# CHAOS ANALYSIS FOR EKG TIME SERIES DATA

CHETAN MEHTA, MATTHEW MILLER

Abstract. Our project aims at testing various measures of chaos in time series data from an electrocardiogram (EKG). We examine data from three types of patients who we assume to produce varying EKG dynamics. We use time-delay embedding and calculate correlation dimensions to determine if the data is chaotic or random. Our results indicate that there is some evidence of deterministic chaotic behavior in the RR-Interval time series for an athletic patient and an atrial fibrillation patient. The correlation dimension for both types indicated chaotic behavior in five and six dimensions over a specific range of $r$.

## 1. Introduction

Measuring the electrical impulses of the heart through an electrocardiogram (EKG) has become the primary method for detecting problematic heart conditions such as heart attacks or cardiac arrhythmia. Gaining more insight into the dynamical behavior of heartbeat irregularities would have meaningful applications in cardiology–especially if an "irregular" heartbeat could be characterized as being chaotic or deterministic, as opposed to being merely random. Using times series from electrocardiogram (EKG) data, one is able to apply methods of chaos theory to investigate possible determinism in EKG data. A study by Zhang *et al* examined what they believed to be deterministic behavior in the EKG data of patients with "abnormal heart conditions, such as ventricular tachycardia [and] branch block."[1] Using eight different subjects from the MIT/BIH arrhythmia database, Zhang used a wavelet transformation to extract the QRS complexes and generate a "QRS complex irregularity time series." Their results indicate that the EKG data possessed some chaotic and deterministic behavior, indicated by positive maximum Lyapunov exponents and correlation dimension between five and six, both of which hint at the presence of a chaotic attractor.

Our project will compare filtered EKG data for three different types of patients: a normal patient, an athletic patient, and an atrial fibrillation patient. We will use two methods of chaos theory to investigate the deterministic behavior of the heartbeats of these three subjects. First, we will use time-delay embedding to construct phase portraits of the EKG data for each patient in order to examine the presence of a potential attractor. Secondly,

---

*Date*: December 6, 2007
Math 53: Chaos!
Professor Alexander Barnett
Dartmouth College, Department of Mathematics.

[1]Zhang *et al*, p. 467

we will use time-delay embedding for $m$ dimensions to calculate the correlation dimension for each patient type, which we will then compare with the correlation dimension of a random time series and that of the Henon map, a well-known chaotic attractor.

## 2. Electrocardiogram Data

An EKG measures the electrical impulses of heart activity, and is composed of four components: the p-wave, the QRS complex, and the t-wave. We will focus primarily on the QRS complex, a measure of the depolarization of the ventricles[2]. The maximum of the QRS complex is the R-peak, the "spike" in a typical EKG. Often conditions of the heart, in particular cardiac arrhythmia, are detected by irregular behavior in either the amplitude, duration, or frequency of the QRS complex.

