Acta Genet Med Gemellol 36:51-59 (1987)

Spectral Analysis of Twin Time Series Designs

P.C.M. Molenaar[1] **D.I. Boomsma**[2]

[1] *Department of Psychology, University of Amsterdam, and* [2] *Department of Experimental Psychology, Free University, Amsterdam, Netherlands*

Abstract. The genetic analysis of physiological time series has to accommodate the presence of autocorrelation. This can be accomplished by means of orthogonal transformation of the series, thus enabling the use of standard genetic analysis techniques for the sequence of uncorrelated transforms. In view of the oscillatory character which typifies various physiological time series, it is customary to invoke spectral techniques for the analysis of these series. It can be shown that spectral analysis is an orthogonal transformation that asymptotically resembles principal component analysis. Consequently, standard genetic analysis methods for the uncorrelated spectral transforms may be used. This approach will be illustrated with simulated and real (heart rate) data for univariate twin time series. Furthermore, it will be indicated that the proposed analysis can be readily generalized to multivariate time series.

Key words: Spectral analysis, Twin data, Heart rate, Time series

INTRODUCTION

The study of heart rate in cardiovascular psychophysiology is guided by the assumption that heart rate not only reflects adaptation to metabolic demands, but also central nervous system functioning. Especially with the introduction of advanced measurement equipment the study of heart rate shifted from a crude measure of arousal to a sensitive index of cognitive processing, mental stress or attentional load. Within a psychophysiological context, there are two approaches in studying heart rate: 1) description of directional changes in heart rate level, and 2) description of heart rate variability.

Variability of heart rate is influenced by several factors, the most important ones being respiration, changes in blood pressure and changes in the thermoregulatory system.

In addition, heart rate variability is also influenced by mental and physical exercise.

Heart rate variability may be quantified using spectral analysis. In spectral analysis heart rate variability is expressed as a function of underlying frequency, providing information on the sources of variability. That is, the spectrum or power density function that results from spectral analysis shows the contribution of oscillations within a range of frequencies to the total variance of the signal, thus comparable to the partioning of variance in ANOVA. Spectral analysis of a sequence of interbeat intervals (ie, time between successive R waves in the electrocardiogram) can reveal significant peaks in the spectrum that might reflect the presence of important oscillatory processes underlying the sequence. For example, the spectrum of a series of interbeat intervals during rest usually shows peaks that correspond to the dominant respiratory frequency of the subject (around 0.25 Hz) and to baroreceptor resonance (around 0.10 Hz). A peak at still lower frequencies (< 0.05 Hz) is attributed to the thermoregulatory system. The amplitude of the respiration band often is used as an index of vagal control on the heart.

Heart rate variability has been used to study stress- and task-related fluctuations, for example it has been found that heart rate variability decreases during mental load. Heart rate variability spectra also have been used in patho-physiological studies of the cardiovascular system, for example in diabetics.

METHOD

Application of spectral analysis to a sequence of interbeat intervals requires the sequence to be weakly stationary. This means that the sequence has to have both a constant mean function as well as an autocorrelation function that is invariant along the time axis. Stationarity of the autocorrelation function automatically implies that the power density function or spectrum also is stationary, as there exists a one-to-one relationship between autocorrelation and power density. Biological signals, however, can exhibit complex time-dependent variation, yielding an autocorrelation, and thus a spectrum, that varies with time. However, if the state of a subject remains relatively constant, such as during a resting condition, the observed signal will show stationary behavior, that is the autocorrelation will be stable.

Stationary levels of autocorrelation (spectra) obtained within the same experimental condition or state differ substantially from one person to another. Apparently, the functional systems underlying biological signals are not constant across subjects and consequently this intersubjective variation may reflect genetic influences. In order to assess these influences, a genetic analysis is called for in which the autocorrelation of biological signals figures predominantly. On the one hand, this autocorrelation offers a description of the functional dynamics of the underlying biological system and as such is the proper measure for our genetic analysis. On the other hand, the typical autocorrelation structure of biological signals requires the use of special techniques. We therefore elaborated an earlier approach to the genetic analysis of biological signals [3], constituting of the following steps:

1) Each original autocorrelated signal is transformed into a sequence of uncorrelated variables. The transformation is accomplished by means of time-dependent base functions or eigenvectors which describe the dynamics of the original signal. The transformation

results in a sequence of uncorrelated variables that can be conceived of as amplitudes of the associated base functions in a given signal. The square of these amplitudes gives the amount of variance explained by the corresponding base function.

