The Canadian Journal of Statistics
Vol. 21, No. 1, 1993, Pages 1–19
La Revue Canadienne de Statistique

The digital rainbow: Some history and applications of numerical spectrum analysis*

David R. BRILLINGER

The University of California, Berkeley

Key words and phrases: Fourier analysis, inverse problem, nuclear magnetic resonance, periodogram, seismology, spectrum, time series.
AMS 1985 subject classifications: Primary 01A99,42-03, 60G12, 62M15.

ABSTRACT

Statistical concepts and techniques are basic to scientific investigation. One concept that enjoys both a theoretical and a physical existence is the spectrum. A spectrum may be described as a display of the intensity or variability of a phenomenon versus period or frequency. Spectra are particularly useful in the study of systems subject to resonance, but have many other uses. This paper begins with some of the historical development of the field, describing a sequence of contributions by Michelson, Schuster, Einstein, Fisher, Bartlett, Tukey, and Whittle. The paper next presents collaborative applications to the study of the free oscillations of the earth, to the dispersion of seismic surface waves and to nuclear-magnetic-resonance spectroscopy. Finally, there is mention of open problems and opinions on future directions.

RÉSUMÉ

Les concepts et les techniques statistiques sont à la base de toute étude scientifique. Le concept de spectre existe tant dans un cadre théorique que dans un cadre physique. Un spectre peut être décrit comme étant la représentation de l'intensité ou de la variabilité d'un phénomème en fonction de la période ou de la fréquence. Les spectres sont particulièrement utiles lors de l'étude de systèmes soumis à une résonance, mais ont également bien d'autres emplois. Cet article débute par un historique des développements dans le domaine, décrivant les contributions de Michelson, Schuster, Einstein, Fisher, Bartlett, Tukey et Whittle. Cet historique est suivi d'applications des spectres à l'étude des oscillations libres de la Terre, à la dispersion des ondes sismiques de surface et à la spectroscopie de résonance magnétique nucléaire. Enfin, questions et opinions pour des recherches futures sont mentionnées.

1. INTRODUCTION

Gerhard Herzberg's research field is the analysis of the spectra of molecules in order to determine their structure. He did this experimentally by passing light through a prism, as Newton had so many years earlier. On the other hand, statisticians have been concerned primarily with spectra as theoretical parameters. It is noteworthy for a concept to have these distinct existences.

To begin, this paper presents a historical development. A recent highlight of the history of spectra is a just-noticed paper by Albert Einstein. This paper, written in 1914, laid out

*This paper is based on the Herzberg Lecture presented at Carleton University 14 February 1992. The lecture is meant as a vehicle for scientists speaking on topics of general interest to academics and to the public at large. This research was supported in part by the National Science Foundation Grant DMS-8900613.

a practical definition of the power spectrum and a corresponding estimation procedure; see Yaglom (1987a, b). An estimate of the spectrum of sunspot numbers, computed as Einstein might have, is provided here. The work of A.A. Michelson, A. Schuster, R.A. Fisher, M.S. Bartlett, J.W. Tukey, and P. Whittle is highlighted. The second part of the paper provides applications of spectral techniques to the phenomena of free oscillations of the earth, of seismic surface waves, and of nuclear-magnetic resonance spectroscopy. The applications each show a progression from a spectrum analysis to a conceptual-model-based analysis.

By presenting a combination of history and examples it is hoped to interest other statisticians in the topic and to display how statistical concepts can interact with scientific ones.

2. HISTORICAL DEVELOPMENT

Spectrum analysis has its roots in physical science. In 1666, when Newton employed a prism to cast the rainbow on the wall, he coined the term spectrum and began the formal study of the subject. His work was part quantitative, with the counting of the number of colors and the measurement of the widths of the bands; see Topper (1990). In 1801 J.F.W. Herschel measured the temperature at various positions along the image; see Sobel (1989). Herschel's work is more in line with the idea of the spectrum as measuring intensity. A related important step was provided by Gouy (1886), who proposed the representation of light by a Fourier expansion.