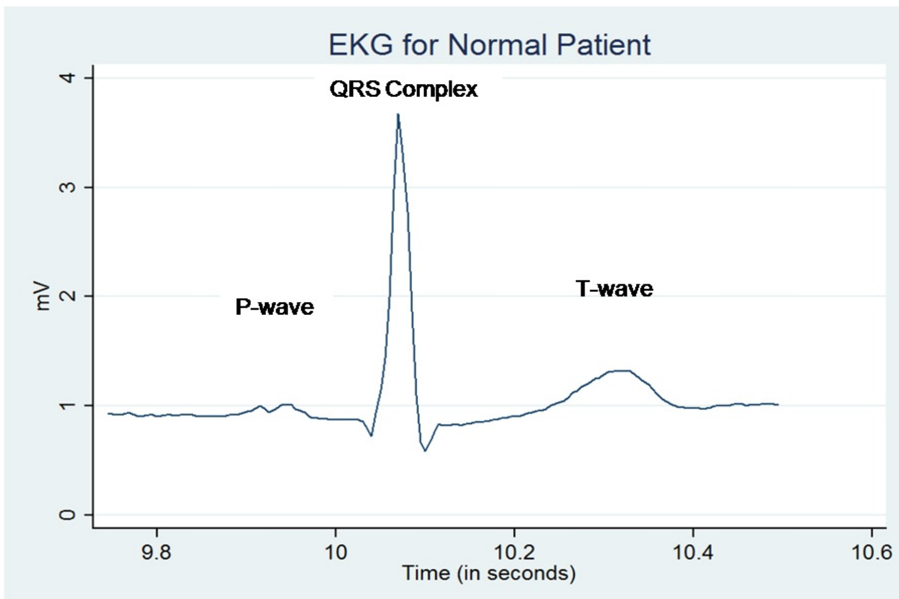


Figure 1. An EKG example

### 2.1. RR-Interval Extractions.

In order to examine possible chaotic behavior in EKG data, we needed to construct a relevant time series from the EKG, allowing us to apply chaos methods for time series data. First, we defined our EKG time series as follows:

$$\vec{x_t} = \{x_0, x_1, \ldots, x_T\}, t = 0, 1, \ldots, T \quad (1)$$

Where $T$ is the maximum time in the series. One useful measure in detecting heartbeat irregularities is the RR-Interval, which we define as the duration between the R-peaks.

[2] http://en.wikipedia.org/wiki/Electrocardiogram

Given the variety in some of our EKG data, we first need to filter the data so we are able to extract the R-peaks. To smooth out any fluctuations within the data, we use the following filter. Let $x_t*$ be the filtered values of $x_t$.

$$x_t* = x_t - \frac{1}{2J+1}\sum_{p-J}^{p+J} x_p \tag{2}$$

Where $J$ ranges from 40 to 100. This filter smooths the EKG so we are more accurately able to extract the R-peaks, and thus the RR-intervals. Using a simple algorithm on the filtered data, we located the times of the R-peaks by isolating the times where $x_{t-1}* < x_t* > x_{t+1}*$ & $x_t* > F$, where $F$ is a threshold above which the only EKG maximums are the R-peaks. Let $\{t_i\}$ denote the set of all $t$ where $x_t$ satisfy the above criterion. The durations of the RR-Intervals, $z_i$, were calculated by $z_i = t_i - t_{i-1}$ for $i = 0, 1, \ldots, N$. The set $\{z_i\}$ is our set of times for the RR-Intervals. We will use this time series for our analysis.

2.2. **Patients and Conditions.** It is important to note the types of patients we tested and the conditions under which each was tested in order to make clear the circumstances under which the EKGs were recorded. All patients were tested with the same electrode placement on their bodies.[3]

2.2.1. *Normal.* Our first patient type was a normal patient. By this specification we mean they were sitting down, not recently engaged in any athletic activity, and do not suffer from any known heart condition.

2.2.2. *Athlete.* We also recorded EKG data for athletes. Athletic subjects were in-season athletes, whose EKG was taken directly after finishing a running test of approximately one mile. Their heartrate data was measured while they were lying down. Again, none were suspected of suffering from any type of heart condition.

2.2.3. *Atrial Fibrillation.* Our third patient type were patients who suffered from atrial fibrillation (AF), a cardiac arrhythmia in which an EKG demonstrates missing p-waves or "irregularity" in the RR-Intervals, as demonstrated in Figure 2.

It is suspected that such cardiac arrhythmia may lead to chaotic or deterministic RR-Interval durations. The EKGs for these patients were not recorded by us directly.[4]

---

[3]This placement was as follows: the green electrode was placed on the upper right area of the chest (below the shoulder), red electrode on the inside of the left thigh, black electrode on the inside of the right thigh.

[4]The EKG data for AF patients was provided by Eric Posmentier, Department of Earth Sciences, Dartmouth College. The data was in fact recorded with the same arrangement of the electrodes.

