2

Methodological Aspects of Heart Rate Variability Analysis

Tom Kuusela

CONTENTS

2.1 Introduction

There are numerous methods and approaches available for time-series analysis. This chapter presents selected methods, which are or at least could be useful when analyzing biosignals, especially heart rate variability (HRV). The methods described here differ greatly, and some of them are commonly used characterization methods. One must always select the method to use, case by case.

The analysis methods are often divided into *linear* and *non-linear* methods. The term "non-linear methods" is in fact not quite appropriate and is actually even misleading, because those methods themselves are no more non-linear than common spectrum analysis using fast Fourier transform (FFT) algorithm. The idea behind the naming convention is merely related to the fact that the system producing the time series is non-linear, which in the case of biological systems might well be true. Sometimes these methods are described using the term *chaos-theoretical approach*. This is true only of some of the methods in which the assumption that the system's trajectory in a phase space is actually a chaotic small-dimensional attractor. There is very little proof for the existence of this in biological systems. In particular, if we are studying the cardiovascular blood pressure regulatory system and typical time series such as the sequence of RR intervals, we can find in the literature a growing number of results that indicate this system to be stochastic rather than chaotic.

Biosignal time series, in general, are very difficult to analyze, in many ways. Many methods demand that the time series has thousands of data points, which may be impossible or difficult to achieve, depending on the measurement method used. When the time series is long, notable changes in the biological system will occur almost inevitably and the time series will start to drift in the parameter space of the system due to either internal or external influences. Such instability has a negative effect on all analysis methods in which a specific characteristic of the whole time series is to be compressed into a single statistic. Of course, calculations will produce some result, but what is the meaning of the statistic? Other disturbing factors in these signals are notable noise levels, strong discreteness of the signals due to a limited amplitude resolution of the digitization and regular strong periodic signals, for example, breathing modulation on top of the otherwise possibly chaotic or stochastic signal. It should be remembered that most of the methods described in this chapter were originally developed in physics for analyzing very different, and often practically ideal systems.

Despite warnings stated earlier, an analysis of biosignals using such methods may still yield useful information. Even though the calculated statistic might not exactly

describe the characteristic it was originally designed to represent, it could still play an important role when searching for a correlation with other clinically interesting parameters. However, basic assumptions and limitations accompanying each method should be clearly understood when interpreting results and searching different explanatory models.

2.2 ECG Recording and Generation of RR Interval

2.2.1 RR Interval Time Series

HRV analysis is based on the RR interval time series, the sequence of intervals between successive fiducial points of R peaks of QRS complexes in the electrocardiogram (ECG). It should be noted that the RR interval time series are actually an event series and not an equally sampled continuous signal. This fact is important, especially when performing frequency domain analysis. It should also be noted that HRV analysis does not measure the rhythms of the sinoatrial node, which is the pacemaker of the heart, since it is not based on P-P intervals. Therefore, HRV reflects fluctuations in the atrioventricular conduction superimposed on the P-P interval. Nevertheless, it has been shown that beat-to-beat changes in RR intervals reflect the variability of the sinoatrial node quite accurately. In theory, it would be better to use P-P intervals for HRV analysis, but in practice, the amplitude of the P wave is small, which makes the accurate determination of P wave peak difficult, especially in the presence of noise. In turn, this would affect the accuracy of P-P interval measurement.

2.2.2 Time and Amplitude Resolution of ECG Recording

The sampling frequency and amplitude resolution of ECG recording are important parameters in HRV analysis. The sampling frequency determines the temporal resolution of R peak recognition of QRS complexes, and, therefore, it determines directly the accuracy of the measurement of RR intervals. Too low a sampling rate produces digitization noise and introduces errors in all HRV measures, whether in time and frequency domains or in non-linear analyses. In most cases, a sampling rate of 200 Hz corresponding to a temporal resolution of 5 ms is high enough, and higher sampling rates do not give any better results. If the overall variability of RR interval is very low, as it can be, for example, with some heart failure patients, HRV analysis requires a higher sampling rate, since changes in the length of RR intervals cannot be captured below that resolution.

The amplitude resolution of the analog–digital conversion of ECG signal is rarely a limiting factor. The commonly used resolution of 12 bits is normally sufficient. However, the effective amplitude resolution is always less than the resolution of the conversion because the full dynamic range of input signal in the converter is rarely utilized in order to avoid signal clipping or saturation in the case of a large ECG signal or a wandering baseline.

2.2.3 Duration of the Recording

The duration of the recording is determined by the method used for HRV analysis, the aim of the study, and stationarity issues. Typically, frequency domain methods are preferred

for short-term measurements and time domain methods for long-term measurements of HRV. Some non-linear methods are likewise suitable for short-term analyses and others for long-term analyses. It is important to note that although the same mathematical method is used for the analysis of both short-term and long-term ECG recordings, the physiological interpretation can be different, that is, they cannot necessarily be considered as the surrogates for each other just with a different duration. There are also non-linear methods where results depend directly on the number of data points, and thus, it is inappropriate to compare the HRV measures obtained from recordings of different durations, with each other. This is also true of most linear methods.

2.2.4 Stationarity of the Recording

The stationarity of the signal is of great importance. There are many methods available for evaluating the stationarity of a signal, depending on how we define stationarity. Stationarity can mean, for instance, that there is no shifting in the base level of the signal or that the amplitude distribution, spectrum and autocorrelation function of the signal do not change, as a function of time. More generally, we can say that a signal is truly stationary if parameters that define the working point of the system remain constant in time. This evaluation is extremely difficult to undertake in the case of physiological systems, since we have a limited knowledge of the background dynamics and most physiological parameters are unknown.

Stationarity is strongly linked to the duration of the recording. The longer a recording is, the less stationary it is, because inevitably there are changes in the physiological state of the subject. Whatever method is used for the analysis of long recordings, the result is always a complicated and often a non-intuitive average over physiological states. Unless the method is designed for such a situation (e.g., see DFA), this kind of analysis should be avoided. Hence, the best approach is to divide the recording into shorter segments and perform separate analyses on them. This confers several advantages: shorter segments are often more stationary and results more reliable; the temporal change of results is a measure of stationarity (or lack thereof), which can be used to investigate the time evolution of the system; and finally, changes over segments can be utilized when estimating the statistical relevance of the results. The last advantage is most useful because normally we get a single statistical estimate as a result, but have no possibilities to quantify the statistical significance of this number.

2.2.5 Removing Trends

Most non-stationarities of the signal are not evident until they are revealed by a more advanced analysis. The trend in RR interval time series is easily seen, and it is often interpreted as a sign of non-stationarity, which can be removed by subtracting the trend from the data. However, *a priori* there is no way to separate the "real" signal and its trend; the signal is *never* a sum of them. In theory, even large trends can be an essential part of the dynamical behavior of a stationary system. But in practice, many HRV analysis methods are useful only if there is no significant trend in the data irrespective of its origin and in these circumstances, the trend should be removed from the data. Normally, only the linear trend is removed, since the removal of non-linear trends may create a significant bias. From the point of view of spectral analysis, the trend removal decreases the contribution of the

lowest frequencies; thus, further analysis is focused on faster oscillations. And finally, it must be remembered that the removal of a trend does not restore the stationarity of data. The only way to overcome the problem of non-stationarity is to minimize all internal and external disturbances during the study.

2.2.6 Ectopic Beats, Arrhythmias and Noise

The ECG signal can contain technical artifacts or QRS complexes originating outside the sinoatrial node, which introduce errors, since HRV analysis is based on assessment of the variability of sinus rhythm. Artifacts can seriously affect HRV analysis and thus cannot be ignored.

Technical artifacts, such as missed beats (due to problems in the R peak detection) or electrical noise (bad contact in the electrode, movement artifacts), can be easily edited by cleaning the data by a proper interpolation based on the preceding and successive QRS intervals. When interpolating the RR interval time series, it is important that the cumulative time is not altered; RR interval is an event series where the time stamp of each beat is the sum of all preceding beats. For example, the respiratory modulation on RR interval encounters a small phase shift if the time line changes after an interpolated segment. Similarly, time synchronization with other signals should be preserved.

The editing of ectopic beats is more problematic. Ectopic beats are usually premature beats and they produce a very short RR interval followed by a compensatory delay and a prolonged RR interval. Such a short–long pair can be edited without affecting the cumulative time of the beats by lengthening the first interval and shortening the second interval, so that they have equal intervals. However, ectopic beats and most arrhythmias result in reduced stroke volume and cardiac output, leading to transient drops in the blood pressure. This activates autonomic reflexes and induces changes in the efferent autonomic activity. Since these true physiological responses can last 10–30 beats, editing of just the two intervals on either side of ectopic beats does not remove all of the changes in HRV produced by them.

The raw recorded ECG signal should always be inspected. QRS complex identification should be carefully verified. If there are any artifacts or ectopic beats present, the best option is to select a signal segment free of them. If an artifact-free recording is not available, editing can be considered. If an excessive editing of ectopic beats is needed, it should be recognized that HRV analysis can then lead to erroneous results. Very few analysis methods are so insensitive to ectopic beats that the ECG needs no editing at all.

2.3 Time Domain Analysis

The most common time domain estimate of HRV is the standard deviation of RR intervals (SDNN; normal-to-normal deviation of intervals measured between consecutive sinus beats). SDNN can be calculated, for instance, over 5 min segments (called SDANN) or over 24 h. These two estimates should not be compared because HRV is not a stationary process, that is, a process in which the mean and the variance are independent of the record

length. In long-term recordings, the low-frequency (LF) variations contribute a major proportion of the overall HRV power and also to the SDNN. Since HRV normally decreases at higher levels of the heart rate, SDNN can be normalized against this effect by dividing it by the mean RR interval.

Some commonly used HRV measures are based on the differences between RR intervals, such as the root mean square of successive differences of RR intervals (RMSSD), the number of pairs of adjacent RR intervals differing by more than 50 ms (NN50 count) and the ratio of NN50 count to the count of all RR intervals expressed as a percentage (pNN50). Since all of these measures use RR interval differences, they reflect mainly high-frequency (HF) variations of heart rate and are almost independent of long-term trends. Furthermore, these measures are highly correlated with each other and can thus be considered as surrogates of each other.

All time domain HRV estimates are easily calculated; they do not need time-consuming computation. Also, they do not require stationarity in the same manner as most frequency domain and non-linear analyses do. The main limitation of time domain methods is their lack of discrimination between effects of sympathetic and parasympathetic autonomic branches.

2.4 Frequency Domain Analysis

The main idea behind the frequency domain analysis of HRV is the observation that HRV is composed of certain well-defined rhythms, which are related to different regulatory mechanisms of cardiovascular control. Time domain HRV measures are mainly markers of overall HRV, although some of them can contain information about heart rate oscillations in certain frequency bands (e.g., RMSSD mainly quantifies fast changes in RR intervals). In order to get more detailed information on the dynamics and frequency components of HRV, more advanced analysis methods, such as power spectral density (PSD) analysis, have to be applied. PSD analysis decomposes the signal into its frequency components and quantifies their relative (and also absolute) intensity, named *power*. In other words, it provides estimates of the PSD function of the heart rate, namely, the distribution of frequency components. There are two commonly used methods to compute the PSD function: Fourier transform and autoregressive (AR) modeling.