2.1. Early Numerical Work.

Michelson (1892) was concerned with finding a length standard. To this end he caused particular substances to emit light. Then, via a mirror, he superposed that light on itself with a time delay. When the superposed light was viewed appropriately, fringes could be seen. The phenomenon is referred to as interference and is understood if one views the original signal as $\cos \lambda t$ and hence after superposition as $\cos \lambda t + \cos \lambda(t+u)$, t denoting time, u delay and λ frequency. As Michelson changed u the clearness of the fringes varied and was recorded. This gave a function $V(u)$, the visibility curve. A variety of examples of the visibility curves he found are presented in Michelson (1892, 1902).

Supposing $f(\lambda) = g(\lambda - \lambda_0)$ to denote the spectrum (for the moment undefined) of the light source at frequency λ to be narrow and centered at λ_0, Michelson (1891) argued that the visibility curve is given by

$$V(u) = \sqrt{\{\int \cos u\lambda \, g(\lambda)\, d\lambda\}^2 + \{\int \sin u\lambda \, g(\lambda)\, d\lambda\}^2}. \tag{1}$$

Michelson called the problem of determining $g(\cdot)$ from $V(\cdot)$ an *inverse problem* and obtained answers by guessing. One of his most important inferences was that the red hydrogen line was a doublet. This inference of splitting led ultimately to important developments in quantum mechanics. Rayleigh (1892) however pointed out that the inverse problem of (1) did not necessarily have a unique solution.

2.2. The Periodogram.

By 1898 Michelson had developed a "harmonic analyser" (Michelson and Stratton 1898). In Michelson (1913) it was employed to compute Schuster's periodogram for

