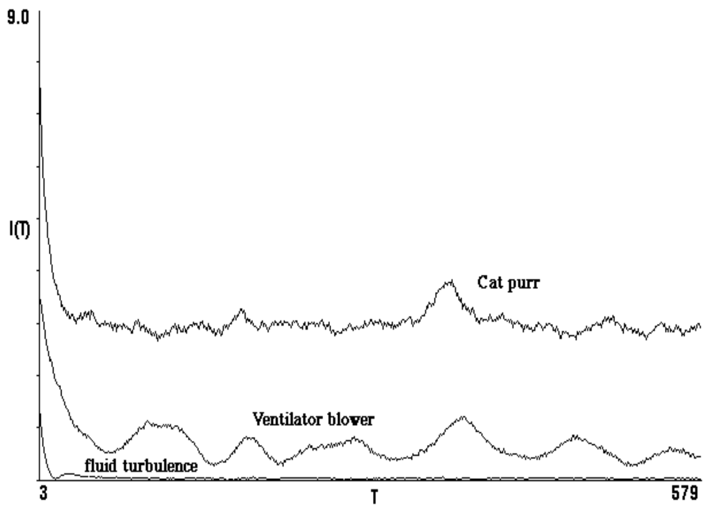


Figure 11 shows a plot of *I(T)* for four simulated systems. The sine wave is the prototypical linear signal. The plot shows that if a state vector (the amplitude) is known, the behavior of a sine wave is perfectly known for all time.

I(T) for a polyspectral (a nonlinear but not chaotic) signal is also shown. While it has some randomlike characteristics, the average mutual information of this signal is high and does not decrease with time.

Finally, *I(T)* for two chaotic systems are also shown in figure 11 -- the Lorenz system and Duffing's oscillator. Average mutual information drops off rapidly as the time lag increases (which, by definition, shows that these systems have positive entropy). Essentially, the signal rapidly becomes random with respect to itself. A plot of a truly random signal (or a signal that is random in a two-dimension state space) drops to near zero even at $T = 1$.

The first local minimum of *I(T)* determines an optimum value of *T*. For Duffing's oscillator, $T_{opt} = 9$. For a chaotic system, as *T* increases past this point, ambiguities in the correlation between *x(n)* and *x(n + T)* arise -- they start to appear to be random with respect to each other. The state space portrait begins to loose resolution and the fractal nature of the attractor starts to become blurred. Information is being lost. *I(T)* for a truly random signal is zero for all time lags, except 0. Thus, mutual information provides one rationale why chaotic signals appear as broad-band noise to a linear processor.



These calculations are sensitive to the number of values used in the time lag vectors and to the number of dimensions. For a clean signal of dimension three, a sample size of 32,768 is the minimum number of samples that will yield good results. For dimensions greater than three or four, many more samples may be required to get an accurate estimate of *I(T)*.

Figure 15 shows the average mutual information for several real signals. The tape recorder hiss has a higher average mutual information than the other signals because it is somewhat more deterministic -- the range of possible values is very small, so knowledge of any point provides significant knowledge about where any other value lies. The other two systems, the automobile engine and the cat purr, rapidly decorrelate (in a multidimension nonlinear sense) from themselves and I(T) is much lower.

In a manner analogous to using the Fourier power spectrum as a signal detection tool, average mutual information can be used to detect signals. With the FFT, one asks “is there a Fourier coefficient that has a significantly higher amount of energy than other coefficients.

We have conducted detection tests using average mutual information as the detection metric on simulated spread spectrum signals down to -40 db SNR and have worked with a real signal of this type at an estimated SNR of 5. Spread spectrum signals are easily detected with chaotic techniques.

III. **Find the Minimum Embedding Dimension, d_E.**

If a dynamic system can be described by *n* independent variables, then the full behavior of the system can be observed in a *n* dimensional "state" space. But, the attractor of the system may be contained in a subset of the state space with dimension d_A, and may be described in a state space, *d*, that is much smaller than *n*. This minimum embedding dimension d_E is, at most, the first integer greater than $2d_A$; it may be less.

The dimension of the underlying dynamics, *d*, determines how many Lyapunov exponents are useful. Determination of the minimum embedding dimension, d_E, is of practical interest because the computation burden rises dramatically as dimension increases. Further, noise fills all dimensions, so computations carried out in a higher than necessary dimension will be corrupted by noise. If d_E is too small, the trajectory may cross itself and neighbors at a point in this area may be indistinguishable in the lower dimension. Generally, by making $d_E > 2d_A$ self intersections can be avoided (Mañé and Takens)[8,9].