2) For each given base function the intersubjective variation can be analyzed by means of standard genetics techniques. In the case of twins, maximum likelihood estimates of the proportions of genetic and environmental variance may be obtained from the mean squares of amplitudes between and within MZ and DZ twin pairs.

3) Estimates of the proportion of genetic and environmental variance of the original signal can be obtained by means of inverse transformation in which the results obtained in the second step are averaged across base functions.

SPECTRAL ANALYSIS OF GENETIC INFLUENCES

Consider the following basic genetic model for an observed biological signal y(t):

$$y(t) = G(t) + E(t) \qquad t = 0,1, ..., n$$

where G(t) and E(t) are latent time series of genetic and non-shared environmental influences and are mutually uncorrelated. In general, the latent time series G(t) and E(t), and consequently the observed series y(t), will be autocorrelated. According to the first step in our approach to the genetic analysis of biological signals, then, the observed series y(t) is transformed into a sequence of uncorrelated variables:

$$y^*(k) = T_{\phi_k}[y(t)] \qquad k = 0,1, ..., n$$

where T denotes a transformation by means of base functions $\phi_k(t)$ such that

$$\text{cor}\,(y^*(k), y^*(k')) = 0, \qquad \text{if } k \neq k'$$

In an earlier publication [3], the transformation T was determined by means of principal component analysis. As biological signals typically are observed across several hundreds of consecutive time points, however, transformation by means of principal component analysis is no longer computationally feasible.

Fortunately, it can be shown [4] that principal component analysis of sufficiently long, stationary time series converges to spectral analysis in which

$$\phi_k(t) = \exp[-i2\pi\omega_k t] = \cos[2\pi\omega_k t] - i \sin[2\pi\omega_k t] \qquad i = \sqrt{-1}$$

Accordingly, the base functions or eigenvectors $\phi_k(t)$ as obtained from a principal component analysis of the dynamics of y(t) converge to the complex-valued exponentials that span up the frequency domain. Hence, with stationary biological signals a spectral analysis is formally equivalent to a transformation by means of principal component analysis.

We are now in a position in which the transformation of y(t) into a sequence of uncorrelated variables can be carried out efficiently because the required base functions are analytically given. The transformation in question is known as the discrete Fourier transform:

https://doi.org/10.1017/S000156600000458X Published online by Cambridge University Press

$$y^*(\omega_k) = \sum_t y(t) \exp[-i2\pi\omega_k t] \qquad k = 0,1, ..., n$$

where

$$S_y(\omega_k) = \text{var}[y^*(\omega_k)]$$

is the power density at frequency ω_k, ie, that part of the variance explained by oscillations with frequency ω_k.

As cor $(y^*(\omega_k), y^*(\omega_{k'})) = 0$ if $k \neq k'$, the intersubjective variation at each frequency ω_k can be analyzed by means of standard genetics techniques. Notice, however, that $y^*(\omega_k)$ is complex-valued. Therefore special care has to be taken in order to arrive at valid estimates and their sampling distribution (see [1] for an extensive overview of these issues). Accordingly, the observed power density at each frequency ω_k is decomposed as:

$$S_y(\omega_k) = S_G(\omega_k) + S_E(\omega_k)$$

where S_G and S_E denote the power densities of G(t) and E(t). In addition, the proportion $h^2(\omega_k)$ of genetic variance at each frequency ω_k can be determined in a straightforward manner. In the final, third step of our approach, the proportion of genetic variance of the original signal y(t) is obtained by inverse discrete Fourier transformation:

$$h^2 = 1/N \sum_u h^2(\omega_k) \exp[i2\pi\omega_k u]$$

at $u = 0$, where u is the lag between time t and time t'.

APPLICATIONS TO SIMULATED DATA

We will present a few illustrative applications of the proposed analysis to simulated data in order to show the reliability of our approach. A computer program has been written which generates simulated data according to the basic genetic model for time series and which carries out the three steps in our spectral analysis of genetic influences. A detailed description of the simulation program is given in [3].