2.4.1 Fourier Transform

The Fourier transform can be used to convert the time domain data into the frequency domain data and back. The Fourier transform is a one-to-one transform, that is, no information is lost or added; the data just has two different *representations* (Marple, 1987). The normal Fourier transform is defined for continuous functions over the whole real axis. In the case of RR interval time series, the original Fourier transform must be replaced with the discrete version of the transform. The discrete Fourier transform has some unique features which one should be mindful of when using it.

For the equally sampled time series $x(t_k)$, where $t_k = k\Delta$ is the time moment of the data sample, Δ is the sampling interval (the inverse of the sampling rate; see comments below

on the resampling of time series) and $k = 0, 1, 2, \ldots, N-1$, the discrete Fourier transform $X(f_n)$ is

$$X(f_n) = \Delta \sum_{k=0}^{N-1} x(t_k) e^{-2\pi i f_n t_k} = \Delta \sum_{k=0}^{N-1} x(t_k) e^{-2\pi i k n/N}, \text{ where } f_n = \frac{n}{N\Delta}. \quad (2.1)$$

When the time series $x(t_k)$ consists of real values, as it is in our case, $n = 0, \ldots, N/2$. The discrete Fourier transform (Equation 2.1) gives us $N/2$ *complex numbers* (real and imaginary parts*), thus there is an equal number of data points in $x(t_k)$, and $X(f_n)$ since the first and last $X(f_n)$ has only the real part. The power spectral density is $\mathrm{PSD}(f_n) = |X(f_n)|^2$ corresponding to the *squared amplitude* of the frequency component f_n.

The discrete Fourier transform (Equation 2.1) has several important features. First, the frequency scale is discrete, since only f_n components are possible. The *resolution* of frequency scale depends inversely on the number of data samples N and the sampling interval Δ. The highest-frequency component ($n = N/2$), called the Nyquist critical frequency, is $f_c = 1/(2\Delta)$. If the time series $x(t_k)$ is a pure sine wave of a frequency exactly equal to one of the frequencies f_n, only $X(f_n)$ is non-zero, as we can expect. However, if the frequency f of the oscillation is between two adjacent f_ns, there are many non-zero spectral components $X(f_n)$ around f. This is called leakage from one frequency to another in the power spectral estimate, and it is characteristic only of the discrete Fourier transform. This problem can be partially overcome by using *data windowing*.

2.4.1.1 Data Windowing

The discrete Fourier transform (Equation 2.1) can be interpreted as a Fourier transform of the product of the infinitely long time series with a square window function, which turns on at $t = 0$ and off at $t = (N-1)\Delta$. Because of this rapid switching, its Fourier transform has substantial components at higher frequencies, causing the leakage from one frequency to another. To remedy this situation, we can multiply the time series by a window function that changes more gradually from zero to maximum (in the middle of the time series) and then back to zero. There are numerous window functions (named *Welch*, *Parzen*, *Hanning* and so on), but for the purpose of HRV analysis, there is effectively little or no difference between any of them.†

Smoothing

The statistical reliability of the PSD is very low. In fact, the standard deviation (SD) of the PSD estimate is always 100% of the value, and it is independent of N, that is, we cannot get more precise estimates by increasing the number of data points; we just get more discrete frequencies f_n. This means that the amplitude of each spectral component is always unreliable from a statistical point of view. Luckily, in HRV analysis, we are not interested in the amplitude of some specific frequency component but in the spectral power over a certain frequency range. If we want to look at the details of the PSD, the

* Complex number $X(f_n)$ can be interpreted as a two-dimensional vector. Each $X(f_n)$ carries information on the amplitude of the spectral component (the length of the vector) and the phase of the corresponding oscillation (the angle of the vector).

† The difference of the window functions lies in subtle trade-offs among the various figures of merit that can be used to describe the narrowness of the spectral leakage functions.

statistical reliability can be increased by smoothing the data using certain methods, for example, the triangular weighting function, but unfortunately, one pays the price of a lower effective frequency resolution. For a 5 min recording, the smoothing range should not be more than 0.01 Hz.

2.4.1.2 FFT Algorithm

The discrete Fourier transform can be computed directly using Equation 2.1, but such an operation would be very time consuming. The FFT is a very effective *algorithm* to evaluate the discrete Fourier transform (Marple, 1987; Kay and Marple, 1981). It has one limitation: the number of data point N must be an integer power of 2. Since normally the length of the time series is not a power of 2, the data can be padded with zeros up to the next power of 2. This operation corresponds to the interpolation of the original data. Since the actual N used in the FFT calculation is different from the original one, the f_ns are also slightly different.

2.4.2 Autoregressive Modeling

The AR model is based on the idea that the future values of a time series depend linearly on the previous values (Akaike, 1969). This kind of approach is totally generic, and it can be used to model large sets of different systems. The AR model is determined from the following equation:

$$x_k = \sum_{j=1}^{p} a_j x_{k-j} + u_k, \tag{2.2}$$

where x_k is the sample of the time series, a_j is the model parameter, p is the model order and u_k represents the noise, the part of the signal that cannot be explained by the previous values of the signal. The computational task is to find optimal model order p such that parameters a_j of the model describes the system as well as possible. When we have such a model, it is possible to compute the corresponding PSD in the form

$$\mathrm{PSD}(f) = \frac{a_0}{\left|1 + \sum_{j=1}^{p} a_j z^j\right|^2}, \quad z = e^{-2\pi i f \Delta}, \tag{2.3}$$

where f is the frequency and Δ is the sampling interval, as usual. The PSD (Equation 2.3) is capable of modeling sharp peaks since all free parameters are in the *denominator*, which can be made arbitrarily small.* The PSD approximation (Equation 2.3) is also called an *all-poles model* and *maximum entropy method*. There are many algorithms to find the best choices for the parameters a_j, but Burg's method is commonly used. Because the AR spectrum is based on a modeling approach, there are no restrictions for the maximal number of frequency components, that is, PSD can be evaluated using as high a frequency

* In fact, the denominator can be, at least in theory, even zero at certain frequencies. These zeros (also called the *poles* of Equation 2.3) correspond to the spectral peaks.

resolution as one desires (f in Equation 2.3 is actually a continuous variable), in contrast to the Fourier spectrum where the frequency resolution is determined by the number of data samples. Furthermore, the AR model makes it possible to resolve the central frequencies (the frequency location of each peak) analytically. In clinical setups, this feature is rarely useful, but in research work, it could be interesting.

2.4.2.1 Model Order

The model parameters can be estimated if the model order p is fixed. It could be imagined that the model would be best when using as high a model order as possible. However, in practice, time-series data always include some noise (the term u_k in Equation 2.2), so an excessively high model order leads to the model capturing the noise, producing a spectrum with spurious peaks. Too high a model order can also split peaks. Again, too low a model order will strongly smooth spectral peaks, their positions can be shifted and some peaks can even be missing. Some methods have been developed based on information theory for estimating the optimal model order, such as Akaike information criteria (AIC), final prediction error (FPE) or minimum description length (MDL) (Akaike, 1969; Rissanen, 1983; Partzen, 1974). The analytical assumptions underlying these methods are hardly ever fulfilled when analyzing the RR interval time series. Thus, they provide only limited assistance in determining the model order. Therefore, it is recommended that the AR method be used in conjunction with the Fourier transform method (or some other direct method) to help choose the correct model order and to avoid getting spurious spectral features.

2.4.3 Resampling

The RR interval time series is an event series, as stated previously, and each interval has a time stamp, which is the cumulative sum of all preceding intervals. The data can be interpreted as a hypothetical continuous function that has been sampled *unevenly* in time at the moments of R peaks. In spectral analysis, either in Fourier transform or in AR modeling, the input data ($x(t_k)$ in Equation 2.1 or x_k in Equation 2.2) are assumed to be the data sampled evenly in time. Hence, the original RR interval data must be converted to such a form by interpolating each interval and resampling this apparently continuous function.* Interpolation can be linear (each interval is interpolated by a straight line) or based on splines.† The method of interpolation makes no essential difference in HRV analysis.

The resampling frequency must clearly be higher than the *effective sampling rate* of the RR interval data, but it is not critical. The effective sampling rate is equal to the mean heart rate, typically around 1 Hz (= 60 bpm). In practice, a resampling rate of 2–5 Hz is adequate, and there is no need to use a higher rate. In Fourier transform analysis, the highest-frequency component is $f_c = 1/2\Delta$, and it increases by a resampling of the data. However, resampling cannot increase the information content of the data, and thus, the highest relevant frequency component is still half of the (effective) sampling rate according to the general sampling theorem (Nyquist criteria); with 60 bpm mean

* The spectrum of event series can be computed correctly without any interpolation and resampling by using the Lomb–Scargle (also called Lomb periodogram) method. This method makes it possible to calculate the spectrum up to frequencies *above* the Nyquist frequency (= half of the sampling rate or mean heart rate).

† Splines are second-order or higher-order polynomials. If the order is 2, then the spline curve is a parabola with three free parameters, two of them are fixed by demanding that it goes through the data points of the interval and the last one is fixed by demanding that the derivative of the curve is continuous at the data points.

heart rate, $f_c = 0.5$ Hz. This is also true in the case of AR modeling, although the PSD estimate can be evaluated up to any frequency. One should notice that when increasing the resampling rate, the order of the AR spectrum must be increased correspondingly.

Sometimes the resampled data are low-pass filtered using a cutoff frequency of 1 Hz (or half of the resampling frequency). Although this operation can, in theory, improve the statistical reliability of the data, it has no real significance in HRV analysis.

2.4.4 Spectral Powers

The total power (TP) of the RR interval data is represented by the area under the PSD curve from zero to the highest relevant frequency f_c, and it is equivalent to the variance of the signal. In a typical short-time (≅5 min) spectral analysis, the spectral power is divided into three frequency bands: *high frequency* (HF; 0.15–0.40 Hz), *low frequency* (LF; 0.04–0.15 Hz) and *very low frequency* (VLF; 0–0.04 Hz). In long-term recordings, an *ultralow frequency* (ULF; 0–0.003 Hz) component can be calculated. The HF component corresponds to heart rate variations related to the respiratory sinus arrhythmia and HF fluctuations are mediated almost exclusively by fluctuations of efferent parasympathetic activity. The central role of sympathetic nervous system on the LF component is well known, but fluctuations in the LF band are also markedly influenced by parasympathetic nervous system. The LF band is important when characterizing baroreflex sensitivity using spectral methods (see Chapter 3). Oscillations at frequencies in the VLF band are often related to the vasomotor tone of thermoregulation or to the dynamics of hormonal systems, but a precise origin of oscillations in this band is still unknown.

The magnitude of HRV in each frequency band is expressed as power. Since the unit of PSD function is either ms^2/Hz or s^2/Hz depending on the unit of RR interval time series, the unit of HRV power is ms^2 or s^2. Because the TP and power in each frequency band vary considerably even among healthy subjects of equal age, direct comparisons of power between two subjects can be misleading. For this reason, powers are usually expressed in normalized units, by dividing each power by TP less VLF power. Therefore, normalized powers are not absolute measures either, since they partially reflect the relative powers of LF and HF components.*

The confidence we can have in the accuracy of the power in VLF, LF and HF bands varies, since it is related to the number of full periods of oscillations in each band. If we have a 5 min recording, there are roughly 0–12 periods of the oscillation in the VLF band, 12–45 periods in the LF band and 45–120 periods in the HF band. Although the recommended minimum length of recording for a short-term spectral analysis is 5 min, sometimes it is necessary to use shorter segments due to artifacts, ectopic beats and so on. As a rule of thumb, one can use the criterion that the minimum number of periods be six.† In order to have a reliable power estimate, for instance, in the whole LF band, we need at least 2.5 min worth of usable recording. Similarly for ULF power, the minimum recording length is about 1 h.