A new method of determining minimum embedding dimension is used in our processor[10]. As dimension is increased, attractors "unfold." Points on trajectories that appear close in dimension *d* may move to a distant region of the attractor in dimension *d+1*. These are "false" neighbors in *d* and the method measures the percentage of false neighbors as *d* increases. Trajectories that are close in *d* are tallied, and the number of these trajectories that become widely separated in *d+1* are calculated. Over the data one tallies

$$\left[\frac{R_{d+1}^2(\nu,\rho) - P_\delta^2(\nu,\rho)}{P_\delta^2(\nu,\rho)}\right] = \frac{\left|\xi(\nu + T\delta) - \xi^{(\rho)}(\nu + T\delta)\right|}{P_\delta(\nu,\rho)} > P_{\tau o\lambda}$$

where R_d is the Euclidian distance between a point and its nearest neighbor and R_{tol} is the criteria for declaring whether the neighbors that are close in d are distant in d+1.

A second criteria is necessary because the nearest neighbor is not necessarily "close." If the nearest neighbor to a point is false but not close, then the Euclidian distance in going to d+1 will be μ $2R_a$. So, the second criteria is

$$\frac{R_{d+1}(\nu)}{P_A} > P_{\tau o\lambda}$$

where

$$R_A^2 = \frac{1}{N}\sum_{\nu-1}^{N}[\xi(\nu) - \bar{\xi}]^2$$

A nearest neighbor is false if either test fails.

For a noiseless signal, the number of false neighbors becomes zero when the minimum embedding dimension d_E is reached. A noisy signal drops off dramatically, but does not become zero. For SNRs as low as 6 to 10 dB, the percentage of false neighbors drops below 1% for $d \geq d_E$.

[8] Takens, F. in Dynamical Systems and Turbulence, Warwick 1980, eds. D. Rand and L.S. Young, Lecture Notes in Mathematics 898, (Springer, Berlin), 366 (1981).

[9] Mañè R., in Dynamical Systems and Turbulence, Warwick 1980, eds. D. Rand and L.S. Young, Lecture Notes in Mathematics 898, (Springer, Berlin), 230 (1981).

[10] Kennel, Matthew B., R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phy. Rev. A* **45** pp. 3403-3411, 15 March 1992.

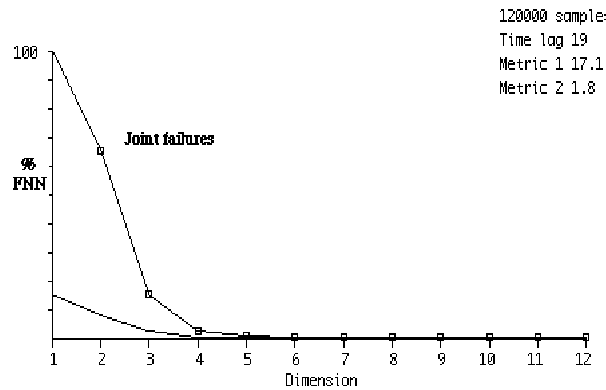
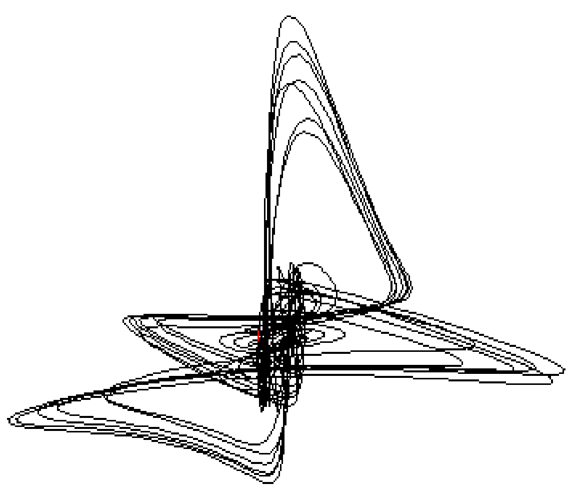


Figure 13 shows these calculations for Duffing's oscillator. The minimum embedding dimension appears to be 4, which is consistent with our knowledge of the system.

IV. **Embed the Time Lagged Variables in the State Space.**



Once T and d_E are determined, the time lagged vectors are plotted in the state space to form the portrait. If the dimension is low enough, computer generated graphics may be used to display the portrait. Dimensions up to four have been plotted, where the fourth dimension is represented by colors. If the dimension is greater than four, the portraits cannot be effectively visualized.

Figure 16 shows the reconstruction of a chaotic trajectory for Duffing's oscillator. It is different from figure 9 partly because this portrait is shown in a three dimension space with a rotated coordinate system. Portraits reconstructed from time lagged scalars do not completely resemble portraits developed from the original vectors. But, the invariant properties are unique and can be used to identify and exploit the signal. Figures 16 through 21 show reconstructed state space attractors for a variety of systems.