Time series y(t) were generated for 10 MZ and 10 DZ twin pairs. Each series consisted of 512 time points. Three different time series were simulated for each twin pair, in which, respectively:

1. The environmental series E(t) is autocorrelated, whereas the genetic series G(T) is not.
2. G(t) is autocorrelated, but E(t) is not autocorrelated, and
3. Both G(t) and E(t) are autocorrelated.

In case a time series is not autocorrelated, it is called a white noise series. For example, in the first data set G(t) is white noise. Each data set was generated in such a way that the proportion of genetic variance was 0.5.

https://doi.org/10.1017/S000156600000458X Published online by Cambridge University Press

RESULTS

Henceforth, a power density function will be called a spectrum. Fig. 1A, then, shows the observed spectrum for the first data set in which G(t) is white noise and E(t) is autocorrelated. In this plot, the abscissa denotes frequencies: from zero up to 0.5 cycles per second. The ordinate denotes power density, that is the spectral value S at each frequency. Figs. 1B and 1C show the main results of our technique, namely the decomposition of the total spectrum $S_y(\omega_k)$ into the underlying genetic and environmental spectra $S_G(\omega_k)$ and $S_E(\omega_k)$. In these figures, the continuous lines represent the true spectra used in the simulation, whereas the broken lines represent estimates of these spectra. Notice that with this data set the genetic spectrum $S_G(\omega_k)$ has the typical flat appearance of a white noise spectrum. In contrast, the environmental spectrum $S_E(\omega_k)$ varies considerably across frequencies, as is typical of autocorrelated series. Estimated heritability was 0.47, which is close to the true value of 0.5.

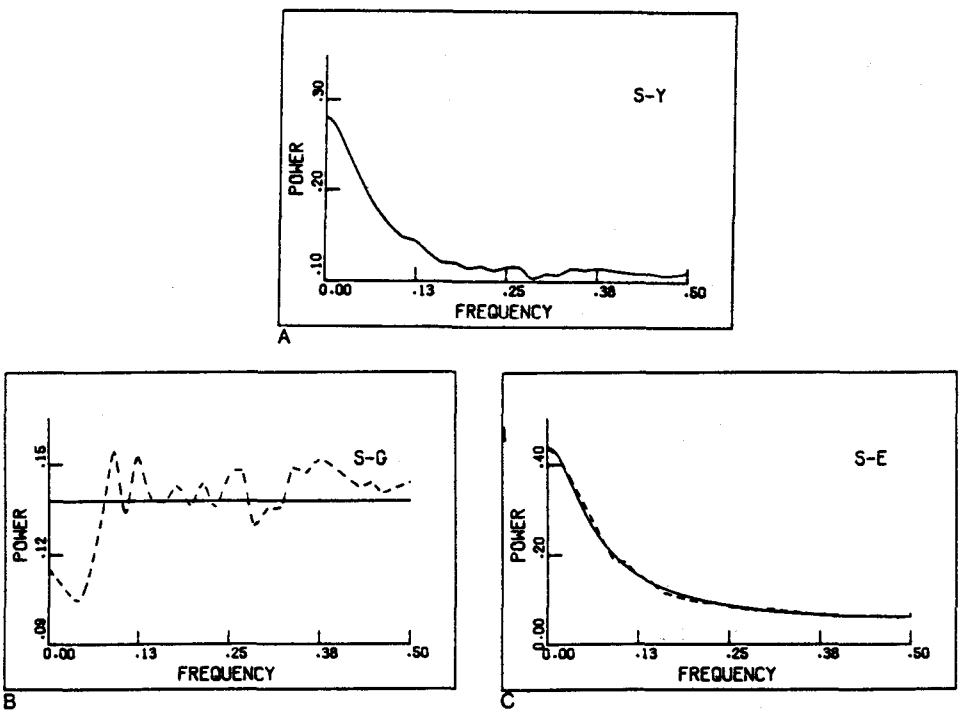

Fig. 1. **A:** Total observed spectrum for simulated data y(t) = G(t) + E(t) where G(t) is white noise and E(t) is autocorrelated. **B:** Estimated (broken line) and true (solid line) genetic spectrum. **C:** Estimated and true environmental spectrum.
Note that ordinates for the 3 spectra are differently scaled.