A typical example of spectral HRV analysis performed with FFT and AR modeling is presented in Figure 2.1. The length of the RR interval time series was 5 min, and it consisted of 349 intervals. The mean RR interval was 856 ms, corresponding to a heart rate of 70.2 bpm. The spectral analysis was done over the whole time series. Before any analysis,

* If we have two cases with equal absolute LF powers but different HF powers, the normalized LF powers are not equal.

† This is actually an arbitrary number of periods, but in practice, it has proven to be a useful minimum number.

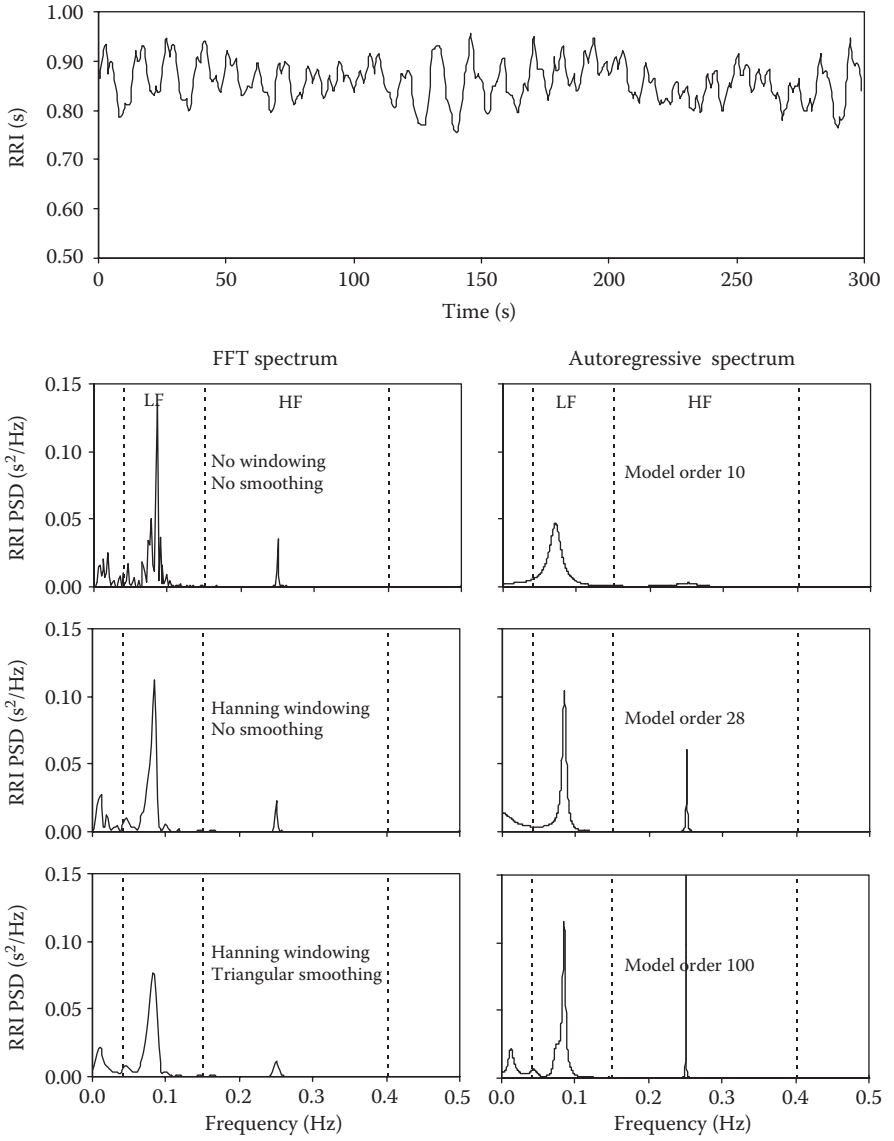

FIGURE 2.1
RR interval time series and corresponding FFT and autoregressive spectrum.

the linear trend was removed. The time series was linearly interpolated and resampled using a sampling frequency of 2 Hz. According to the mean heart rate, the highest relevant frequency f_c was 70.2 (beats/min)/60 (s/min)/2 = 0.585 Hz, well above the upper end of the HF band. The frequency resolution of the FFT spectrum is 0.00195 Hz.* There are clear peaks in both the LF and the HF bands in all versions of the FFT spectrum (left panels). The HF component is very sharp since metronome-controlled breathing was used.

* After resampling at the sampling rate of 2 Hz (the sampling interval $\Delta = 0.5$ s), the 5 min time series has 600 data points. This has been zero padded to the next power of 2; here, it is 1024. The frequency resolution is $1/N\Delta = 1/(1024 \times 0.5\text{ s}) = 0.00195$ Hz.

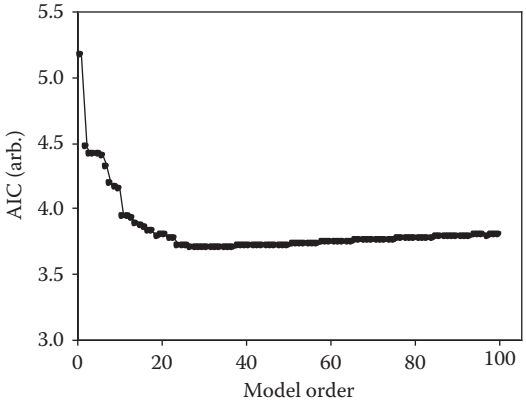

FIGURE 2.2
The Akaike information criteria (AIC) of the time series of Figure 2.1 as a function of the model order. The minimum of AIC is reached at the model order of 28.

Without windowing or smoothing, the FFT spectrum clearly includes HF components, which disappear after using the Hanning window function. By also applying triangular smoothing (range 0.01 Hz), the spectrum has even less detail, but the statistical reliability of the spectral features has increased. The AR spectrum (right panels) with the model order of 10 cannot capture all essential features. If the model order is increased to 28 (the optimal model order according to AIC, that is, the order at which AIC is at its minimum value, see Figure 2.2), the spectrum resembles the FFT spectrum. If the model order is further increased up to 100, more spectral details can be seen, especially in the lowest frequencies. It is remarkable how similar the spectral details below 0.05 Hz are in both FFT (windowed and smoothed) and AR spectra (model order 100). Although AIC has a minimum value at 28, it increases very slowly with increasing model order. Thus, it is not always obvious which model order should be selected.

2.4.6 Effects of Respiration

The respiratory component of HRV depends strongly on the breathing volume and its frequency. The HF component decreases markedly as breathing volume decreases. In addition, the power of respiratory peak increases as its frequency decreases. If the breathing frequency is below 0.15 Hz, it becomes measured not as the HF component but as a part of the LF component. The effect of the breathing frequency on HRV is presented in Figure 2.3. The breathing volume was fixed in all cases. When the breathing frequency was fixed at 0.10 Hz (the upper right panel in Figure 2.3), the amplitude of the respiratory peak was approximately 0.25, ten times higher than at the breathing frequency of 0.25 Hz (the upper left panel; note the different scale in the y-axis). At a breathing frequency of 0.05 Hz, the peak was still three times higher (the lower left panel).

In practice, the HF component can be used as a measure of parasympathetic tone and vagal activity only in situations in which the breathing frequency and volume are carefully controlled. Controlled respiration at a constant rate, however, can induce some stress, which might affect the function of the autonomic nervous system and therefore interfere with HRV. Using a respiratory frequency that is as close as possible to the natural breathing rate of the subject, one can minimize the stress. This can be accomplished by first measuring the

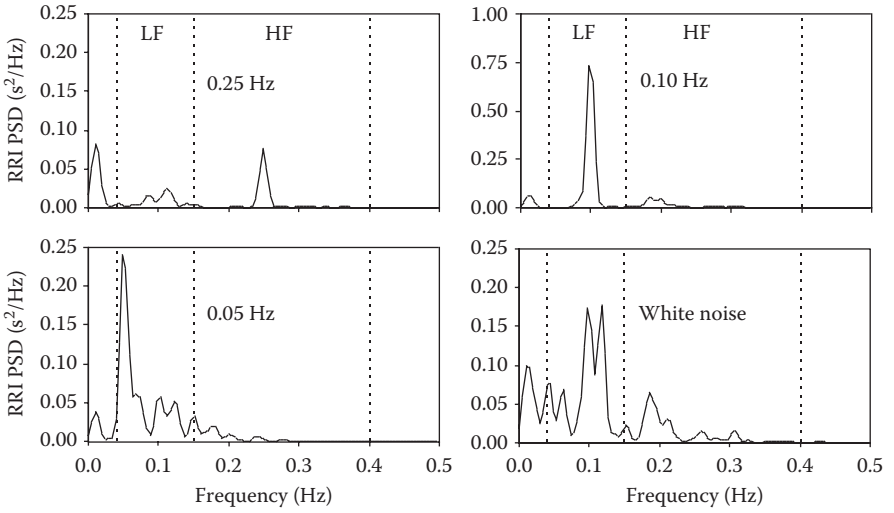

FIGURE 2.3
RR interval FFT spectra at three different breathing frequencies of 0.25, 0.10 and 0.05 Hz and using white noise breathing. Note that the *y*-scale is different in the case of 0.10 Hz.

breathing rate and then adjusting the frequency of metronome accordingly. Another possibility is to use white noise breathing, that is, a metronome guides the subject to breathe at variable rates with constant distribution (see the right bottom panel in Figure 2.3).

2.4.7 Time–Frequency Analysis

When considering perturbation of the autonomic nervous system, usual steady-state spectral analysis methods are no longer useful, since they cannot determine exactly when critical frequency components change in amplitude or frequency. Conversely, if information about temporal variations of frequency components were available, spectral features of slowly changing system could be analyzed in more detail, for instance, to estimate the stationarity of a time series. There are many different approaches to perform a time–frequency analysis, and most of them are based on various integral transforms of the original data. The main limitation on all of them is the trade-off between temporal resolution and statistical reliability of spectral components. It is not possible to determine the amplitude of a certain frequency locally because we always need a time window to measure it. A shorter time window can capture more rapid changes within the power spectra in time, but a longer time window can give a more reliable estimate of the amplitude or spectral power.