Appendix B shows how the selection of T affects the state

space portrait of Duffing's chaotic oscillator. At $T = 0$, there is no time lag and every datum is plotted against itself. As T increases, the attractor begins to unfold. For $T = 6$ and $T = 7$ the characteristics of the attractor are well defined. After this point, the portrait becomes progressively indistinct. In effect, the reconstructed signal is losing information about itself as T increases.

Figure 15 has many places where the trajectories appear to cross. In reality, this cannot be. If the trajectories really crossed, the state vectors would be identical at the intersection. If the state vectors were truly identical, the trajectories would never diverge. But, they do.

There are two explanations. First, we know that this system has three degrees of freedom. Thus, displaying a portrait of the trajectory in two dimensions discards information about the third dimension. The trajectories do not really cross, we have just chosen too low a dimension to display the system adequately. Real signals (figure 20, for example) have attractors that appear to be much more ill defined than the simulated signals for several reasons. First, we usually display the portraits as three-dimension systems projected on the two-dimension page. But, these systems are probably more than three dimensions. The portraits are reconstructed from real signals and reflect the messy real world and less than optimum collection circumstances. So, we personally like real attractors better than the simulated ones.

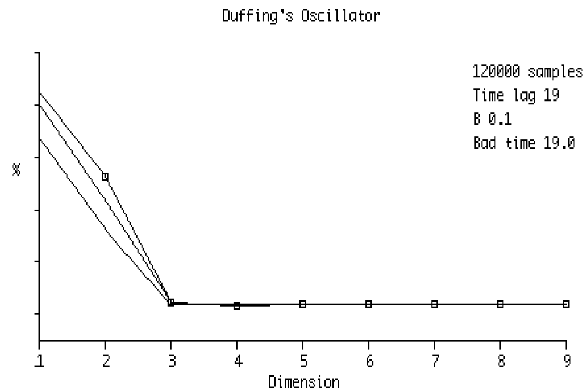
V. **Determine the local embedding dimension, d_L.**

The global embedding dimension d_E is necessary to completely unfold the attractor. However, local evolutions may be adequately described in fewer dimensions[11]. If such a local dimension d_L exists and if $d_L < d_E$, then all the important dynamical behaviors can be captured in d_L. The local dimension does not vary for different regions on the attractor. It is is the invariant degrees of freedom that describes the deterministic behavior for all finite regions on the attractor.

Predictive models are usually made for finite horizons. So, a model in d_E would contain

[11] Abarbanel, H.D.I. and M.B. Kennel, "Local False Nearest Neighbors and Dynamical Dimensions from Observed Chaotic Data", Phys. Rev. E., **47**, 3057-3068 (1993).

more degrees of freedom than are necessary for finite time predictions. If $d_L < d_E$, then the model can be simplified.



The method for calculating the local dimension is straight forward. Nearest neighbors for many neighborhoods are found. If the trajectories for these neighbors separate faster than is expected for close neighbors, it can be presumed that the separation is caused by the projection into too few dimensions. Referring back to figure xx, the trajectories associated with close neighbors remain close for some reasonable period. Those that apparently cross because of the projection onto the page separate rapidly. The task is simply to define a criteria for how fast sufficiently embedded trajectories should separate and then test a large number of neighborhoods against this criteria.

A neighborhood is selected and a local coordinate system is defined using the principle components of the neighborhood. This results in coordinates lying along the eigendirections for the neighborhood (the direction where most of the trajectories are "headed"). This procedure is is relatively simple and easy to compute. The nearest neighbors in the local coordinate system are then selected.

For all pairs of neighbors, the percentage in each neighborhood for which the trajectories remain "close" is computed. "Close" is arbitrarily defined as some size relative to the overall size of the attractor. The embedding dimension is then increased and the computations are repeated. The dimension where the percentage of "bad" predictions becomes independent of the embedding dimension is d_L. The beauty of this scheme is that one does not care if the predictions are good, bad, or indifferent. The only issue is when do the predictions become independent of the embedding dimension. Hence, an easy to calculate prediction estimator can be used.

Duffing's oscillator is described by three first-order ordinary differential equations. The global embedding dimension for this system is 6, based on the earlier global false nearest neighbor calculations. But, figure xx shows that $d_L = 3$. The metric is 10% of the mean size of the attractor and 19 samples. That is, if the trajectories separate by more that 10% of the attractor size in less than the nonlinear decorrelation time (T_{opt} from average mutual information), the prediction is bad.