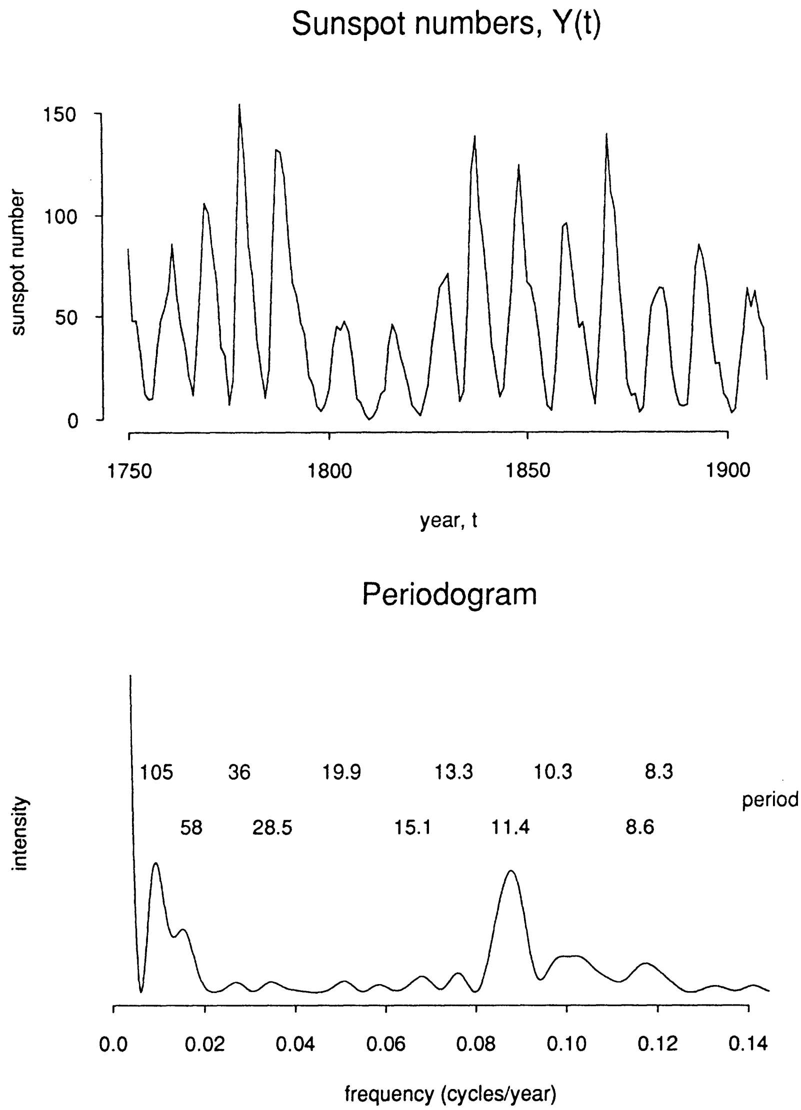


FIGURE 1: The top graph shows monthly relative sunspot numbers for 1750 to 1910. The bottom graph is the periodogram computed for those data. The periods listed are those that Michelson (1913) indicated.

some sunspot data studied by Kimura (1913). For data $Y(t)$, $t = 0, \ldots, T-1$, and period τ, the Schuster periodogram is given by

$$P^T(\tau) = \left(\sum_{t=0}^{T-1} Y(t) \cos \frac{2\pi t}{\tau}\right)^2 + \left(\sum_{t=0}^{T-1} Y(t) \sin \frac{2\pi t}{\tau}\right)^2. \tag{2}$$

The square root of this quantity was introduced in Schuster (1898). The basic components, in (2), are seen to be correlations of the time-series data with cosine and sine functions respectively. In consequence the periodogram may be expected to highlight periodicities in a series $Y(t)$.

Figure 1 presents the stretch of data analyzed by Michelson, and an attempt at reproducing the periodogram he computed. There were 161 observations in the series. The periodogram of Figure 1 was derived numerically, while Michelson employed the harmonic analyzer. Because Michelson's analyzer could handle only 80 observations and there are 161 here, the reproduction is necessarily approximate. The numbers in the figure are the periods that Michelson mentions, and he seems to mention one for each bump in his estimate. The main hump, in the center of the picture, is near the traditional period of 11 years. The large value near frequency 0 occurs because Michelson did not remove the mean of the data prior to the Fourier analysis.

2.3. The Dark Ages.

Michelson (1913) listed 11 periods for the sunspot series. In a like manner Beveridge (1922) lists 19 periods for a wheat price index. Figure 2 provides Beveridge's data and its periodogram. Many peaks are apparent. Another early researcher eager to ascribe peaks in a periodogram to periodicities was Brownlee (1917), who listed 7 periods for measles data for 1838 to 1913 and in fact remarked:

> It might also be said, in the language of VOLTAIRE, that, if these periods were not found, they would require to be invented.

Frightening words to a statistician. Of such period-chasing, Tukey (1980) remarks:

> More lives have been lost looking at the raw periodogram than by any other action involving time series!

Something was clearly amiss with the naive use of the periodogram.

2.4. Progress to Understanding.

In fact Schuster (1898) had been concerned with whether peaks in the periodogram might simply be due to chance. For an individual τ he assessed their significance (prob-value) via the result

$$\operatorname{Prob}\left\{\frac{P^T(\tau)}{\operatorname{ave}(P^T)} > x\right\} \simeq e^{-x}, \tag{3}$$

where $\operatorname{ave}(P^T)$ refers to the average of all the periodogram values.

Later Fisher (1929) recognized that the periodogram was being examined not just at a single period, but as a function of τ. He derived the more pertinent expression for

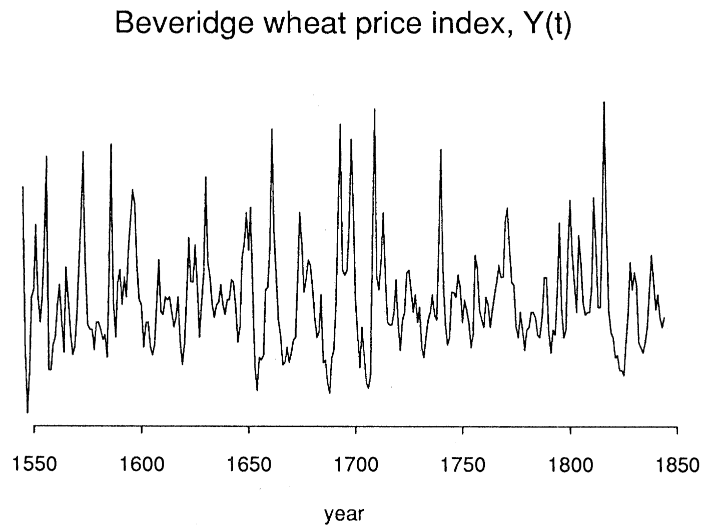
$$\operatorname{Prob}\left\{\max_{\tau} \frac{P^T(\tau)}{\operatorname{ave}(P^T)} > x\right\}. \tag{4}$$