2.4.7.1 Windowed Fourier Transform

The most simple and straightforward time–frequency analysis is called the *windowed Fourier transform* (WFT, also called short-time Fourier transform or STFT). In this approach, the time series is covered by a reasonably short time window that slides over the data in small steps. In each window, the PSD is calculated as described above (e.g., using data windowing, zero padding, trend removing and resampling). As a result, one gets a set of PSDs. In Figure 2.4, we present the results of a WFT performed on the same RR interval data shown in Figure 2.1 as a three-dimensional plot. The width of the window was 60 s

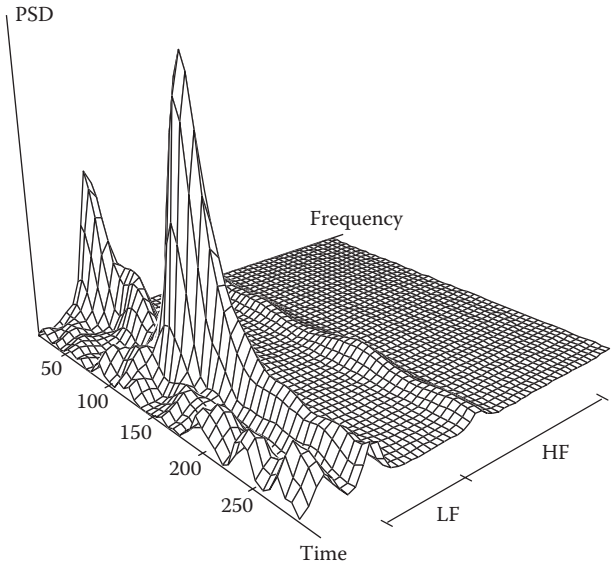

FIGURE 2.4
Time–frequency analysis of the RR interval time series using windowed Fourier transform method.

and the step 5 s. The RR interval has prominent low-frequency oscillations around 140 s (see the upper panel in Figure 2.1), and this can be also seen as a large peak in Figure 2.4. Furthermore, the respiratory modulation (the shallow peak in the HF band) seems to be almost constant in time. Although the window step is 5 s, the real-time resolution of the analysis is still approximately 60 s. If there had been an abrupt change in the spectral features of the signal, it would be seen in all the windows covering that moment in time. The same analysis is presented as a contour map in Figure 2.5. This representation reveals that there are a lot of variations in the VLF and LF bands.

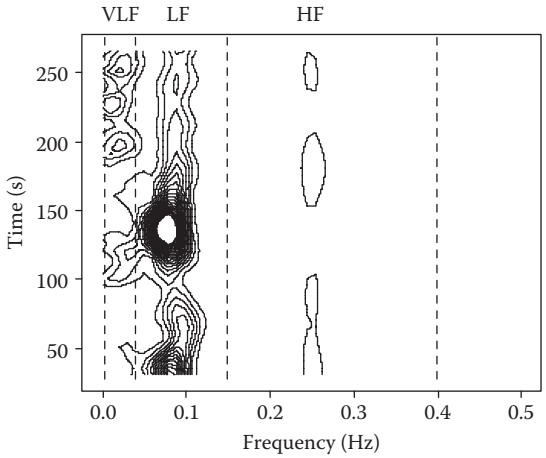

FIGURE 2.5
Same as Figure 2.4 but displayed as a contour map.

Since the number of periods within the window of WFT depends on the frequency, the statistical reliability is different for each frequency component as discussed in Section 2.4.4. If the width of the window is set for the lower frequency components, it becomes unnecessarily wide for higher frequencies. This feature can be avoided if the duration of the window is inversely proportional to the analyzed frequency. This kind of analysis is called *selective discrete Fourier transform algorithm* (SDA) and can be implemented by normal discrete Fourier transform, but it is computationally intensive (Keselbrener and Akselrod, 1996).

2.4.7.2 Wigner–Ville Distribution

The Wigner–Ville bilinear distribution (WVD) of the continuous function $x(t)$ is defined by

$$\mathrm{WVD}(t,f)=\int_{-\infty}^{\infty} x\left(\tfrac{t+\tau}{2}\right)x^{*}\left(\tfrac{t-\tau}{2}\right)e^{-2i\pi f\tau}\mathrm{d}\tau\,. \tag{2.4}$$

The WVD maps a one-dimensional function of time into a two-dimensional function of time and frequency (Claasen and Mecklenbräuker, 1980a,b; Hyung-Ill and Williams, 1989; Novak and Novak, 1993). For a discrete time series $x(n)$, the distribution (Equation 2.4) can be written in the form

$$\mathrm{WVD}(n,k)=\sum_{\tau=-\infty}^{\infty} W_N(\tau)e^{2\pi ik\tau/N}\left[\sum_{\mu=-\infty}^{\infty} W_M(\mu)K(\mu,\tau)x(n+\mu+\tau)x^{*}(n+\mu-\tau)\right], \tag{2.5}$$

if the kernel function K and W is equal to 1. This distribution is also called a *smoothed windowed* WVD because of the window functions W_M and W_N. The function W_M is a rectangular window that has a value of 1 for the range of $-M/2 < \mu < M/2$, and W_N is a symmetrical window (often also rectangular), which has non-zero values for the range $-N/2 < \tau < N/2$. The parameter N determines the frequency resolution of the WVD, while the parameter M determines the level of temporal smoothing.

The frequency resolution of the spectrum calculated by the WVD is two times the resolution of the FFT spectrum when using the same length of time window, therefore making WVD more suitable than FFT for shorter time series. However, WVD has another problem: if there are two frequency peaks f_1 and f_2 close to each other, the WVD spectrum also contains small but significant erroneous components at the frequencies $f_1 - f_2$ and $f_1 + f_2$. In the case of multiple dominant frequency peaks, the background of WVD spectrum can be contaminated by many spurious frequency peaks of low amplitude. This feature of the WVD method forces us to be careful when interpreting the finer details of the spectrum. The cross-terms can be suppressed by introducing a suitable non-trivial kernel function K in Equation 2.5, but unfortunately at the price of lower frequency resolution. One example of such a kernel function is the so-called *exponential distribution*.

2.4.7.3 Complex Demodulation

The complex demodulation (CDM) method is a common non-linear method used to define the amplitude of a time series at a specified frequency or frequency band as a function of time (Kim and Euler, 1997). In other words, for a given frequency ω, we assume the signal is of the form

$$x(t) = A(t)\cos\left(\omega t + \phi(t)\right) + z(t), \tag{2.6}$$

for which we need to determine the amplitude $A(t)$ and the phase $\phi(t)$. The term $z(t)$ contains all other oscillating components (having a frequency different from ω) and possible noise. In the CDM method, the original real value signal $x(t)$ is rewritten into complex format

$$x(t) = 0.5A(t)\left\{e^{i\left(\omega t+\phi(t)\right)} + e^{-i\left(\omega t+\phi(t)\right)}\right\} + z(t), \tag{2.7}$$

where i is the imaginary unit. In the next step, all frequency components are shifted by $-\omega$. This operation is equivalent to multiplying $x(t)$ by the term

$$y(t) = 2e^{-i\omega t}, \tag{2.8}$$

which gives us

$$x'(t) = A(t)e^{i\phi(t)} + A(t)e^{-i\left(2\omega t+\phi(t)\right)} + 2z(t)e^{-i\omega t}. \tag{2.9}$$

Here, we see that the frequency of the first term is zero and the frequency of the second term is twice that of the component under study. The last term does not contain frequencies around zero, because the component $z(t)$ did not originally contain the frequency ω. If the signal $x'(t)$ is fed into a low-pass filter with a cutoff frequency at zero, we get the signal

$$x''(t) = A(t)e^{i\phi(t)}, \tag{2.10}$$

from which we can easily calculate the slowly and time-dependently changing amplitude $A(t)$ since $A(t) = |x''(t)|$. Likewise, we can get the amplitude of any desired frequency signal as a function of time (by varying ω). If the cutoff frequency is $\Delta\omega$ and not exactly zero, the CDM method picks from the signal only the amplitude of those components with a frequency between $\omega - \Delta\omega$ and $\omega + \Delta\omega$. In this way, it is possible to pick the part of the signal that belongs, for example, to the LF or HF band.

Therefore, theoretically, the CDM method can give the amplitude at every moment of time, but in reality, the temporal resolution depends upon the characteristics of the low-pass filter. If the desired frequency band is to be limited as steeply as possible, one needs to use a higher-order filter, but in that case, one needs more data points to perform filtering, and the time resolution will be lower. In practice, time resolution of the CDM method applied to the RR interval time series is approximately 15 s when analyzing oscillations in the LF band. Any faster changes in the amplitude of oscillations cannot be clearly distinguished. This temporal resolution is, however, clearly better than the resolution obtained with other methods.

The CDM amplitudes of the RR interval data (the same time series as in Figure 2.1) in the LF and HF bands are shown in Figure 2.6. Again we see that there are significant changes in the LF component, but the HF component is almost constant.

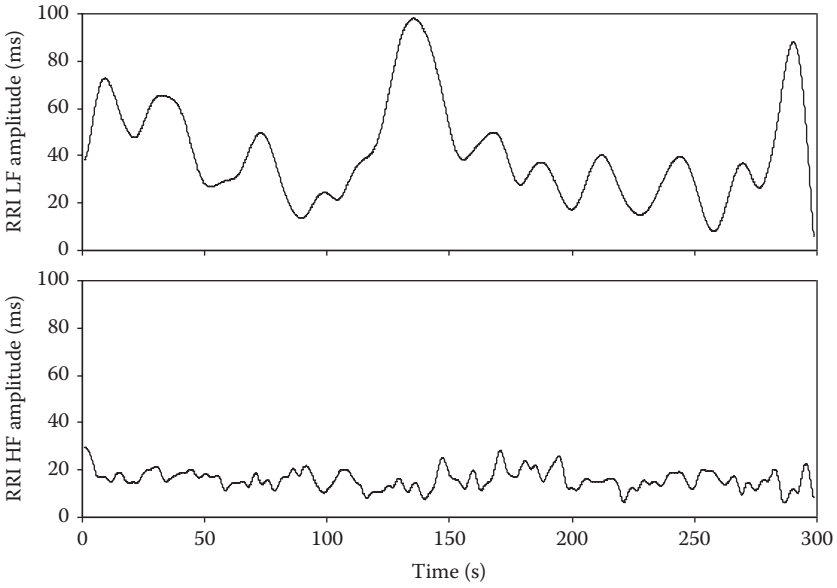

FIGURE 2.6
The amplitude of the LF and HF components of the RR interval time series (the same data as in Figure 2.1) as a function of time based on CDM method.

2.4.7.4 Other Methods

In the method of sliding window analysis, the Fourier transform can be replaced with AR modeling. If there are significant changes in the structure of HRV data as a function of time, the optimal model order should be obtained individually in each window. A similar approach called *time-variant autoregressive modeling* is suitable for online monitoring (Bianchi et al., 1997). In this method, a new set of AR parameters is computed whenever a new sample value is available, and the weight of the previous samples is controlled by means of a forgetting factor.

The *Wavelet transform* is a general tool for analyzing the temporal changes in time series (Figliola and Serrano, 1997). The continuous wavelet transform (CWT) of the time series $x(t)$ is defined as

$$\mathrm{CWT}(a,\tau) = \frac{1}{\sqrt{a}} \int x(t) \Psi\left(\frac{t-\tau}{a}\right) \mathrm{d}t \,, \tag{2.11}$$

where $\Psi(t)$ is the basic (or mother) wavelet. The wavelet transform, similar to STFT, maps a time function into a two-dimensional function of a and τ. The parameter a is called the scale*; it scales the wavelet function by compressing or stretching it. τ is the translation of the wavelet function along the time axis. There are an infinite number of valid wavelet functions, but all of them are well localized.† By contrast, the STFT uses truncated sine waves, which are not well localized. The shape of the wavelet function must be selected according to the application. Furthermore, by increasing the window width in STFT, the

* Inverse of the scale can be interpreted as a frequency.
† The wavelet is well localized since it goes rapidly to zero in infinity.

number of periods increases, but in the wavelet transform, basic shape of the wavelet is same; it is only compressed or stretched. All of these features of wavelets make them ideal for capturing rapid temporal changes. Although wavelet analysis has been applied to the HRV signal, results are not superior to those obtained using more traditional approaches (Wiklund et al., 1997).