This confirms that a prediction model derived solely from observations of the data would need only three degrees of freedom. In this case, we know that a priori, because we generated the signal from a simulation we wrote. But, in most interesting real world data analysis, one does not have that advantage.

VI. **Compute the Fractal Dimension of the Attractor, d_a.**

The fractal dimension of the attractor[12], d_a, provides information on how much of the state space is filled by the system. One interpretation of d_a is that it measures how many degrees of freedom are significant. Another interpretation of the fractal dimension is that it provides a measure of how an object's bulk scales with its size: bulk = $size^{da}$. Bulk that can be associated with volume and size is then interpreted as Euclidean distance. A plane, for example, has dimension two because the area = d^2.

The fractal dimension of the attractor, d_a, may be estimated using Ruelle's approach by calculating the number of spheres or boxes, $N(r)$, of size r that capture all points as r approaches zero:

$$d_a \approx \frac{\log(N(r))}{\log(1/r)} \quad \text{as } r \to 0$$

Grassberger and Procaccia[13] defined a relatively easy approximation (the correlation dimension) that may be done on a PC for high SNR signals of low dimension. One major issue is the sensitivity of these calculations to signal SNR. The amount of data required to do the calculations may dramatically increase as SNR decreases. In fact, if the diameter of the attractor is an order of magnitude greater than the smallest r, then $N = 10^{d/2}$[14]. The computation burden in high dimensions, even with this algorithm, is significant. Also, real signals may not have even distributions across small r and it is necessary to use local smoothing to reduce wide fluctuations in the region of interest.

---

[12]F. Hausdorff, "Dimension and ausseres Mass," *Math. Annalen*, vol 79, pp. 157-179, 1918.

[13]P. Grassberger and I. Procaccia, "Measuring the Strangeness of Strange Attractors", *Physica* **9D**, 189 (1983).

[14]D. Ruelle, Proc R. Soc. A 427 (1990) 241.

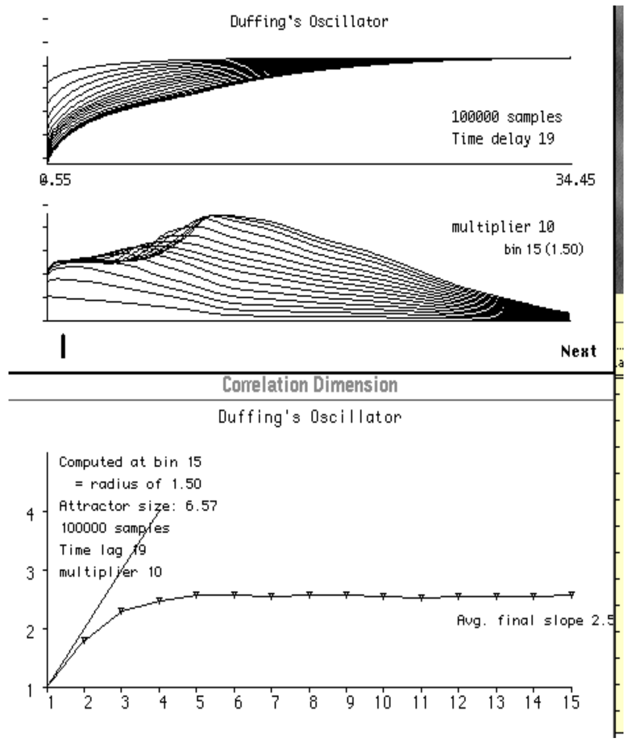


Figure 21 shows these calculations for Duffing's oscillator. The raw counts of *N(r)* are on top and the local slope of *log(N(r))/log(1/r)* for all *r* is in the middle. The final estimate of d_A for Duffing's chaotic oscillator, are the slopes at the cursor. The point chosen is the smallest r for which the counts are reliable.

The final estimate of d_A is 2.5. This is consistent with our other knowledge of the system. That is, the fractal dimension meets Ruelle's criteria, $2d_A + 1 < d_E$. A correlation dimension of 2.5 is also consistent with the local dimension of 3.

As a practical matter, however, these computations are easily corrupted by noise. Even a small amount of contamination is sufficient to blur the attractor and make the results meaningless.

VII. **Develop the Map.**

The map is a function that moves a point in state space to the next point as a function of time by using a local polynomial. The Taylor series expansion of the polynomial is the Jacobian of the underlying dynamics. Recent approaches retain the higher order terms to better fit local curvatures in the attractor.