In the second data set G(t) is autocorrelated whereas E(t) is white noise. Fig. 2A shows the observed spectrum obtained with this data set. Figs. 2B and 2C depict the resulting decomposition of $S_y(\omega_k)$ into the underlying genetic and environmental spectra. Estimated heritability was 0.52, which is again close to the true value of 0.5.

https://doi.org/10.1017/S000156600000458X Published online by Cambridge University Press

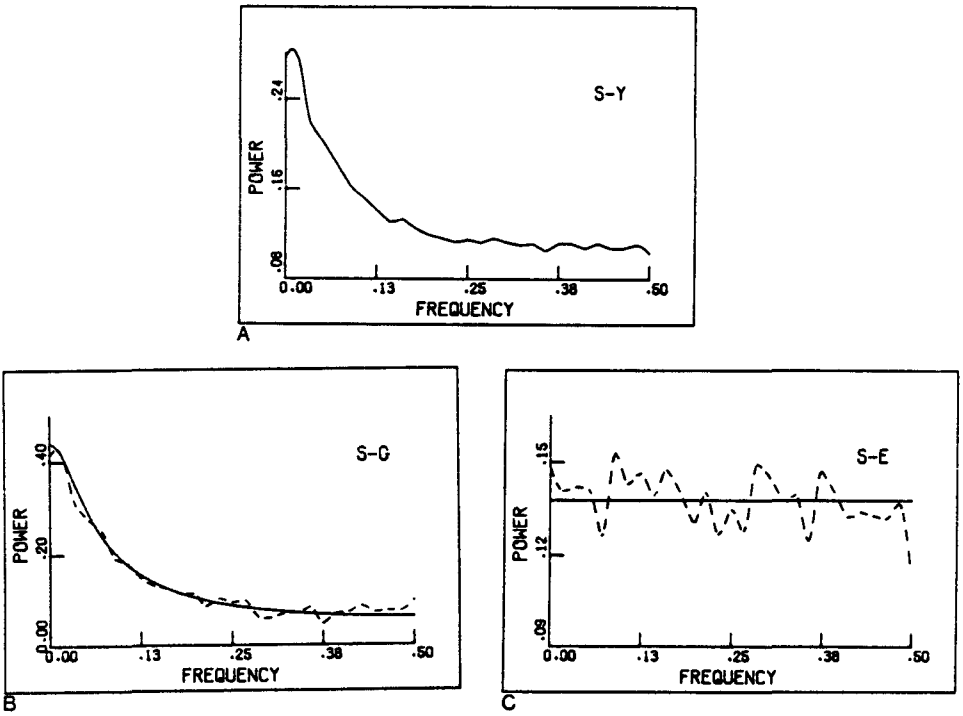

Fig. 2. **A:** Total observed spectrum for simulated data $y(t) = G(t) + E(t)$ where $G(t)$ is autocorrelated and $E(t)$ is white noise. **B:** Estimated and true genetic spectrum. **C:** Estimated and true environmental spectrum.

In the third and final data set both $G(t)$ and $E(t)$ are autocorrelated. Fig. 3A shows the obtained spectrum $S_y(\omega_k)$ while Figs. 3B and 3C give the underlying genetic and environmental spectra. Estimated heritability was 0.51.

These figures show that, even with a small sample of 10 MZ and 10 DZ twin pairs, the genetic and non-shared environmental spectra underlying biological signals can be reliably estimated.

APPLICATION TO HEART RATE DATA

We will now illustrate the proposed spectral analysis by an application to heart rate data. For 10 MZ and 10 DZ twin pairs between the age of 14 and 19, ECG was recorded during an 8.5 minutes period of rest. Subjects were seated in a comfortable chair and were instructed to relax as much as possible. ECG was measured by Ag-AgCl electrodes placed on sternum and lateral margin of the chest. The ECG signal was digitized at 250 samples/second via 12 bit A-D converter and stored on disk. These data were then used to determine successive R-R intervals (interbeat intervals). Mean interbeat was 918 ms (SD = 18.8). Differences between subjects in mean interbeat interval of course correspond to differences in the number of beats during each 8.5 minute period. No corrections on the interbeat interval sequence were carried out before invoking the spectral analysis.

https://doi.org/10.1017/S000156600000458X Published online by Cambridge University Press