2.5 Non-Linear Analysis

2.5.1 Approximate Entropy and Sample Entropy

All frequency domain analyses are based on the recognition of certain predetermined patterns. For instance, in Fourier transforms, the pattern is a sinusoidal wave, and in wavelet analysis, the pattern is a certain wavelet function. Another alternative for characterizing the variability of heart rate is to measure the regularity or complexity of the fluctuations without specifying the form of repeating patterns. Entropy is a general approach for quantifying the regularity or information content of the data (Pincus and Goldberger, 1994; Pincus, 1995; Bettermann and van Leeuwen, 1998; Cysarz et al., 2000). There are many ways to determine the entropy of a time series, but most of them require noise-free data and very long recordings. *Approximate entropy* (ApEn) has been developed for measuring the complexity of relatively short time series and it is most useful for HRV analysis where long noise-free recordings are difficult to obtain (Pincus and Goldberger, 1994). Furthermore, ApEn calculations are not based on specific assumptions regarding the internal structure or dynamics of the system.

ApEn is computed as follows. First we construct the so-called pseudo phase space vectors from the initial time series $x(i)$, in which $i = 1 \ldots N$, N being the number of data points

$$u(i) = \left[x(i), x(i+1), x(i+2), ..., x(i+m-1) \right], \tag{2.12}$$

m being the so-called *embedding dimension*.* Vectors $u(i)$ can be interpreted as m-point patterns. First, we select one m-point pattern and search for similar m-point patterns. Two patterns are similar when the maximum distance d between the corresponding components is less than the *tolerance* r (Figure 2.7):

$$d\left[u(i), u(j) \right] = \max \left\{ \left| u(i+k) - u(j+k) \right| : 0 \le k \le m-1 \right\} \le r \,. \tag{2.13}$$

The (normalized) number of similar vectors $u(j)$, which are at a distance r from $u(i)$, is

$$C_i^{(m)}(r) = \left\{ \text{the number of index } j \text{ for which, } j \le N-m+1, \, \mathrm{d}\left[u(i), u(j) \right] \le r \right\} \Big/ (N-m+1). \tag{2.14}$$

* The use of pseudo phase vectors is a starting point of many non-linear analysis methods. The idea of these vectors is that they can be used as a *replacement* for the set of true dynamical variables, all of which we cannot measure either for technical reasons or since we actually do not know them. Under certain assumptions, the dynamics of the pseudo phase vectors is similar to the real dynamical variables.

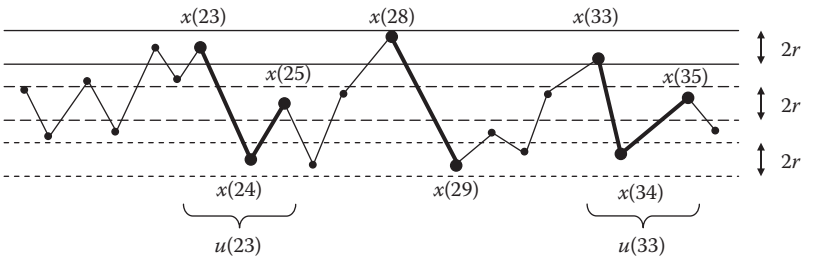

FIGURE 2.7
Approximate entropy. An example showing for embedding dimension $m = 2$, a search for similar pseudo phase space vectors. For vector [$x(23)$, $x(24)$], two nearby vectors, [$x(28)$, $x(29)$] and [$x(33)$, $x(34)$], are found to be similar, that is, the distance between $x(28)$ and $x(33)$ and the distance between $x(29)$ and $x(34)$ are both < tolerance value r. Therefore, both $u(28)$ and $u(33)$ increase the quantity $C_{23}^{(2)}(r)$. However, when m is increased to 3, only vector [$x(33)$, $x(34)$, $x(35)$] is similar to [$x(23)$, $x(24)$, $x(25)$] and increases $C_{23}^{(3)}(r)$.

Due to normalization, the maximum value of C is 1, and C can be regarded as the probability of finding similar m-point patterns. The same analysis can be performed for all m-patterns. The logarithmic average of probabilities over all m-patterns is

$$\Phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r). \tag{2.15}$$

Approximate entropy is defined as

$$\mathrm{ApEn}(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r). \tag{2.16}$$

One sees that ApEn measures the (logarithmic) conditional probability that similar m-patterns are similar also when looking for (m + 1)-patterns. In other words, ApEn is the averaged probability of finding m-patterns minus the averaged probability of finding (m + 1)-patterns (see Figure 2.7). If the last probability is equal to the first one, ApEn = 0, i.e., the time series is absolutely regular in this sense. However, if the last probability is zero, ApEn gets its maximum value, i.e., the time series is totally irregular.

ApEn(m,r,N) is dependent on three parameters: the length m of the vectors being compared, the tolerance parameter r, and the number N of data points. This means that direct comparisons always require fixing of parameters. For HRV analysis, $m = 2$ is the value normally used. When the number of data points is increased, ApEn approaches its final value asymptotically. In practice, $N > 800$ and $m = 2$ give a reliable result. ApEn depends strongly on the tolerance parameter r. If r is chosen such that it is a fraction of the SD of the data, ApEn does not depend on absolute variability.* The most frequently used value of r is 15 or 20% of SD.

ApEn is sensitive to smallest trends in the data, because comparison of patterns is based on the absolute values of data. The trend can be removed before ApEn analysis but with

* If r is a fraction of SD, ApEn does not depend on the unit of data, and therefore, it is also possible to compare the ApEn values of different signals, such as RR interval and systolic pressure time series.

cautions as described in Section 2.2.5. One alternative is to use the differentiated data (the difference of successive RR intervals) that eliminate all slow trends, but this operation behaves like a special high-pass filter in the frequency domain. ApEn is not sensitive to changes in single data values if the tolerance parameter is fixed, but if the tolerance parameter is bound to the SD as recommended, the situation could well be different. Ectopic beats especially, if not edited, can alter the SD significantly.

When calculating the number of similar vectors to get ApEn, the similarity of the vector to itself is included in the calculation. This ensures that $C_i^{(m)}(r)$ is non-zero, which is essential for calculating the logarithm. This causes ApEn to give a result most of the time, which implies greater regularity of the signal than may be present.

Sample entropy (SampEn) is calculated in a way that removes the previously described bias (Richman and Moorman, 2000). When calculating the number of nearby vectors, a comparison to the vector itself is prevented:

$$C_i'^m(r) = \left\{\text{the number of index } j \text{ for which, } j \neq i, j \leq N-m+1, \mathrm{d}\left[u(i), u(j)\right] \leq r\right\} \Big/ (N-m+1). \tag{2.17}$$

The average of the probabilities Φ is also defined without logarithms

$$\Phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} C_i'^m(r). \tag{2.18}$$

Now SampEn is defined as

$$\mathrm{SampEn}(m, r, N) = \ln\left(\Phi^m(r) / \Phi^{m+1}(r)\right). \tag{2.19}$$

The interpretation and use of SampEn remain exactly the same as for ApEn. However, the dependence on the tolerance parameter r and the number of data points N is different. ApEn reaches its maximum with a certain value of r, but SampEn decreases monotonically as r increases. SampEn is, in principle, also independent of the number of data points N, but with small values of N, its statistical reliability is naturally poor. When r and N are large enough, SampEn and ApEn yield the same result. SampEn provides a more reliable estimate of the complexity of a signal compared to ApEn. It may be used for considerably shorter time series than the ApEn, (<200 points).

Regularity and complexity are often interpreted as being contrary to each other: increased regularity means lower complexity and vice versa, but this is not always true. The degree of regularity can be quantified by evaluating the appearance of repetitive patterns and characterized by entropy measures. Complexity is, however, intuitively associated with *meaningful* structural richness, which can exhibit relatively high regularity. Entropy-based measures grow monotonically with the degree of randomness and they reach highest values from totally uncorrelated random data or white noise. Such data are unpredictable but not actually complex. Thus, entropy measures may lead to misleading results when they are applied to physiological time series such as the heart rate signal. For example, atrial fibrillation (AF) is associated with highly erratic fluctuations with statistical properties resembling uncorrelated noise. The entropy value of such

a signal is high. By contrast, healthy cardiac rhythms that are regulated by multiple interacting feedback mechanisms will yield lower entropy values. This inconsistency is obviously related to the fact that entropy measures are based on single-scale analysis, but many biological systems operate across multiple spatial and temporal scales and hence their complexity is also multiscaled.

To overcome the aforementioned difficulties in interpreting entropy measures, *multiscale entropy* (MSE) analysis has been introduced (Costa et al., 2002, 2005). In this method, the coarse-grained time series determined by the scale factor τ is defined as

$$y^{(\tau)}(j) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x(i), \tag{2.20}$$

where $x(i)$ is the original time series and $1 \le j \le N/\tau$. For a scale of 1, the time series $y(j)$ is simply the original time series. The length of $y(j)$ is equal to the length of the original time series divided by the scale factor. In the next step, the entropy of the coarse-grained time series is calculated as a function of the scale factor. In principle, the entropy can be calculated using any method that is reliable for a time series of variable length, and SampEn is a good choice when analyzing the RR interval time series. In order to have good statistical reliability at higher scales, the number of data points must be greater than 10,000. This limits the use of MSE in many clinical studies.

As an example, the MSE method has been applied to three different subjects: one healthy, one with congestive heart failure (CHF) and one with AF (see Figure 2.8). At the scale of 1, the healthy and CHF cases cannot be separated, but if the scale is 5 or 6, large separation can be obtained. Entropy of AF time series is highest at a scale of 1, but it decreases monotonically as the time scale increases, similar to white noise. At a very large time scale, CHF and AF time series cannot be separated any more. We conclude that MSE can distinguish these cases, but it needs computation of the entropy as a function of scale; no single entropy value at a fixed scale is enough.

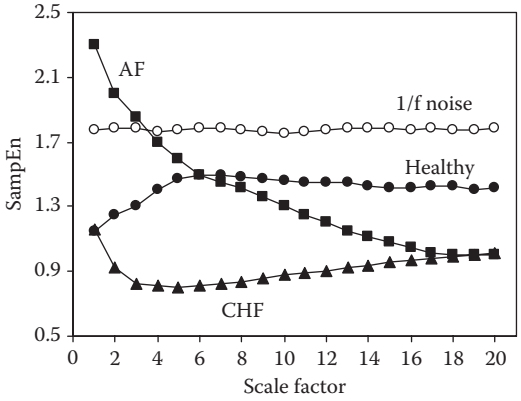

FIGURE 2.8
Multiscale entropy analysis of an RR interval time series derived from healthy subjects, congestive heart failure (CHF) subjects and subjects with atrial fibrillation (AF). As reference, the results of simulated 1/*f* noise are presented.