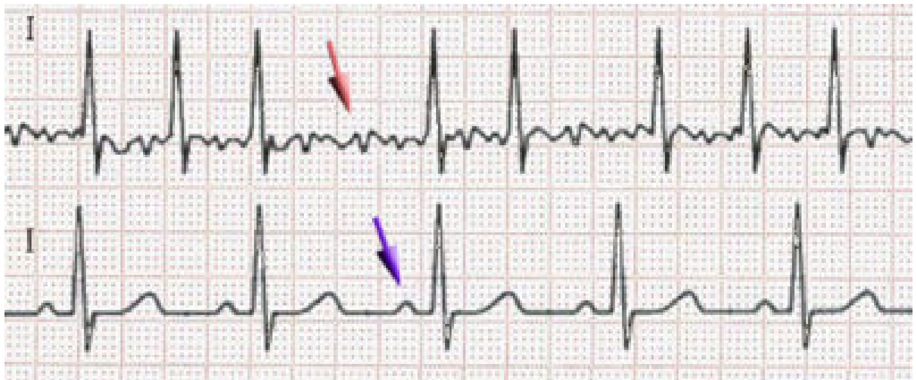

Figure 2. EKG for an AF patient

## 3. Chaos Analysis

Our chaos analysis will consist of time-delay embedding and correlation dimension comparison for a number of dimensions. Before we analyze phase portraits for EKG data we will consider the correlation dimension plots, generated from time-delay embedding in eight dimensions, from the two extremes: random data and the Henon map.

3.1. **Time-Delay Embedding.** A practical method of examining possible deterministic behavior in time series data is a time-delay plot, a technique that graphs the value at a given time versus the value at a time period $(m-1)\tau$ after, where $m$ is the embedding dimension. It is possible that chaotic behavior is present in a time-delay plot of higher embedding dimension. The highest dimension we can plot is $m = 3$. We use $\tau = 8\bar{z}$ where $\bar{z}$ is the mean of the set $\{z_i\}$.[5] The three dimensional time-delay plots for our three patient types are below.

3.1.1. *Normal.* At first glance the time-delay plot for the normal patient does not exhibit any particular chaotic behavior. This is somewhat expected, since we do not anticipate the normal patient to exhibit any symptoms of heartbeat irregularity. If a patient were to demonstrate perfectly normal EKG dynamics, their time delay plot would be a collection of points at one time $t_{RR} = t_{RR} + \tau$, corresponding to the duration of every RR-Interval. It is unsurprising that this patient exhibits some variation in RR-Interval duration, since it would be very unlikely for even the most healthy patient to generate a perfectly normal EKG. The delay plot in Figure 3 uses a $\tau$ as suggested by Zhang *et al.* An alternate time-delay plot, with a simple delay of $\tau = 1$, was "cigar"-shaped for $m = 2$. This structure would indicate significant dependence on the preceding RR-Interval. This "linear" dependence deteriorates in higher embedding dimensions.[6]

[5] Zhang *et al*, p. 470

[6] Only for the normal patient were there significant changes in time-delay embedding given the two different $\tau$.