This step develops a function that maps small displacements in the orbit into small displacements at the next time step by defining a local polynomial that maps $y(n)$ to $y(n+T_2)$. The map describes the distance between nearby orbits and how the distance between these neighbors change over time. The distance between neighbors at T_0 is $z^r(n;0)$. The corresponding distance at time T_2 is:

A Taylor series expansion of *F* contains the Jacobian of the underlying dynamics. Until recently, only the first term in the Taylor series was retained. Our approach is to retain terms that are up to fifth order in the Taylor series. The terms in the ill-conditioned Jacobian matrix are solved by a least squares fit. The problem is computation intensive because a N_b x d_E matrix must be inverted d_E times. It is desirable to include the higher order terms in the neighborhood mapping because the corresponding Jacobian provides a better fit to local curvatures of the data set[15] and all the Lyapunov exponents can be derived. Small errors caused by excluding the higher order terms can lead to large changes in the computed Lyapunov exponents.

VIII. **Compute the Global Lyapunov Exponents,λ.**

The Lyapunov exponents describe the rate at which close points diverge. If one or more Lyapunov exponents is positive, the system is chaotic[16]. The Lyapunov exponents are invariant with respect to initial conditions and smooth changes of coordinates. Therefore, they are another way of classifying a chaotic system. The practical interest in Lyapunov exponents is that they are a measure of predictability and the limits on predictions.The algorithmic requirements are that all exponents must be found, the results must be accurate for small sample sets and the algorithm and results must be robust to noise.

The Lyapunov exponents may be calculated from the Jacobian of the map by the QR decomposition technique discussed by "EKRC."[17] The Lyapunov exponents are a measure of how quickly the trajectories of very close points in state space diverge. If the Lyapunov exponents are all zero or negative, the trajectories do not diverge and the system is stable. If one or more Lyapunov exponents is positive, the trajectories diverge and the system is unstable. A requirement for chaos is that at least one Lyapunov exponent be positive[18]. For the chaotic regime of Duffing's oscillator used as an example in this paper, $\lambda \approx .01$ (Moon, 1987). Equally important is the relationship of the Lyapunov exponents and an ability to predict the behavior of a system. The more exponents one can correctly find, the better predictions will be[19].

Noise, however, corrupts the local Jacobian and can affect the accuracy of the calculation of the Lyapunov exponents, as shown in figure 25 (Abarbanel: 1990). For $d_E = 3$ the exponents are robust for very small levels of noise but rapidly loose credibility as the noise level rises. Thus, one must work with the best data possible, many samples, or use a noise-reduction algorithm, if the results are to be believable.

IX. **Compute the Local Lyapunov Exponents,λ.**

As defined above, Lyapunov exponents are a global invariant because they describe the effect of infinitesimal perturbations over infinite time. Recent approaches examine how perturbations grow in finite time[20,21]. The "local" Lyapunov exponents measure the average divergence of trajectories in different regions of state space for finite lengths. For example, in figure 17, there are regions where the trajectories do not diverge. The distance between them

---

[15]Reggie Brown, Paul Bryant and Henry D.I. Abarbanel, "Computing the Lyapunov Spectrum of a Dynamical System from Observed Time Series", Physical Review Journal, March 15, 1991.

[16]J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57** (1985) 617.

[17]Eckmann, J.P., S.O. Kamphorst, D. Ruelle, and S. Ciliberto, *Phys. Rev.* **A34**, 4971 (1986).

[18]Grebogi, C.E. Ott, S. Pelikan, and J. A. Yorke, Plasma Preprint UMLPF 84-014 October 1983.

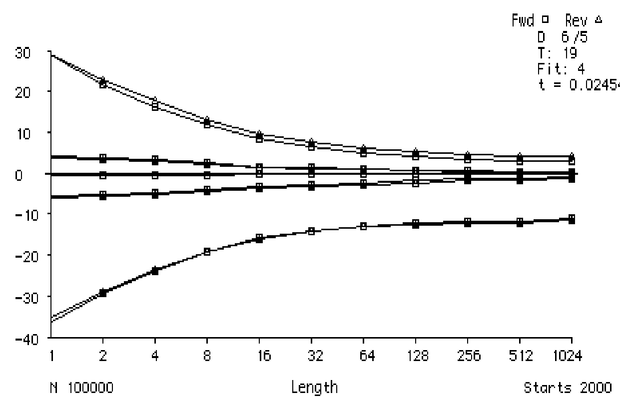
[19]Abarbanel, Henry D. I., "Determining the Lyapunov Spectrum of a Dynamical System from Observed Data", Presented at the SIAM Conference on Dynamical Systems, Orlando, Florida, 8 May 1990.

[20]Abarbanel, H. D. I., R. Brown, and Matthew Kennel, "Variation of Lyapunov Exponents on a Strange Attractor", *Journal of Nonlinear Science*, **1**, 175-199 (1991).