2.5.2 Scaling Exponents

2.5.2.1 Detrended Fluctuation Analysis

When analyzing a longer time series lasting several hours, identification of oscillatory components or repeated patterns are not the best approach for examining the HRV data. One method is to characterize the internal correlations of the signal. These correlations are expressed by scaling properties and fractal structures. Detrended fluctuation analysis (DFA) presents a possibility for characterizing this as a function of correlation distance (Peng et al., 1993, 1995; Iyengar et al., 1996).

To calculate DFA, we must first form an integrated version of the original time series $x(i)$, where $i = 1 \ldots N$, which gives us

$$y(k) = \sum_{i=1}^{k} \left(x(i) - < x > \right), \tag{2.21}$$

where $<x>$ is the mean of the original time series and $k = 1 \ldots N$. Next, we divide the time series $y(k)$ into equally spaced segments with length n as shown in Figure 2.9. For each segment, we calculate separately the local trend by fitting a regression line $y_n(k)$ to the segment. The RMS (root-mean-square) fluctuation of the integrated time series is calculated by removing the linear trend of each segment. Thus,

$$\mathrm{DFA}(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left[y(k) - y_n(k) \right]^2} . \tag{2.22}$$

In the summation, we must take into account that when the index k is stepped, $y_n(k)$ must be updated when moving into the next segment. DFA is calculated for several different segment lengths, that is, n values. Typically, DFA increases when the segment length increases. If log(DFA) increases linearly as a function of log(n), the time series follows (fractal) scaling law, and in this case, the slope α of the linear change, the *scaling exponent*,

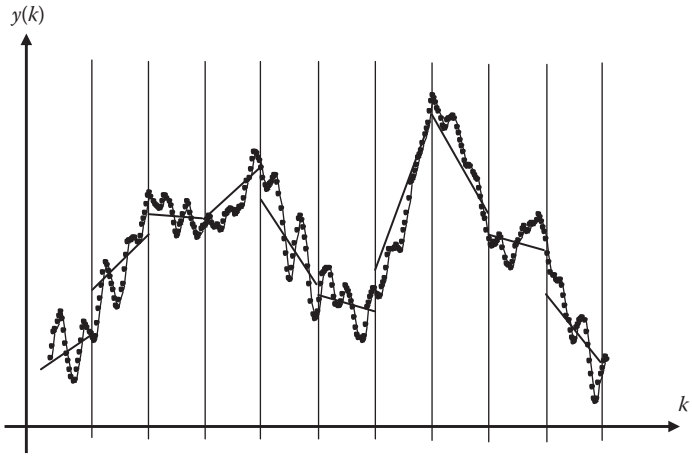

FIGURE 2.9
The integrated time series and the local trends.

TABLE 2.1

Scaling Exponent α of Certain Type of Time Series

Scaling Exponent	Description of the Signal
$0 < \alpha < 0.5$	Small value followed most probably by a larger value and vice versa
$\alpha = 0.5$	Completely uncorrelated time series, that is, white noise
$0.5 < \alpha < 1.0$	Small value followed most likely by a small value and large value followed most likely by a large value (correlated)
$\alpha = 1.0$	$1/f$ type noise
$1.0 < \alpha < 1.5$	Noise of variable type
$\alpha = 1.5$	Brownian $1/f^2$ noise (integral of white noise)

defines the type of scaling. Different values of α correspond to specific types of time series as presented in Table 2.1.

A typical DFA of long RRI time series is shown in Figure 2.10. A heartbeat has short-range correlations reflecting the baroreflex mechanism, as well as long-range correlations, which are related to the efforts to keep the variation of the beat cycle within certain limits. The measurement of the long range correlations requires that the time series under investigation be preferably at least a few hours long so that the statistical reliability would be at least reasonable. The limit for short- and long-range correlations is set typically to 10 or 11 beats (corresponding to 2.4 on the logarithmic scale). The long-range scaling exponent α_L for a healthy patient is ≈ 1, which corresponds to $1/f$ behavior. The short-range scaling exponent α_S may vary, but it is usually between 0.5 and 1.5. Many factors affect it, such as the functioning of the baroreflex mechanism, breathing modulation and so on. With a longer time series, there always exists the possibility that the measured correlations are not at all a characteristic of the system but rather reflect environmental effects.

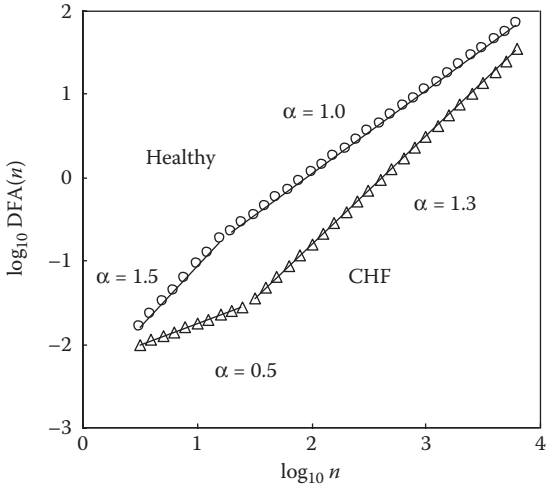

FIGURE 2.10
DFA as a function of the number of segments for healthy and CHF subjects.

2.5.2.2 Spectrum Power Law Exponent

Long-range correlations of a time series may also be analyzed using the spectrum of the signal. In this case, we can study the lowest-frequency components of the spectrum and try to characterize its shape using simple exponential law. If we presume that for a certain frequency spectrum we have $1/f^{\beta}$, the scaling exponent β can be calculated by presenting the spectrum on a log–log scale and by fitting a line over the desired frequency range (Iyengar et al., 1996; Bigger et al., 1996). The slope of the line gives the spectrum power law exponent. The value of the exponent varies between 0 and 2. The border line case 0 corresponds to a flat spectrum, that is, white noise, and the value 2 corresponds to Brownian noise. Usually, the frequency range 0.0001–0.01 Hz of the spectrum is studied, and this corresponds to an oscillation period of 1 min to several hours. For the above definition to make sense, the HRV time series must be several hours long. The spectrum is calculated almost without exception using the FFT algorithm. Because the spectrum has a rather irregular form, especially at the lowest frequencies, the use of some smoothing method is desirable. Replacing the regression line with a less sensitive fitting method may also improve the reliability of the result.

2.5.3 Fractal Dimensions

2.5.3.1 Correlation Dimension

The dynamics of a system can be described by measuring its attractor (the path toward which the system converges) *dimension*. Especially for chaotic systems, the attractor can be fractal, in which case its dimension is not an integer. Knowing the dimension of the attractor may also help getting useful information about the characteristics of underlying systems. The correlation dimension (CD) is one of the simplest methods for estimating the attractor dimension (Grassberger and Procaccia, 1983; Kantz and Schreiber, 1995; Yum et al., 1999). The CD is sometimes referred to with the designation *D*2.

The basis for the calculations once again is in the reconstruction of time series in the multidimensional phase space $x(i)$, where $i = 1 \ldots N$, by using the vectors of the pseudo phase space $u(i) = [x(i), x(i+1), x(i+2), \ldots, x(i+m-1)]$, where m is the embedding dimension. Next, we calculate for each vector $u(i)$, how many attractor points are at a distance r as measured from the observation point

$$C_i^m(r) = \left\{\text{the number of index } j \text{ for which, } j \le N-m+1,\ \mathrm{d}\left[u(i), u(j)\right] \le r\right\} \Big/ (N-m+1), \tag{2.23}$$

where the distance d is defined (i.e., differing from the ApEn method) as the normal Euclidian distance

$$\mathrm{d}\left[u(i), u(j)\right] = \left(\sum_{k=1}^{m} \left| u(i;k) - u(j;k) \right|^2 \right)^{1/2}. \tag{2.24}$$

Next, we calculate the mean of the quantities $C_i^m(r)$ over all vectors, from which we compute the so-called correlation integral

$$C^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} C_i^m(r). \tag{2.25}$$

CD is defined as a limit

$$\mathrm{CD}(m) = \lim_{r \to 0} \lim_{N \to \infty} \frac{\log C^m(r)}{\log r}. \tag{2.26}$$

In practice, with limited data sets, these limits cannot be calculated with certainty and therefore the CD is defined as the slope of the regression line calculated from a log–log representation and over a range with the required linearity.

When calculating CD, the embedding dimension m must be selected so that it is at least $2D$, where D is the dimension of the system under study, that is, the number of real dynamical variables. In order for the correlation integral to describe the attractor accurately, the number of data points should exceed 10^m. For example, when studying the blood pressure regulation system, we may assume that the number of dynamical variables is >4, which means that the time series must be very long. In addition, it is nearly impossible to find such a range of the distance r, in which $\log C^m(r)$ changes linearly as a function of $\log r$, because of the noise contained in the data and non-stationarity of the data. Due to these limitations, the calculation of the correct CD for biosignals is computationally not easily achieved. Despite this fact, computation of the correlation dimension may still be useful. Promising results have been achieved by using $m = 20$ and by searching for the mean slope within $0.01 < C^m(r) < 0.1$. The quantity calculated in the above fashion without forgetting the aforementioned limitations is called the *modified correlation dimension*. This quantity cannot accurately define the real dimension of the system, but nevertheless, it does give a measure of the complexity of the system, that is, when CD increases, the system becomes more complex.

2.5.3.2 Pointwise Correlation Dimension

Pointwise correlation dimension (CDi) is defined in a very similar way compared to CD, but instead of labeling the time series with a single value, it is calculated as a function of time (Farmer et al., 1983; Mayer-Kress et al., 1988). This gives us the possibility of evaluating changes in the system characteristics as a function of time, which is very important in non-stationary cases. CDi is sometimes referred to as D2i.

When searching for the regression line, we must once again select the range in which the relation is linear and use a high enough embedding dimension value. When calculating the CDi, we can set $m = 20$ (similar to computing the CD above) and select the area $0.01 < \mathrm{CDi}(r) < 0.1$. In addition, we must note that even though CDi is calculated at each point and it can therefore in principle follow changes in data, the calculation of CDi at any point requires its calculation at all other points. For this reason, CDi is not applicable to non-stationary time series. Thus, in practice, it is advisable to use an additional condition that states that when calculating the regression line at each point in a log–log representation, the correlation factor of the achieved line must exceed a certain limit (e.g., 0.8), and if this criterion is not fulfilled, the CDi value at the point in question is not reliable.

2.5.3.3 Dispersion Analysis

Dimension analysis of a time series may also be performed by studying the curve describing the time series itself rather than the dynamical system behind the signal. This approach

toward analyzing RR intervals is similar to image analysis. Because complex behavior of the dynamical system manifests itself in complicated patterns in the measured time series, a study of the curve's fractal structure will also give information regarding the system itself.

In dispersion-based analysis, we first calculate from the time series, the standard deviation

$$\mathrm{SD}(1) = \frac{1}{N}\sqrt{N\sum_{i=1}^{N} x^2(i) - \left(\sum_{i=1}^{N} x(i)\right)^2}\,, \tag{2.27}$$

where $x(i)$ is the time series having N data points. Next, we compute the mean of two consecutive data points, (for the entire time series) resulting in a new time series of $N/2$ data points. For this new time series, we again calculate the standard deviation SD(2). This is continued with group sizes 4, 16, 32 and so on until there are less than 4 data points left in the time series. By repeating the above process, we have a series of standard deviation values SD(m). When we plot log SD(m) as a function of log m, a line can be plotted through the points if the original time series curve was fractal. For this fractal object, the dimension is FD−DA = 1 – the slope of the line (Bassingthwaighte and Raymond, 1995; Yum et al., 1999). Fractal dimension defined in this way can have values between 1 and 1.5, where 1 represents the case of a steady-state signal and 1.5 represents maximal fractal characteristics.