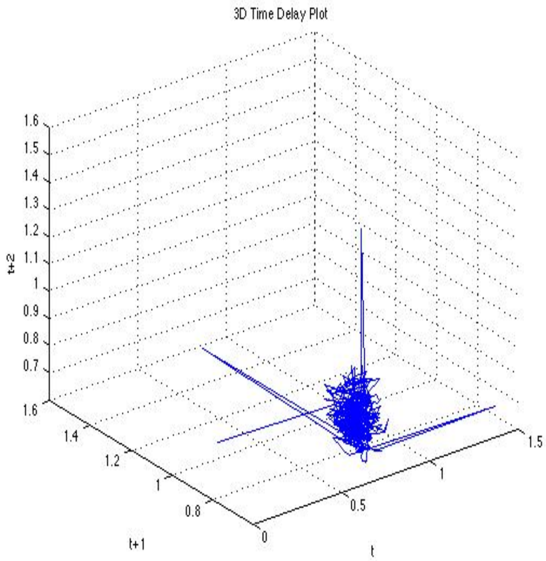


FIGURE 3. Time-Delay Plot, Normal Patient, $m = 3$

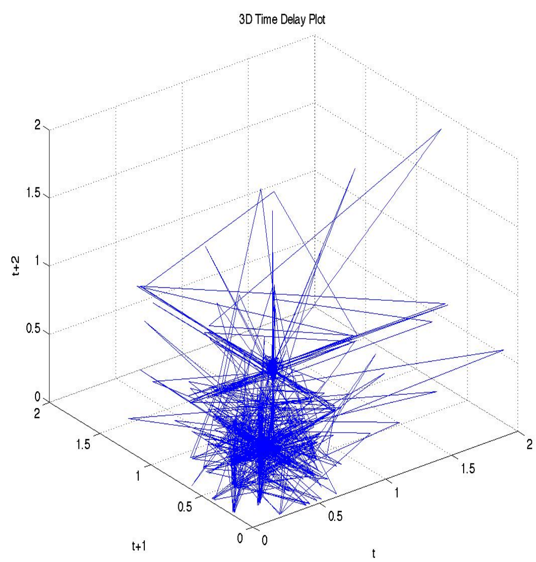


FIGURE 4. Time-Delay Plot, Athletic patient, $m = 3$

3.1.2. *Athlete.* The time-delay plot for the athlete is pictured in Figure 4. Our goal with measuring EKGs for athletes was to demonstrate what we anticipated to be a case in between the normal patient and the AF patient. There does appear to be some attractive behavior within the time-delay structure, however there is significant noise present that disrupts our ability to further determine whether or not an attractor exists. This may be due to our filter, which was not entirely robust to drastic fluctuations in the EKG. For example, heavy breathing by the athletes post-exercise resulted in an undulating EKG, making our method of RR-Interval extraction more difficult. Further analysis may require subdividing the time series by intervals of high and low variance in voltage, and applying custom filters to more precisely identify RR-Intervals.

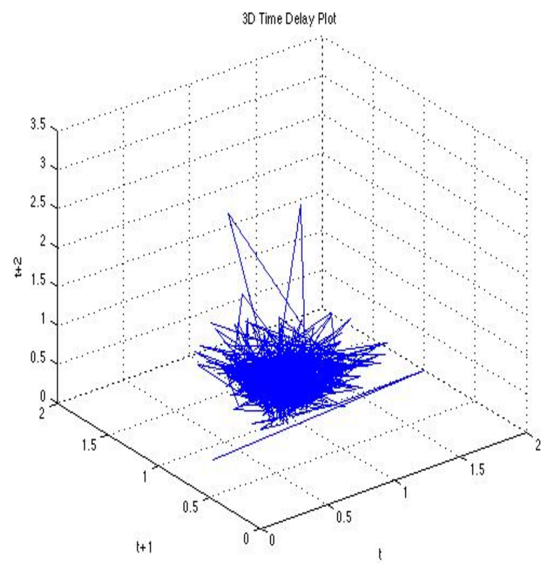


Figure 5. Time-Delay Plot, AF patient, $m = 3$

3.1.3. *Atrial Fibrillation.* The time-delay plot for the AF patient is modeled in Figure 5. This plot is very similar to the time-delay plot for $m = 3$ in Zhang *et al*, demonstrating a relatively dense collection of points in three dimensions. A plot for a randomly generated time series possesses a very similar structure to both our plot and the plot presented by Zhang *et al.* This indicates that, for one and two time-delays, the duration of the RR-Intervals is random and evidently not deterministic.

3.2. **Correlation Dimension.** Time-delay plots give insight into the possible chaotic behavior of time series in $R^2$ and $R^3$ , but it becomes difficult to visualize time-delay structure for higher embedding dimension. A way of further investigating chaos in time series is to examine the correlation dimension of a time series for a number of different embedding dimensions. The figures below plot the $logC(r)$ versus $log(r)$, and the correlation dimension, $d(m)$ versus embedding dimension, $m$. For each dimension $m$ we calculate the average slope using the slopes between consecutive $r$-values. We then took the average of the slopes for a given dimension as a measure of $d(m)$.

3.2.1. *Extremes: Random Data and the Henon Map.* Before we analyze the correlation dimension plots for our EKG time series, it will be useful to look at the plots for the two extreme cases. The image on the left of Figure 6 is the plot of $logC(r)$ versus $log(r)$, where $r$ is a set distance and $C(r)$, the correlation function, is the ratio of the number of $z_k, z_j$ within set distance $r$ from one another and the total number of $z_k, z_j$ in $\{z_i\}$. The slope, therefore, is the correlation dimension.[7]