[21]Abarbanel, H.D.I and M.M. Sushchik, “True Lyapunov Exponents and Models of Chaotic Data”, Int. J. of Bifurcation and Chaos, **3**, 543-550 (1993)

remains fairly constant for long periods. It is only in the center "core" where the fractal nature of the orbits are apparent and the Lyapunov exponents are positive and large. The issue addressed by local Lyapunov exponents is predictability[22]. Again, referring to figure 17, there are large regions of the state space where knowledge of a state vector permits prediction over a reasonable long time. There are other regions (of very large positive local Lyapunov exponents) where there is almost no predictability. These considerations are important for the study of many systems including economics.

Spurious exponents and their elimination



One problem caused by selecting too high a global dimension is that Lyapunov exponents will be generated for all dimensions. The exponents for dimensions above d_E are spurious. In fact, exponents for dimensions greater than the degrees-of-freedom that are active locally are also spurious. Identification of these unwanted, misleading, inaccurate, and meaningless parameters is accomplished by reversing the time series and computing the Lyapunov exponents for both the forward and backward time evolutions[23].

The signs of the exponents for the reverse time series are changed and the values are compared to the forward pass. The true exponents for each dimension will now be identical (or very close) for both the forward and reverse pass. The spurious exponents will have different values.

This is especially evident in noisy data. Because the noise occupies space in the dimensions above the active dynamics (the local dimension), the spurious exponents are affected in unpredictable ways.

Fortuitously, we can address all three issues with a single set of calculations. The local Lyapunov exponents are calculated for a variety of lengths (giving the average local exponents), including a length that is sufficient to cover the attractor (generally giving a good estimate of the global exponents). The time reversed exponents are computed for the same data, identifying the spurious exponents and providing an indication of the local dimension (d_L). The one trap here is that an adequate d_E must be determined prior to performing the calculations. Fortunately, the False Nearest Neighbor test has already provided this.

---

[22]Abarbanel, H. D. I., R. Brown, and Matthew Kennel, "Local Lyapunov Exponents Computed from Observed Data", accepted for the *Journal of Nonlinear Science* (Sept. 1992).

[23]Ulrich Parlitz, "Identification of true and spurious Lyapunov exponents from time series," Int. J. of Bifurcation and Chaos, Vol 2, No. 1 (1992) 155-165.

X. **Compute the Moments of Invariant Distribution.**

The moments of invariant distribution measure the number of times an orbit visits a region of the embedding space. The invariant density of orbits is computed by dividing the total data set into two parts. The second part is further divided into G groups of length L. A G x G matrix is generated and the eigenvalues and eigenvectors are computed. After normalization, a set of G orthonormal functions that are invariants of the mapping can be found[24]. Although not used for classification, invariant moments are used for noise reduction.

NOISE REDUCTION BY PROBABILISTIC SCALED CLEANING

Signals may be generated by dynamical systems that are of higher dimension than signals of interest. These high dimension signals completely fill the lower dimension state space and are called "noise." ***In our paradigm there is no such thing as noise in the traditional sense.*** *In any d dimensional state space there are signals of dimension d or lower that can be characterized and exploited. There are signals of dimension greater than d that fill the state space and corrupt measurements of the lower dimension signals.*

Noise is just a signal from a higher dimension dynamical system. Our approach is to treat both the signal and the noise as deterministic processes that can be separated in an appropriate state space.

Noise probably has components from many dimensions. Elimination of noise from a source of some dimension still leaves components from even higher dimensions. But, from a signal detection perspective, reduction of some portion of the noise has the same effect as boosting detector SNR by improving the signal. The fundamental issue is that methods of improving signal strength may be expensive or intractable. Reducing noise by applying chaos theory has the same effect and may be relatively inexpensive.

Noise reduction by Probabilistic Scaled Cleaning (PSC) determines the invariant properties of a signal's attractor, and then uses this knowledge to separate the signal from other data. PSC can be used to separate a deterministic signal from higher dimension noise, or can be used to separate a complex signal (such as speech) from low-dimension chaotic noise. The only distinction is in which data are called the signal and which are called noise. As long as one of them has an invariant attractor, PSC will work. A videotape made by Professor Abarbanel demonstrating PSC is available either from Randle, Inc. or The Institute for Nonlinear Science.

Procedurally, PSC works as follows[25]. The data to be cleaned are selected and, if available, a reference orbit is selected. A state space reconstruction of the signal is performed using an appropriate time lag. If the embedding dimension is too large, the only penalty is that the computations are more expensive. If the dimension is too low, then the distances used to estimate the distributions will be incorrectly computed. But, as the embedding dimension approaches the correct minimum dimension, the error may become small and the errors that

---

[24] Abarbanel, Henry D. I., "Prediction in Chaotic Nonlinear Systems: Time Series Analysis for Nonperiodic Evolution," Lectures at the NATO Advanced Research Workshop on MODEL ECOSYSTEMS AND THEIR CHANGES, Matrea, Italy September 4 - 8, 1989. INLS preprint 1020.