If the curve under study is not purely fractal, that is, it contains a sizeable amount of noise, the SD(m) values do not exactly fit a line in log–log representation. For automated analysis when this line is not visually verified, one can set a minimal correlation of 0.8 as a measure of linearity.

2.5.4 Return Map

Dynamical systems are usually described by a group of differential equations. If the variables get values only at specific discrete moments in time, as is the case with the RR interval time series, the differential equations can be replaced with discrete equations, for example,

$$\begin{aligned} x_{i+1} &= F(x_i, y_i, z_i, \ldots) \\ y_{i+1} &= G(x_i, y_i, z_i, \ldots) \\ z_{i+1} &= H(x_i, y_i, z_i, \ldots) \\ &\vdots \end{aligned}\,, \tag{2.28}$$

where x, y, z and so on are dynamical variables of the system and F, G, H, ... are functions that define the dynamics. Usually, these functions are not known, but we may try to solve for them by examining the measured time series. If there is only a single variable, the equation is expressed in a simpler form:

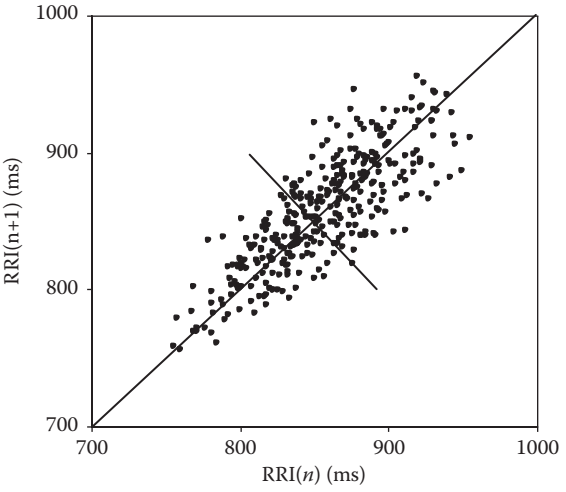

FIGURE 2.11
A return map made from an RR interval time series. Standard deviation along the diagonal (SD2) = 5 ms, standard deviation perpendicular to the diagonal (SD1) = 18 ms.

$$x_{i+1} = F(x_i)\,. \tag{2.29}$$

Because this expression binds the new value x_{i+1} of the variable to its predecessor value x_i, we can solve function F in principle, by pairing successive values of the time series, (x_i, x_{i+1}) for $i = 1$ to $N - 1$, and plotting them on a two-dimensional graph. This kind of a graph is called a return map.* If the dynamics of x is wholly determined by function F and there are enough data points, the method should reveal the shape of the function F.

If the dynamical system behind the time series is not one dimensional, suggesting that more than one variable has an impact on the system, a return map formed on the basis of a single measured variable naturally cannot solve functions F, G and so on. Even in such cases, a single variable return map may prove to be useful, although it is a certain type of projection of the multidimensional system into a single dimension. Figure 2.11 shows the return map of an RR interval time series. The points are typically scattered to form an ellipsoid but can also form complex structures. When the return map is an ellipsoid, it can be characterized by two quantities: the SD in the direction of the diagonal SD2 and the SD in a direction perpendicular to SD2, i.e., SD1† (Huikuri et al., 1996; Woo et al., 1994). These deviations are, by nature, measures of variability, since they quantify the movement of the system in a phase space. However, when the return map has a complex shape, above parameters do not describe the variability very well.

* This is also called Poincaré or Lorenz plot or map.

† For RR interval, SD1 is the same as RMSSD $\sqrt{2}$.

2.5.5 Other Approaches

2.5.5.1 Stationarity Test

An example of a simple stationarity test that mainly measures changes in the baseline signal is as follows: the signal is divided into segments of suitable length and for each segment, the signal average is calculated. When SD of these averages is divided by the SD of complete signal, we get a measure of stationarity (Palazzolo et al., 1998). This measure is small if the signal is stationary. The length of the segment should be chosen so that it is not too long in order for the local changes to be detectable, but it cannot be too short either in order to prevent the averages of the segments from varying too much. For an RR interval time series, a good choice is to use 20 beats/segment. If the measure is < 0.3, the signal can be considered to be reasonably stationary.

2.5.5.2 Symbolic Dynamics

The basic idea of this method is to characterize the original time series with a much simpler and coarser symbolic notation which, however, retains the essential dynamic characteristics of the original time series. This is done by converting the time series into a string of symbols. By this, we reduce the study of dynamics into handling of a symbol string. Naturally we lose most of the information contained in the original time series, but nevertheless, we retain key dynamical features, in a *coarse-grained form* (Voss et al., 1995, 1996; Palazzolo et al., 1998).

The conversion of a time series into a symbol string may be done using several methods. One of them is described in Figure 2.12. The signal is divided into two or more value ranges, depending on how many symbols we wish to utilize. Value ranges may be absolute bands or based on signal averages or SD. For example, if we have four different symbols, we may use the following bands:

A signal ≤ average − SD
B average − SD < signal ≤ average
C average < signal ≤ average + SD
D signal > average + SD

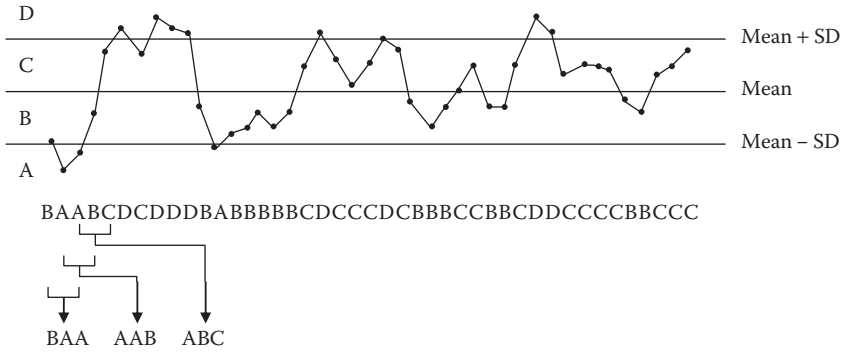

FIGURE 2.12
Conversion of the time series into a symbol row and grouping of the symbol row into words.

After the bands have been selected, the time series can be converted into a symbol string. The next step involves grouping the symbols in the string into *words*. A word is always formed by stepping forward one step in the symbol string. If we choose a word length of 3 and we have 4 different symbols, we get altogether $4 \times 4 \times 4 = 64$ different words. Each word corresponds to a specific graphical representation, which has at least a rough connection to the original dynamics. Different words do not have same probability because HRV dynamics favors certain words. The distribution of the words can thus be interpreted as a probability distribution. The shape of the distribution may itself act as a basis of further analysis, but it is also possible to measure the order related to the distribution in the terms of entropy. The simplest such measure is Shannon's entropy.

2.5.5.3 Multifractal Analysis

A monofractal signal can be described using just one scaling exponent α or Hurst exponent $h = \alpha - 1$, and in these situations, DFA, for example, is a valid method. If a signal is *multifractal*, a set or continuum of Hurst exponents is needed corresponding to a generalized fractal dimension D(*h*) (Ivanov et al., 1999; Amaral et al., 2001). Multifractal analysis is based on the wavelet transform (Equation 2.11) and the scaling properties of its local maximums (see computational details in Muzy et al., 1991). As a result, we get the fractal dimension as a function of the Hurst exponent.

A typical multifractal analysis of an RR interval time series of 25,000 data points of a healthy subject and a CHF subject is shown in Figure 2.13. In the healthy subject, $D(h)$ has a peak at $h = 0.1$ ($\alpha = 1.1$), and in the CHF subject, $D(h)$ has a peak at $h = 0.2$ ($\alpha = 1.1$). Thus, there is only a minor difference, and both cases resemble a $1/f$ type of dynamics (see Table 2.1). However, $D(h)$ of the CHF subject has a very narrow span in h, corresponding almost to monofractal behavior, but the healthy subject has very wide $D(h)$, indicating true multifractal dynamics. This result clearly reflects the common belief that the complexity of the system decreases as heart rate control degenerates.

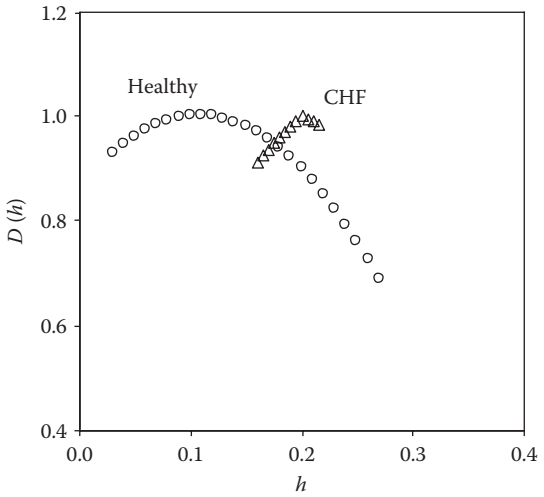

FIGURE 2.13
Fractal dimension $D(h)$ as a function of Hurst exponent h.

2.5.5.4 Stochastic Modeling

The unpredictable portion of the heart rate fluctuation can be due to chaotic dynamics, but there is an alternative explanation: the system includes a real stochastic component. The basic idea behind stochastic modeling is that the unpredictable component is not a perturbation but an essential part of the dynamical behavior of the system. The source of this true noise can be physiological, or it may be a reflection of external disturbances. If the system is truly stochastic, it cannot be described by a deterministic model; rather a stochastic one is needed. Many stochastic systems can be described by the *Langevin* equation

$$\frac{dX(t)}{dt} = g(X(t)) + h(X(t))\Gamma(t), \tag{2.30}$$

where $X(t)$ is the state of the system at moment t, and functions g and h represent the deterministic and stochastic parts of the time evolution. $\Gamma(t)$ is the uncorrelated white noise with zero mean and Gaussian distribution. It has been shown that the *difference version* of Equation 2.30,

$$X(t+\tau) = X(t) + g(X(t)) + h(X(t))\Gamma(t), \tag{2.31}$$

where τ is a finite delay parameter (2–20 min), can model the long-term RR interval time series (Kuusela, 2004; Kuusela et al., 2003). The control functions g and h can be extracted from the data by computing conditional probability distributions. In practice, 50,000–100,000 data points, that is, 12–24 h of RR intervals, are needed in order to determine these distributions reliably. Results from an analysis of 24 h recordings of RR interval are shown in Figure 2.14. At small (large) RR interval values, the deterministic part g is clearly positive (negative); this guarantees that the RR interval is kept within certain limits.* In the case of the healthy subject, typical g function has three zero-crossings (two stable fixed points and one unstable fixed point†) in the middle, but with the CHF subject, g function is flat, indicating a less complex dynamical control. The stochastic parts h have a complicated structure in both cases, but the mean level of the h function is smaller in the CHF subject, which can be interpreted as a lower stochasticity of the system.