Both of the results (3) and (4) were derived for the case of Gaussian white noise. It was for Whittle (1952) to derive the needed result for a stationary time series. He found that in Fisher's result one replaces $P^T(\tau)$ by $P^T(\tau)/f(2\pi/\tau)$, with $f(\cdot)$ being the true spectrum of the series.

2.5. Definition and Estimation of the Spectrum.

In a remarkable paper Einstein (1914), mentioning sunspots as a motivating example, may be seen to define the spectrum and to provide an estimate. He sets down a "mean value"

$$c(u) = \operatorname*{ave}_{t}\,\{Y(t+u)Y(u)\} \tag{5}$$



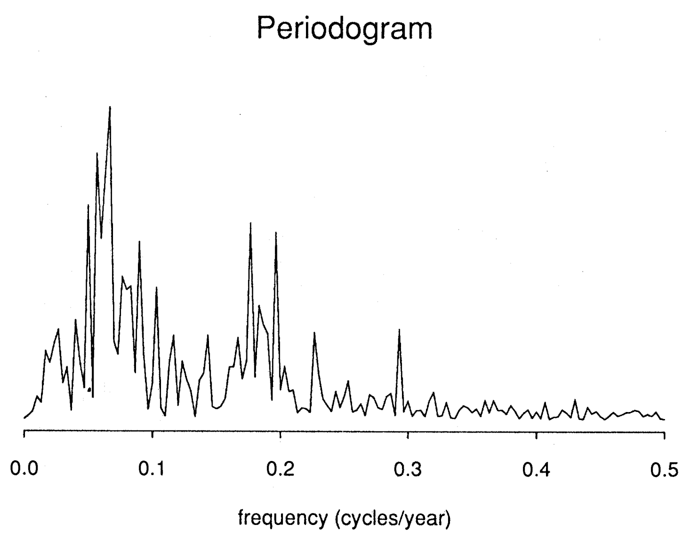


FIGURE 2: The top graph shows the wheat price index data of Beveridge (1922). The bottom graph is the corresponding periodogram, on a linear scale.

and then considers its Fourier transform

$$f(\lambda) = \int \cos \lambda u \, c(u) \, du, \tag{6}$$

referring to the intensity of $Y(\cdot)$.

To estimate $f(\cdot)$, Einstein proposed taking the "mean value" near n of the A_n^2 of the development

$$Y(t) = \sum_n A_n \cos\left(\frac{\pi n t}{T}\right), \qquad 0 < t < T. \tag{7}$$