[25] Marteau, P.-F. and H. D. I. Abarbanel, "Noise Reduction in Chaotic Time Series Using Scaled Probabilistic Methods", *Journal of Nonlinear Science*, **1**, 313-343 (1991).

accrue in the distributions may become less significant. We have not explored this issue in detail. Early studies INLS/UCSD show that even if the estimate of the embedding dimension is low, significant noise reduction is still possible.

PSC analyzes the conditional probability that the state of the system is *S* when the scalar number measured by the sensor is *O*. This conditional probability is $P(S|O)$ and the task is to maximize this conditional probability over the possible values of the true state *S*.

PSC uses a reference orbit from the system to find the state probability density and the translation probability from some state to another at a later time. Essentially, the Markov property of the dynamics is translated into probability densities. When presented with an observed chaos plus "another signal," PSC shifts the observations around in state space until the densities are accurately matched. An estimate of the correct location of chaotic signal in state space results from this. The "other signal" is estimated by subtraction from the observations.

Ideally, the reference orbit is obtained from clean observations of a signal. If clean data are not available, a noisy orbit can be used for "self cleaning." The limits on self cleaning are an issue that remains to be studied.

PSC is, in essence, a maximum posteriori (or MAP) estimator method. It adds to the usual MAP procedures: (1) analysis in d-dimension space appropriate to the problem; (2) an evaluation of probability densities from the reference signal rather than from *a priori* assumptions; and (3) a recurrsive procedure that explores finer and finer state space scales at each pass.

Two other approaches to noise reduction have been developed by Stephen Hammel and Eric Kostelich. Hammel attacks noise reduction by "shadowing" where a clean (true) orbit of a known system (i.e., the map is known) is constructed from a noisy orbit by using a recurrsive process that operates on an entire noisy orbit to yield a new, less noisy, orbit. "The refinement process is then applied to this new, cleaner orbit to produce a third orbit, expected to be less noisy than its predecessor." The process may be repeated up to 30 times, depending on the orbit length and the noise level. The Hammel noise reduction algorithm[26] and its generalizations[27] presumes that the mapping function is known and seeks solutions that map new data in a numerically stable manner that exploits stable and unstable manifolds in the state space. Recent research suggests that the Hammel noise reduction technique may be useful in Kalman filters, which can be sensitive to noise.

Kostelich's approach is to minimize the effect of noise by assuming that the observed trajectory is a noisy representation of the true trajectory and that the trajectory dynamics are near linear over small regions. An ensemble of points is used to compute a linear approximation of the dynamics at each point on the attractor. The observed trajectories are replaced with nearby ones that best satisfy the linear approximations[28].

---

[26]Stephen M Hammel, "A noise reduction method for chaotic systems", **Physics Letters A**, **148**, 8,9 pp 421-428, 3 Sept. 1990.

[27]Abarbanel, H. D. I., S. M. Hammel, P.-F. Marteau, and J. J. ("Sid") Sidorowich, "Scaled Linear Cleaning of Contaminated Chaotic Orbits", UCSD/INLS Preprint, Spring, 1992.

[28]Kostelich, Eric J. and Yorke, James A., "Noise Reduction: Finding the Simplest Dynamical System Consistent with the Data," **Physica D** 41 (1990) 183-196.

INDEX

APPENDIX A - CHAOS READING LIST

Gleick, James (1987) *Chaos: Making a New Science*. Viking: New York, NY.

Ruelle, David (1991) *Chance and Chaos*. Princeton University Press, Princeton, NJ.

Shaw, R., (1984) *The Dripping Faucet as a Model Chaotic System*. Aerial Press: Santa Cruz, CA.

Pool, Robert, "Is It Chaos, or Is It Just Noise?," *Science*: Research News, Vol 243, pp. 25-28, 6 January 1989.

Pool, Robert, "Is Something Strange About the Weather?," *Science*: Research News, Vol. 243, pp. 1290-1293, 10 March 1989.

Stewart, Ian, (1989) *Does God Play Dice? The Mathematics of Chaos*. Basil Blackwell: Cambridge. MA.

Moon, Francis C. (1987) *Chaotic Vibrations: An Introduction for Applied Scientists and Engineers.* John Wiley & Sons: New York, NY.