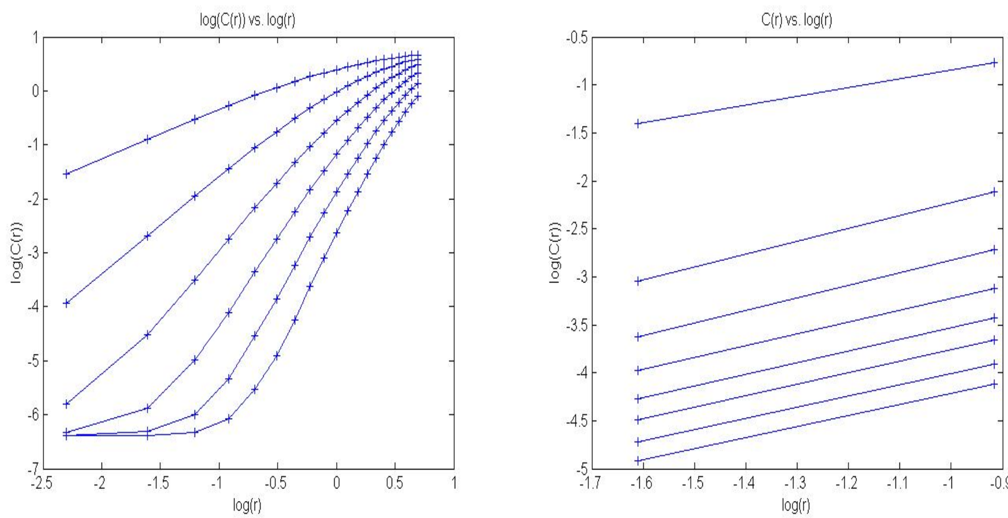


Figure 6. Correlation Dimension for a Random Time Series and the Henon Map

It is evident from these plots that for the random time series, the slopes of the trajectories at a given $r$ are all different. This implies that the correlation dimension does not converge or for any dimension, and we do not have a chaotic attractor. This is certainly logical for the random time series, which by definition exhibits no chaotic or deterministic behavior. For the Henon Map, we see that for all dimensions and all $r$ the slope, and thus the correlation dimension, is the same. We calculated a correlation dimension for the Henon attractor of $\approx 1.3$. This is the expected structure of the correlation dimension for a chaotic attractor. The correlation dimension should converge to some specific value for all embedding dimensions in the presence of an attractor.

[7]Alligood, p. 180

3.2.2. *Normal.* For the normal patient (Figure 7) it is again clear that no chaos is present. The slopes across different dimensions are consistently different, indicating, as expected, the absence of determinism.

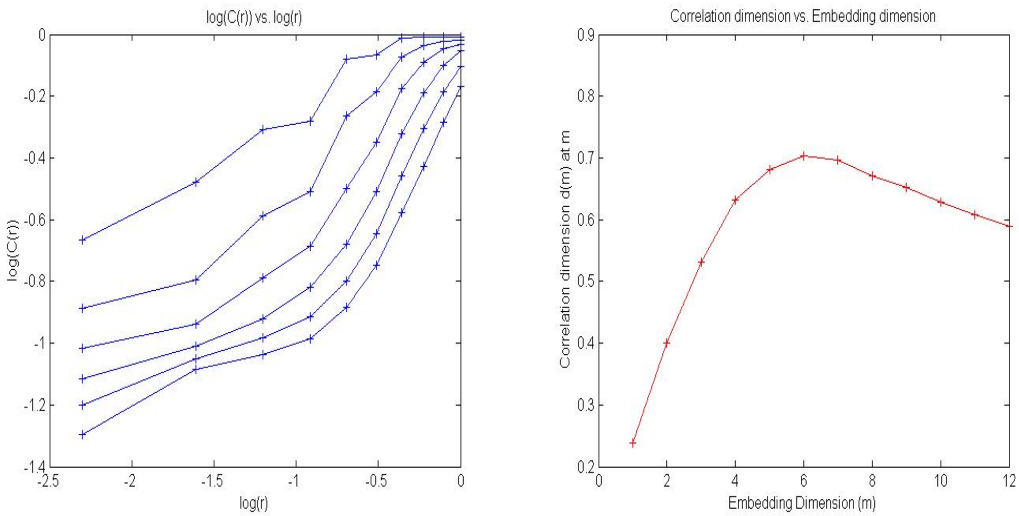


Figure 7. Correlation Dimension, Athlete patient

3.2.3. *Athlete.* The correlation dimension plot for the athletic patient in Figure 8 indicates some evidence of a chaotic attractor. The slopes for some lower dimensions are relatively similar, but not close enough to forcefully imply chaos. However, for later dimensions ($m = 4, 5, 6$) and $-.5 < log(r) < -0.1$, the slopes appear to be very similar, indicating what may be the presence of an attractor in the relevant embedding dimensions. This is consistent with the time-delay plot for the athlete which indicated possible attractive behavior, although for lower dimensions.