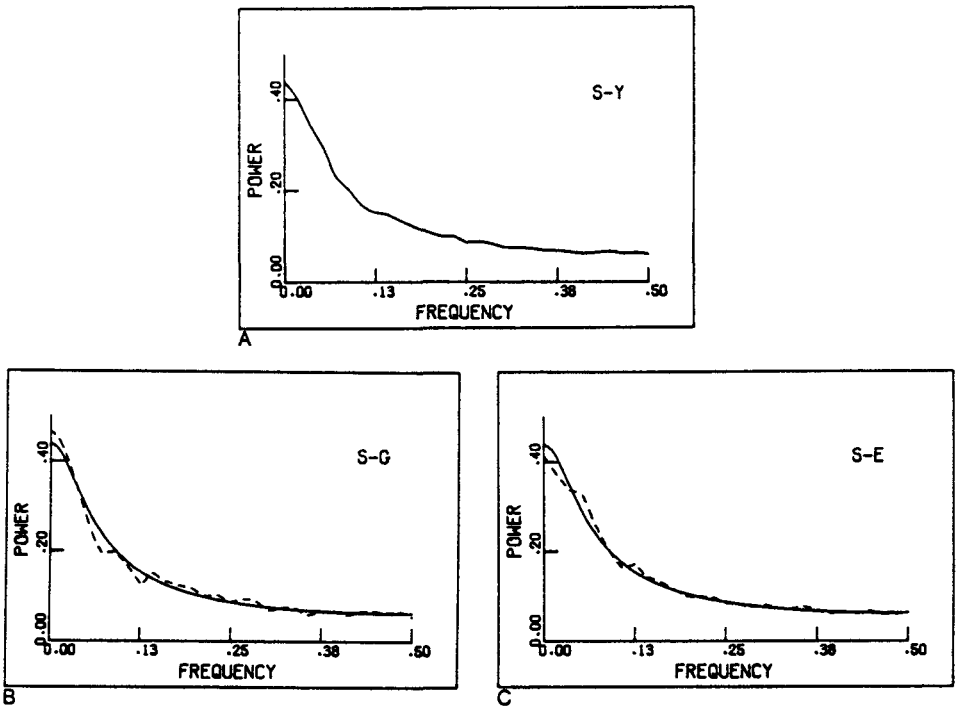

Fig. 3. **A:** Total observed spectrum for simulated data $y(t) = G(t) + E(t)$ where both $G(t)$ and $E(t)$ are autocorrelated. **B:** Estimated and true genetic spectrum. **C:** Estimated and true environmental spectrum.

Fig. 4A shows the observed spectrum of interbeat intervals. Notice the slight peak at 0.10 cycles per second (blood pressure band). Figs. 4B and 4C show the decomposition into underlying genetic and environmental spectra. Each spectrum was standardized in order to show more clearly the morphological features. Estimated heritability was 0.07.

Perhaps more revealing is Fig. 4D, showing the frequency-dependent heritability estimates. There is some interesting variation of heritability across frequencies. For instance, the highest frequency-dependent heritabilities are obtained with frequencies of about 0.35 cycles per second. Such high-frequency oscillations above 0.25 cycles per second may reflect vagal influences on the heart [2]. In contrast, heritabilities at frequencies of about 0.10 cycles per second appear to be close to zero. These 0.10 cycles per second oscillations are supposed to be coupled to the blood pressure system. Finally, oscillations with frequencies close to zero, which are supposed to be coupled to the thermo-regulatory system, show medium heritabilities.

CONCLUSIONS

Our applications to simulated data indicate that a spectral analysis of genetic influence underlying interindividual variability of biological signals can yield reliable results with

https://doi.org/10.1017/S000156600000458X Published online by Cambridge University Press