Thompson, J. M. T., and Stewart, H. B., (1986) *Nonlinear Dynamics and Chaos*. John Wiley & Sons: New York, NY.

Abarbanel, Henry D. I. "Signal Processing on Strange Attractors," presented at the Randle/SAIC seminar on nonlinear signal processing, March 1990.

Casdagli, Martin and Stephen Eubank, eds., *Nonlinear Modeling and Forecasting*, Papers from a workshop sponsored by the Santa Fe Institute and NATO, Addison-Wesley Publishing (1992).

Keefe, Douglas H., Laden, Bernice, "Correlation dimension of woodwind multiphonic tones," *J. Acoust. Soc. Am.* **90 (4)**, pt 1. October 1991, pp 1754 - 1765.

Abarbanel, Henry D. I., Brown, Reggie and Kadtke, James B., "Prediction in Chaotic Nonlinear Systems: Methods for Time Series with Broadband Fourier Spectra," *Phys. Rev. A* **41** 4:1782-1807 (1990).

Abraham, R. H., and Shaw, C. D. (1982) *Dynamics: The Geometry of Behavior*, Part One, *Periodic Behavior*. Aerial Press: Santa Cruz, CA.

Abraham, R. H., and Shaw, C. D. (1982) *Dynamics: The Geometry of Behavior*, Part One, *Chaotic Behavior*. Aerial Press: Santa Cruz, CA.

Leon Glass and Michael C. Mackey (1988) *From Clocks to Chaos*, Princeton University Press, Princeton, New Jersey.

Parker, Thomas S., and Chua, Leon O., "Chaos: A Tutorial for Engineers," *Proceedings of the IEEE*, Vol. 75, No. 8 pp. 982 - 1008, August 1987

Abraham, R. H., and Shaw, C. D. (1982) *Dynamics: The Geometry of Behavior*, Part One, *Global Behavior*. Aerial Press: Santa Cruz, CA.

Berge, P., Pomeau, Y., Vidal, C. (1984) *Order Within Chaos: Towards a Deterministic Approach to Turbulence*. John Wiley & Sons: New York, NY.

David Ruelle (1989) *Chaotic Evolution and Strange Attractors*, Cambridge University Press, Cambridge, UK.

Theiler, James, "Estimating the Fractal Dimension of Chaotic Time Series," The Lincoln Laboratory Journal, Vol. 3, No. 1, pp. 63-86, (1990).

Brown, R., Bryant, P. and Abarbanel, H. D. I., "Lyapunov Exponents from Observed Time Series," *Phys. Rev. Lett.* **65**, 1523 (1990).

Brown, R., Bryant, P. and Abarbanel, H. D. I., "Computing the Lyapunov Exponents of a Dynamical System from Observed Time Series," *Physical Review A* **43**, 6 March 15, 1991, pp. 2787-2806

Abarbanel, Henry D. I., "Determining the Lyapunov Spectrum of a Dynamical System from Observed Data," Presented at the SIAM Conference on Dynamical Systems, Orlando, Florida, 8 May 1990.

Abarbanel, Henry D. I. and Kadtke, James B., "Information Theoretic Methods for Determining Minimum Embedding Dimensions for Strange Attractors," INLS preprint.

Fraser, Andrew M., "Information and Entropy in Strange Attractors," *IEEE Trans. on Information Theory*, Vol. 35, No. 2, pp. 245-262, March 1989.

Fraser, A. M., and H. L. Swinney. "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, 33:1134-1140, Feb. 1986.

S. Wiggins (1990) *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, New York.

Kostelich, Eric J. and Yorke, James A., "Noise Reduction: Finding the Simplest Dynamical System Consistent with the Data," *Physica D* **41** (1990) 183-196.

Hammel, Stephen M. Hammel, "A Noise Reduction Method for Chaotic Systems," *Physics Letters A* **148** (1990) No. 8,9 pp. 421-428.

Ditto, W. L., Rauseo, S. N. and Spano, M. L., "Experimental Control of Chaos"

Feigenbaum, Mitchell J., "Universal Behavior in Nonlinear Systems," *Los Alamos Science*, pp. 4-27, Summer 1980.

Yu. A. Kravtsov, "Randomness, determinateness, and predictability," Usp. Fiz., Nauk **158**, 93-122, May 1989.

A. V. Gaponov-Grekhov and M. I. Rabinovich, "Some Problems in Non-Linear Dynamics: Dynamic Chaos and Structures," Vestnik Akademii Nauk SSSR, 1988 pp. 3-5.

Frederic D. Tappert, Michael G. Brown, and Gustavo Goni, "Weak chaos in an area-preserving mapping for sound ray propagation", *Physics Letters A* **153** 4,5 (1991), pp 181-185.