These A_n are given by

$$\frac{1}{T}\int_0^T \cos\left(\frac{n\pi t}{T}\right) Y(t) \, dt \tag{8}$$

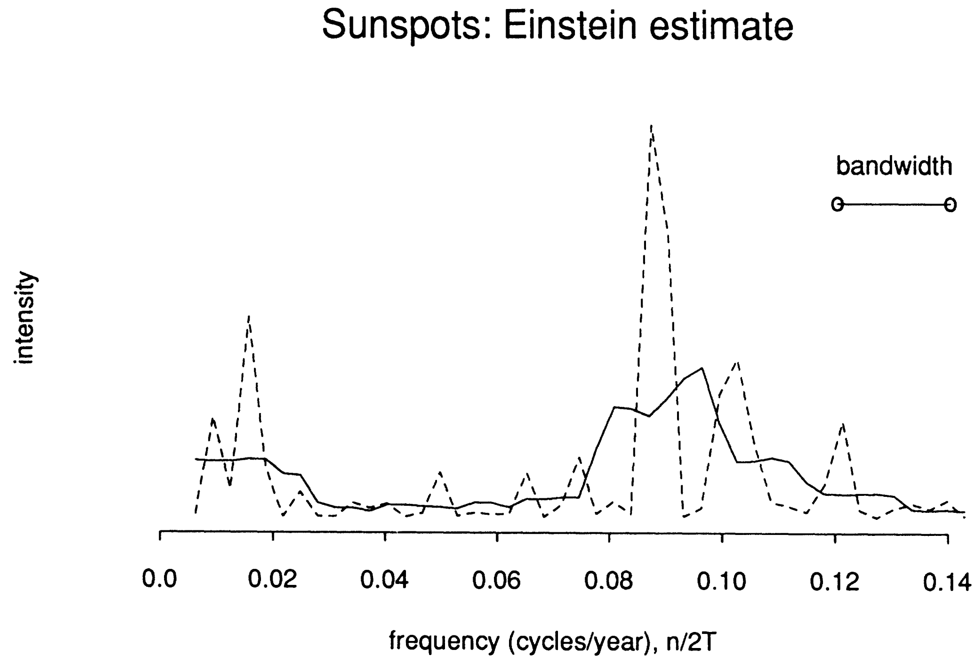


FIGURE 3: An estimate of the spectrum of the sunspot data, smoothing the A_n^2 of (7) over the indicated bandwidth.

Here A_n corresponds to fluctuations at frequency $\lambda = n\pi/T$. For the sunspot data considered above, the results of computing the A_n^2, employing a discrete approximation to (8), are given in Figure 3 as the dashed line. The solid line corresponds to taking a "mean value" over the interval indicated by "bandwidth" on the figure, which produces a much less pronounced 11-year period (0.09 cycles/year).

As provisos, one has to say that Einstein does not make it clear what the "mean value" referred to is, nor whether he had stochastic functions in mind. The notation employed here differs from his.

2.6. The Modern Era.

The modern era of spectrum analysis may be said to begin in the research of Maurice Bartlett and John Tukey in the late 1940s. In particular one may point to the references Tukey and Hamming (1949), Bartlett (1950), and Tukey (1950). The work of these individuals provided an effective estimate of the power spectrum of a stationary time series, given sufficient data. Their research further determined useful approximations to the sampling fluctuations of the estimates. Figure 4 provides an estimate for the sunspot data analyzed earlier. The dashed lines provide approximate 95% confidence limits for each frequency about a heavily smoothed version of the periodogram. From the figure it is apparent that sunspots are far from having a precise period of 11 years. Examination of the series itself shows the lengths of the cycles to vary from 7 to 14 years and for the cycles to be of different shapes.

2.7. Discussion.

It now seems that (6), the formal definition of the power spectrum, is due to Einstein,