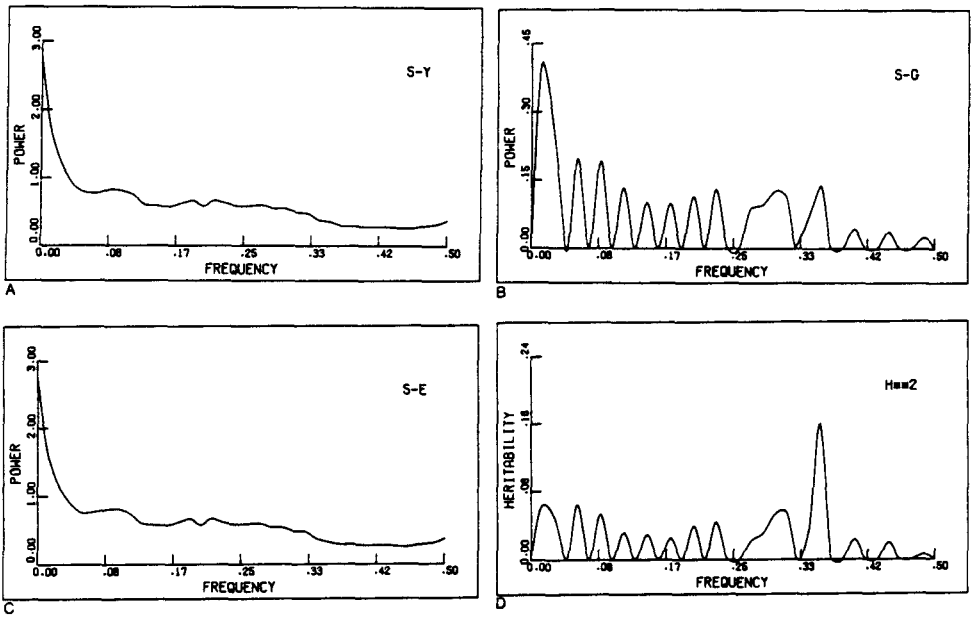

Fig. 4. A: Total observed spectrum of interbeat intervals. B: Estimated genetic spectrum. C: Estimated environmental spectrum. D: Frequency-dependent heritability estimates.

only a few subjects. This is partly due to the availability of numerous repeated measures making up an observed biological signal. As to this, the proposed method is the only method which can easily handle a practically unlimited amount of repeated measures, where at each frequency various suitable genetic models can be fitted by means of maximum likelihood techniques. Furthermore, the statistical theory of spectral estimates thus obtained is well developed, yielding practical guides at each level of the analysis. Hence, the proposed method would seem quite eligible for the genetic analysis of biological signals.

In the final step of our approach, the proportion of genetic variance of the original signal is obtained by inverse discrete Fourier transformation of $h^2(\omega_k)$. In effect, this inverse transformation amounts to averaging across frequency-dependent heritabilities, and therefore may not yield very informative results in case $h^2(\omega_k)$ has substantial frequency-dependent variation. It then will be more informative to consider a frequency-dependent plot of $h^2(\omega_k)$, like Fig. 4D. Such a plot might be called a heritability spectrum, showing the proportions of frequency-dependent genetic variance of a biological signal.

The proposed spectral method can be generalized in order to analyze genetic influences underlying multidimensional biological signals. Multidimensional biological signals are obtained in the study of, for instance, respiratory influences on the heart beat or neural sources of electroencephalografic (EEG) activity at different regions of the head. The genetic analysis of these multidimensional signals is based upon Hermitian matrices of auto- and cross-spectra, where at each frequency a complex-valued factor analysis can be carried out in order to determine common and specific genetic influences. Again, the spectral method is the only method which can easily handle a multidimensional genetic

analysis of biological signals and therefore is of central importance in this field of application.

Acknowledgement. Zygosity determinations were made by the Central Laboratory of the Blood Transfusion Service, Netherlands Red Cross, Amsterdam.

REFERENCES

1. Brillinger DR (1973): The analysis of time series collected in an experimental design. In Krishnaiah PR (ed): Multivariate analysis III. New York: Academic Press, pp 241-256.
2. DeBoer RW, Karemaker JM (1985): The phase between respiration and heart-rate variability. Proceedings 38th Annual Conference on Engineering in Medicine and Biology.
3. Molenaar PCM, Boomsma DI (1987): The genetic analysis of repeated measures II: The Karhunen-Loeve expansion. Behav Genet (in press).
4. Molenaar PCM, Molen MW van der (1985): Global models: A viable compromis between content specificity and ease of application to heart rate changes. In Orlebeke JF, Mulder G, Doornen LJP (eds): Psychophysiology of cardiovascular control. Models, methods and data. New York: Plenum press, pp 375-390.

Correspondence: Dr. Peter C.M. Molenaar, Department of Psychology, University of Amsterdam, Weesperplein 8, 1018 XA Amsterdam, Netherlands.