2.6 Conclusions

Heart rate fluctuations can be analyzed using many different methods and approaches. No single method described here is clearly superior to other techniques, and therefore, it

* When the RR interval is small, the g function and also the derivative in Equation 2.30 are positive; thus, the RR interval increases. Similarly, when the RR interval is large, the g function and the derivative are negative; thus, the RR interval decreases.

† If we omit the stochastic part in Equation 2.30, the zeros of the g function correspond to the fixed points of the system. At the fixed points, the derivative is zero and the system has no tendency to change. Fixed points can be either stable and attract all nearby states or unstable and repel nearby states.

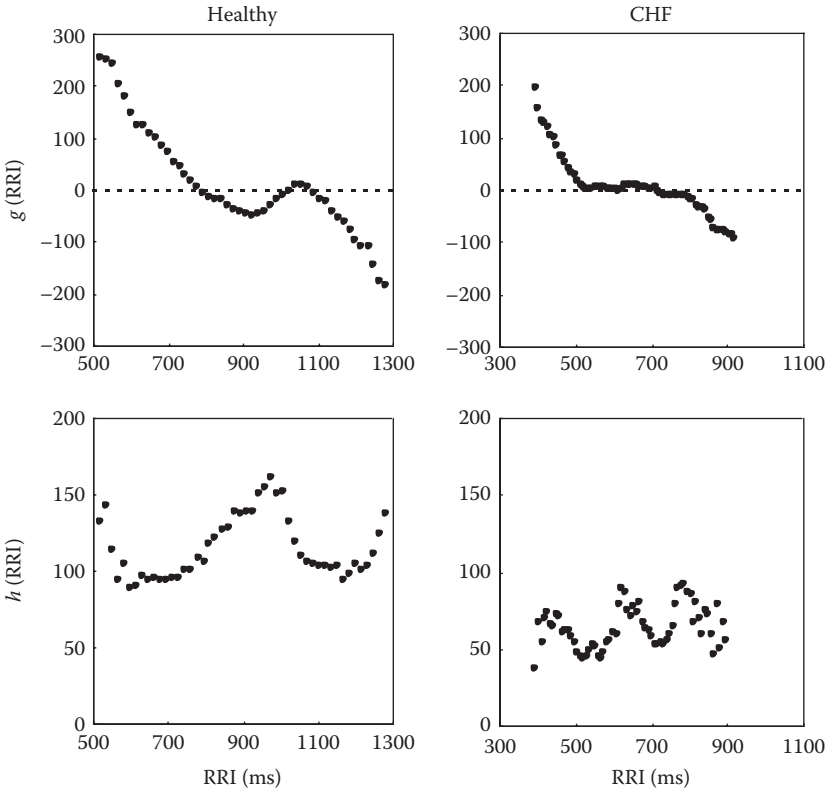

FIGURE 2.14
The deterministic control function *g* and the stochastic control function *h* for healthy and CHF subjects.

is recommended that researchers use several techniques in combination. The physiological interpretation of results is often difficult, especially in the case of non-linear methods and further investigation is needed. However, the time series analysis of RR interval data has proven to be useful and it has already gained significant clinical relevance. It would be very useful to agree on common rules or guidelines about when to use each method, the length of the time series needed, normative values and the kind of editing or filtering operations, so that the results from different laboratories can be compared.

Abbreviations

AF	Atrial fibrillation
AIC	Akaike information criteria
ApEn	Approximate entropy
AR	Autoregressive
CD	Correlation dimension
CDi	Pointwise correlation dimension
CDM	Complex demodulation

CHF	Congestive heart failure
CWT	Continuous wavelet transform
DFA	Detrended fluctuation analysis
FFT	Fast fourier transform
FPE	Final prediction error
HF	High frequency
HRV	Heart rate variability
LF	Low frequency
MDL	Minimum description length
MSE	Multiscale entropy
NN50	Number of pairs of adjacent RR intervals differing by more than 50 ms
pNN50	Ratio of NN50 count to the count of all RR intervals expressed as a percentage
PSD	Power spectral density
RMSSD	Root mean square of successive differences of RR intervals
SampEn	Sample entropy
SD	Standard deviation
SDA	Selective discrete Fourier transform algorithm
SDANN	Normal-to-normal standard deviation of all 5 minute segments averaged over 24 hours
SDNN	Normal-to-normal standard deviation
STFT	Short-time Fourier transform
TP	Total power
ULF	Ultralow frequency
VLF	Very low frequency
WFT	Windowed fourier transform
WVD	Wigner–Ville distribution

References

Akaike H. (1969) Power spectrum estimation through autoregressive model fitting. *Ana. Inst. Stat. Math.* 21:407–419.

Amaral LAN, Ivanov Ch, Aoyagi N, Hidaka I, Tomono S, Goldberger AL, Stanley HE, Yamamoto Y. (2001) Behavioral-independent features of complex heartbeat dynamics. *Phys. Rev. Lett*. 86:6026–6029.

Bassingthwaighte JB, Raymond GM. (1995) Evaluation of the dispersional analysis method for fractal time series. *Ann. Biomed. Eng*. 23:491–505.

Bettermann H, van Leeuwen P. (1998) Evidence of phase transitions in heart period dynamics. *Biol. Cybern.* 78:63–70.

Bianchi AM, Mainardi L, Petruci E, Signorini MG, Mainardi M, Cerutti S. (1997) Time-variant power spectrum analysis for the detection of transient episodes in HRV signal. *IEEE Trans. Biomed. Eng.* 40:135–144.

Bigger JT, Steinman RC, Rolnitzky LM, Fleiss JL, Albrecht P, Cohen RJ. (1996) Power law behaviour of RR-interval variability on healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. *Circulation.* 93:2142–2151.

Claasen TACM, Mecklenbräuker WFG. (1980a) The Wigner distribution—A tool for time-frequency signal analysis. Part I: Continuous-time signals. *Philips J. Res.* 35:217–250.
Claasen TACM, Mecklenbräuker WFG. (1980b) The Wigner distribution a tool for time-frequency signal analysis. Part II: Discrete-time signals. *Philips J. Res.* 35:276–300.
Costa M, Goldberger AL, Peng C-K. (2002) Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett*. 89:068102 1–4.
Costa M, Goldberger AL, Peng C-K. (2005) Multiscale entropy analysis of biological signals. *Phys. Rev*. E71:021906 1–18.
Cysarz D, Bettermann H, Van Leeuwen P. (2000) Entropies of short binary sequences in heart period dynamics. *Am. J. Physiol.* 278(6):H2163–H2172.
Farmer JD, Ott E, Yorke JA. (1983) Dimension of chaotic attractors. *Physica.* D7:153–180.
Figliola A, Serrano E. (1997) Analysis of physiological time sereis using wavelet transforms. *IEEE Eng. Med. Biol.* 16:74–79.
Grassberger P, Procaccia I. (1983) Characterization of strange attractors. *Phys. Rev. Lett.* 31:346–349.
Huikuri HV, Seppänen T, Koistinen MJ, Airaksinen KEJ, Ikäheimo MJ, Castellanos A, Myerburg RJ. (1996) Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction. *Circulation.* 93:1836–1844.
Hyung-Ill C, Williams WJ. (1989) Improved time-frequency representation of multicomponent signals using exponential kernels. *IEEE Trans. Acoust. Speech. Signal Process.* 37:862–871.
Ivanov Ch, Amaral LAN, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE. (1999) Multifractality in human heartbeat dynamics. *Nature.* 399:461–465.
Iyengar N, Peng C-K, Morin R, Goldberger AL, Lipsitz LA. (1996) Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 271:R1078–R1084.
Kantz H, Schreiber T. (1995) Dimension estimates and physiological data. *Chaos.* 5:143–154.
Kay SM, Marple SL. (1981) Spectrum analysis: A modern perspective. *Proc. IEEE.* 69:1380–1418.
Keselbrener L, Akselrod S. (1996) Selective discrete Fourier transform algorithm for time–frequency analysis: Method and application on simulated and cardiovascular signals. *IEEE Trans. Biomed. Eng*. 43:789–802.
Kim SY, Euler DE. (1997) Baroreflex sensitivity assessed by complex demodulation of cardiovascular variability. *Hypertension*. 29:1119–1125.
Kuusela TA. (2004) Stochastic heart-rate model can reveal pathologic cardiac dynamics. *Phys. Rev.* E69:031916 1–7.
Kuusela TA, Shepherd T, Hietarinta J. (2003) Stochastic model for heart rate fluctuations. *Phys. Rev*. E67:061904 1–7.
Marple SL. (1987) *Digital Spectral Analysis with Applications*. Englewood Cliffs, NJ, Prentice-Hall.
Mayer-Kress G, Yates FE, Benton L, Keidel M, Tirsch W, Pöppl SJ, Geist K. (1988) Dimensional analysis of nonlinear oscillations in brain, heart, and muscle. *Math. Biosci.* 90:155–182.
Muzy JF, Bacry E, Arneodo A. (1991) Wavelets and multifractal formalism for singular signals: Application to turbulence data. *Phys. Rev. Lett*. 67:3515–3518.
Novak P, Novak V. (1993) Time/frequency mapping of the heart rate, blood pressure and respiratory signals. *Med. Biol. Eng. Comput*. 31:103–110.
Palazzolo JA, Estafanous FG, Murray PA. (1998) Entropy measures of heart rate variation in conscious dogs. *Am. J. Physiol. Heart Circ. Physiol.* 274:H1099–H1105.
Partzen E. (1974) Some recent advances in time series modeling. *IEEE Trans. Automat. Contr.* 19:723–730.
Peng C-K, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL. (1993) Long-range anticorrelations and non-gaussian behaviour of the heartbeat. *Phys. Rev. Lett.* 70:1343–1346.
Peng C-K, Havlin S, Stanley HE, Goldberger AL. (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos.* 5:82–87.
Pincus SM, Goldberger AL. (1994) Physiological time-series analysis: What does regularity quantify? *Am. J. Physiol. Heart Circ. Physiol.* 266:H1643–H1656.

Pincus S. (1995) Approximate entropy (ApEn) as a complexity measure. *Chaos.* 5:110–117.
Richman JS, Moorman JR. (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278:H2039–H2049.
Rissanen J. (1983) A universal prior for the integers and estimation by minimum description length. *Ann. Stat*. 11:417–431.
Voss A, Kurths J, Kleiner HJ, Witt, A, Wessel N. (1995) Improved analysis of heart rate variability by methods of nonlinear dynamics. *J. Electrocardiol.* 28:81–88.
Voss A, Kurths J, Kleiner HJ, Witt A, Wessel N, Saparin P, Osterziel KJ, Schurath R, Dietz R. (1996) The application of methods on non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. *Cardiovasc. Res.* 31:419–433.
Wiklund U, Akay M, Niklasson U. (1997) Short-term analysis of heart-rate variability by adapted wavelet transforms. *IEEE Eng. Med. Biol.* 16:113–118.
Woo MA, Stevenson WG, Moser DK, Middlekauf HR. (1994) Complex heart rate variability and serum norepinephrine levels in patients with advanced heart failure. *J. Am. Coll. Cardiol.* 23:565–569.
Yum MK, Kim NS, Oh JW, Kim CR, Lee JW, Kim SK, Noh CI, Choi JY, Yun YS. (1999) Non-linear cardiac dynamics and morning dip: An unsound circadian rhythm. *Clin. Physiol.* 19:56–67.