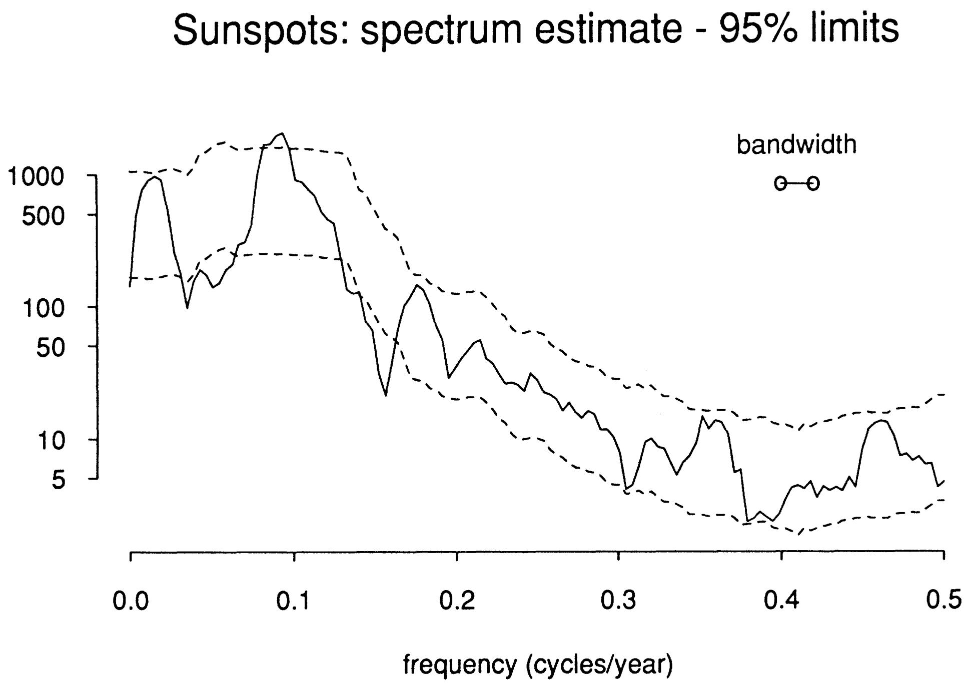


FIGURE 4: An estimate of the power spectrum of the sunspot series with (marginal) confidence bounds indicated.

not Wiener (1930) as had been thought. It also appears that the first effective estimate is due to Einstein, not Wiener (1930) or Daniell (1946) as might be claimed. However, it was Wiener's work that had the influence on the development of the subject. Also some (Masani 1986) doubt Einstein's priority.

It needs to be mentioned that various researchers have made notable contributions to the statistical study of power spectra. One reference to the history and details of others' work is Yaglom (1987b). Another reference on the general history of spectrum analysis is Robinson (1982).

3. SOME APPLICATIONS

Three analyses of scientific data are now presented. This work required close collaboration with substantive scientists. The examples have in common: that frequency analysis elucidates the situation, that a physically based model later provides insight, that a numerical approach is highly flexible, and that a statistical approach handles error and uncertainty.

3.1. Free Oscillations of the Earth.

The empirical spectral analysis of the earth's motion following the great Chilean earthquake of 1960 is viewed by many as *the* success story of numerical spectrum analysis of the sixties. See Tukey (1966), Bath (1974), for example.

After a great earthquake, the Earth "rings" for days at various resonance frequencies (Press 1965). The particular resonance frequencies depend on the structure of the Earth. It is this structure which has long intrigued geophysicists and seismologists. In particular