3.2.4. *Atrial Fibrillation.* For the correlation dimension of the AF patient (See Figure 8), the structure appears, for the most part, very similar to the correlation dimension plot for the random time series. There is some similarity between slopes in higher dimensions, possibly hinting at the presence of an attractor in dimensions five or six (at a range of $-0.75 < log(r) < -0.5$). This is consistent with the findings of Zhang *et al* that there is some evidence of a chaotic attractor for these dimensions.

## 4. Conclusion and Further Considerations

Our results indicate that there is little evidence in favor of determinism or chaos for an RR-Interval time series extracted from EKG time series data from a normal patient. Although there were some slight indications of determinism for the athletic patient, the results from time series for the AF patient were very similar to equivalent measurements

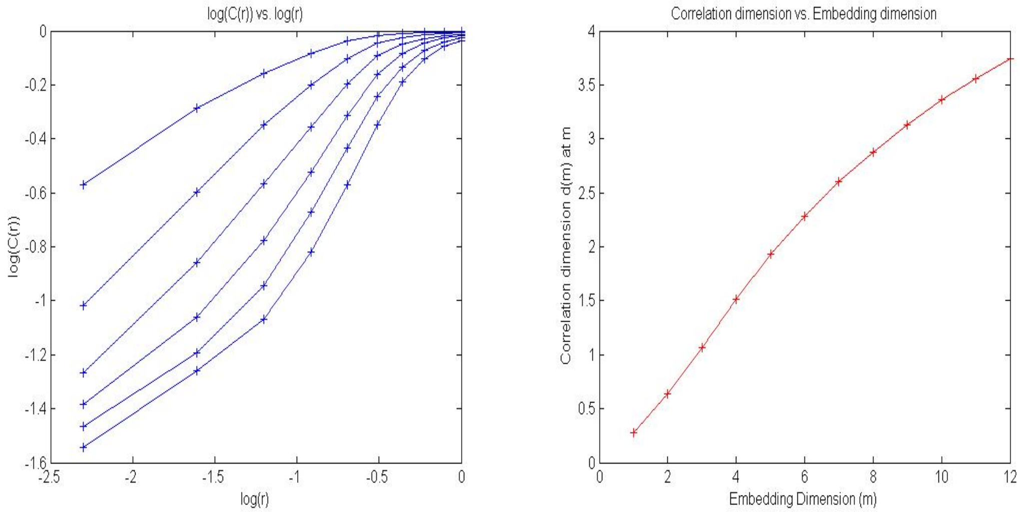


Figure 8. Correlation Dimension, AF patient

conducted on a random time series for most dimensions, with slight evidence from correlation dimension calculations indicating attractive behavior in five and six dimensions. This leads us to conclude that there is greater evidence that the durations of RR-Intervals for an EKG are random rather than chaotic.

Further considerations would be alternate measures of chaos, namely the measurement of the maximum Lyapunov exponents for time series, using the Wolf *et al* algorithm. It would also be useful to extract a different time series from the EKG data, perhaps the RT-interval or the QRS complex amplitude and duration. These time series, once extracted, could be analyzed using our methods of determining chaos.

## References


[1] Alligood, Kathleen T., Sauer, Tim D., & Yorke, James A. (1996). Chaos: An Introduction to Dynamical Systems. New York: Springer.

[2] Atrial fibrillation. Retrieved December 7, 2007, from Wikipedia Web site: http://en.wikipedia.org/wiki/Atrial_fibrillation

[3] Electrocardiogram. Retrieved December 7, 2007, from Wikipedia Web site: http://en.wikipedia.org/wiki/Ecg

[4] Peters, Edgar. Chaos and Order in The Capital Markets: A New View of Cycles, Prices, and Market Volatility. 2nd Edition. New York: John Wiley & Sons, Inc., 1996.

[5] Wolf, Swift, Swinney, Vastano. "Determining Lyapunov Exponents from a Time Series." Physica 16D(1985): 285-317.

[6] Zhang, X. S., Zhu, Y. S., & Zhang, X. J. (1997). New approach to studies on ECG dynamics: extraction and analysis of QRS complex irregularity time series. Medical & Biological Engineering & Computing. 35, 467-474.