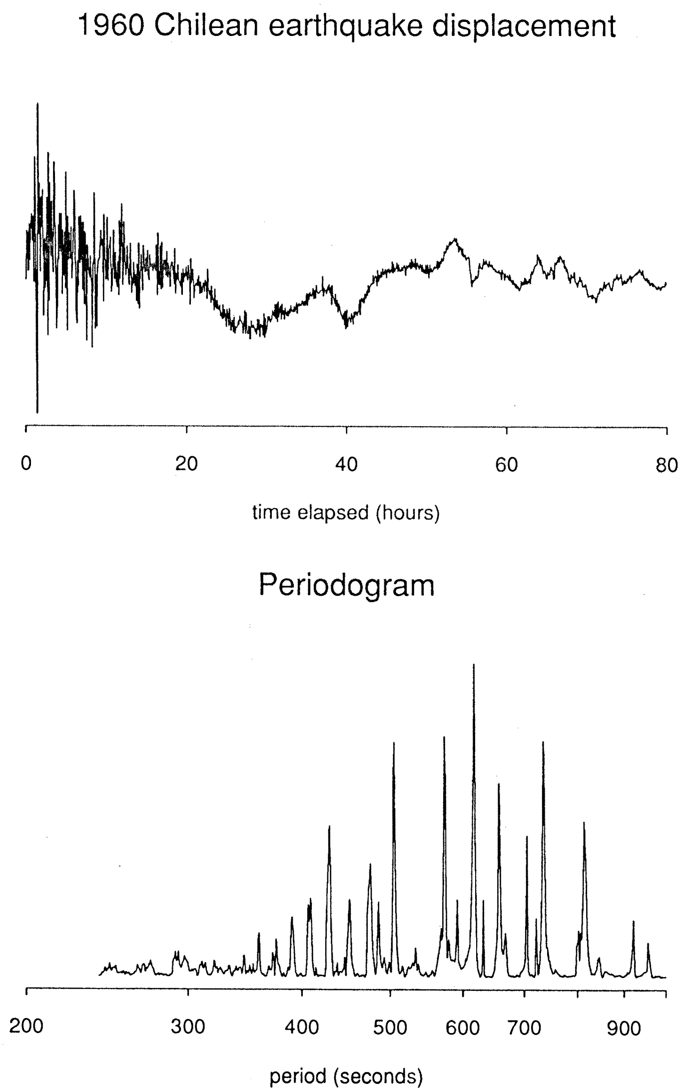


FIGURE 5: The trace of the great Chilean earthquake, corrected for tides, and the corresponding periodogram, on a linear scale.

they have studied the inverse problem of inferring the Earth's structure from the resonance frequencies.

Figure 5 presents a record of the Chilean event recorded at Trieste. The data is described in Bolt and Marusi (1962). The lower figure is the periodogram of that record, graphed against period. (This is the usual seismologist's display.) A variety of peaks are apparent. Values for the periods may be read off, and it seems that the values have varying uncertainty. The periodogram is clearly an important display for assessing the situation, but the trace is far from stationary, and more is needed. In an attempt to estimate the uncertainty (and this is why the seismologist turned to a statistician), one is led to the following steps.

The equations of motion of the earth are approximately linear with constant coefficients

— see Aki and Richards (1980). Such equations have solutions of the form

$$\sum_k \alpha_k e^{-\beta_k t} \cos(\gamma_k t + \delta_k), \qquad t > 0; \tag{9}$$

see Hochstadt (1975). One is thus lead to choose estimates of the α, β, γ, δ so that (9) is near $Y(t)$ for the given data.

A likelihood-motivated analysis of the Chilean data is carried out in Bolt and Brillinger (1979), separately by frequency band. The approximate standard deviations of $\hat{\alpha}_k$ and $\hat{\beta}_k/T$, when derived, are found to have an interesting and intuitive form. Namely they are both proportional to

$$\sqrt{f(\gamma_k)/\alpha_k^2 T^3},$$

where $f(\cdot)$ is the noise spectrum. The estimate of γ_k is more precise for small $f(\lambda_k)$, for large α_k, and for large T.

The derived frequency estimates and associated standard errors can now be taken as input data to the inverse problem of determining the structure of the earth.

3.2. Seismic Surface Waves.

Seismic surface waves are earthquake waves whose energy is trapped near the Earth's surface. They have the property that their velocity of transmission depends on frequency. Figure 6 presents an example of the velocity-frequency relationship for one earth model and Rayleigh waves.

The relationship may be validated by the computation of an empirical dynamic spectrum. This is a display of the estimated intensity of the phenomenon as a function of both time and frequency. A naive way to compute such a spectrum is as

$$\left(\sum_u Y(t-u)\cos\lambda u\right)^2 + \left(\sum_u Y(t-u)\sin\lambda u\right)^2, \tag{10}$$

where u sums over a restricted time interval. Ridges will appear if the different frequency components travel with different velocities. Figure 7 presents, at the top, the vertical displacement trace of the 7 December 1988 magnitude-7.0 Armenian earthquake, as recorded at Berkeley. Below, on the same time scale, is the corresponding dynamic spectrum. One sees the lower-frequency components arriving first, about 1000 seconds after the record starts. In this computation, first a running autoregression was fitted, then (10) was computed based on residuals. The statistic (10) was then corrected for the autoregressive fit. This prewhitening and recoloring is done in order to reduce the bias.

The geophysicist and seimologist hope to learn about the structure of the earth from such data. Figure 8 provides an example of a simple earth model, one of a homogeneous surface layer above an infinite underlayer. For given values of depth h, compressional velocities α_1, α_2, shear velocities β_1, β_2, and densities ρ_1, ρ_2 one can compute curves such as the one of Figure 6; see Bolt and Butcher (1960). This suggests that one should be able to put together the arrival times of the various frequency components with the theoretical arrival times for a given earth model parameter θ, and thereby estimate θ.

In particular one can proceed as follows: determine $\hat{t}(\lambda)$, the time at which the frequency λ component arrives. One knows the distance x from the observatory to the earthquake location. Thus velocities $\hat{U}(\lambda) = x/\hat{t}(\lambda)$ may be computed. Next, suppose one measures the nearness of model parameter θ to the true value by the sum of squares

$$\sum_\lambda \{\hat{U}(\lambda) - U(\lambda|\theta)\}^2,$$

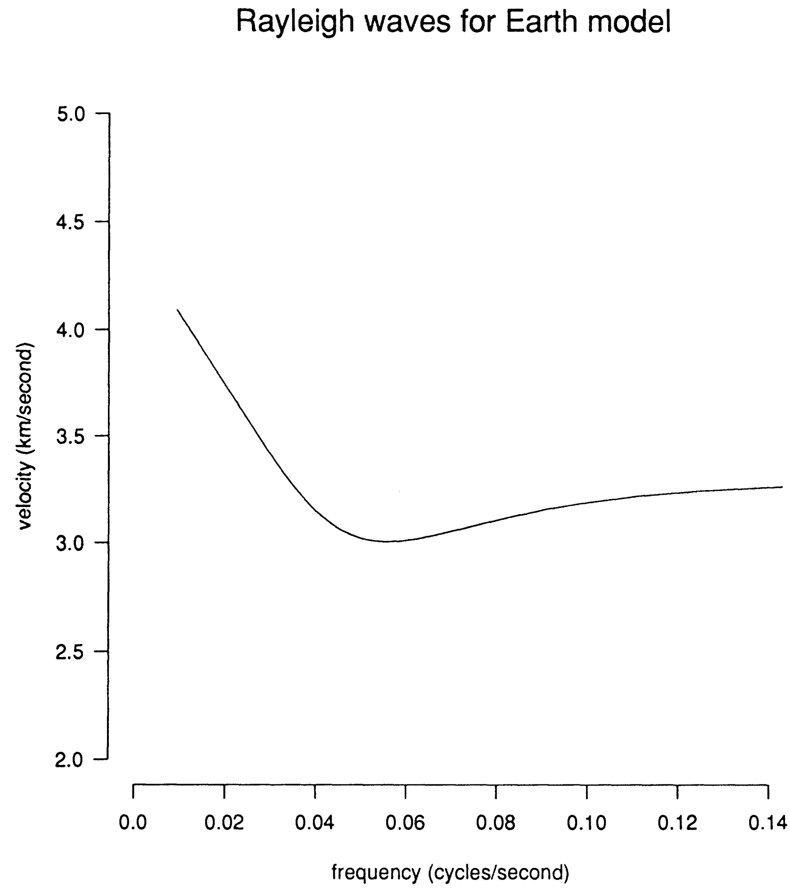


FIGURE 6: Velocity as a function of frequency for an earth model of the form of Figure 8.

summing over the various λ, and then estimates θ by the minimizing value. Figure 9 gives the $\hat{U}(\lambda)$ and the fitted curve $U(\lambda|\theta)$ for the Berkeley station. The fit appears reasonable, particularly at the lower frequencies where the signal-to-noise ratio is greatest. One can aggregate the data for several stations to obtain a group solution and further obtain estimated standard errors for the parameters. Details are provided in Brillinger (1993). A full likelihood procedure for the semiparametric modelling of the seismogram itself is currently under development with B.A. Bolt.

In this example the idea of spectrum has been central to the development of an estimation procedure for parameters which have direct physical interpretation.

3.3. NMR Spectroscopy.

Nuclear magnetic resonance (NMR) is a quantum-mechanical phenomenon. A resonance effect occurs in particular substances when an oscillation frequency of a surrounding radio frequency field coincides with a nuclear precession frequency of the substance. The purpose of the spectroscopy is to infer the structure of the substance. The data consist of the fluctuating voltage response $Y(t)$ to an applied magnetic field $X(t)$. Various inputs $X(\cdot)$ are employed, e.g., a pulse, sequences of fluctuating pulses, and sine waves. Becker and Farrar (1972) is one review paper.

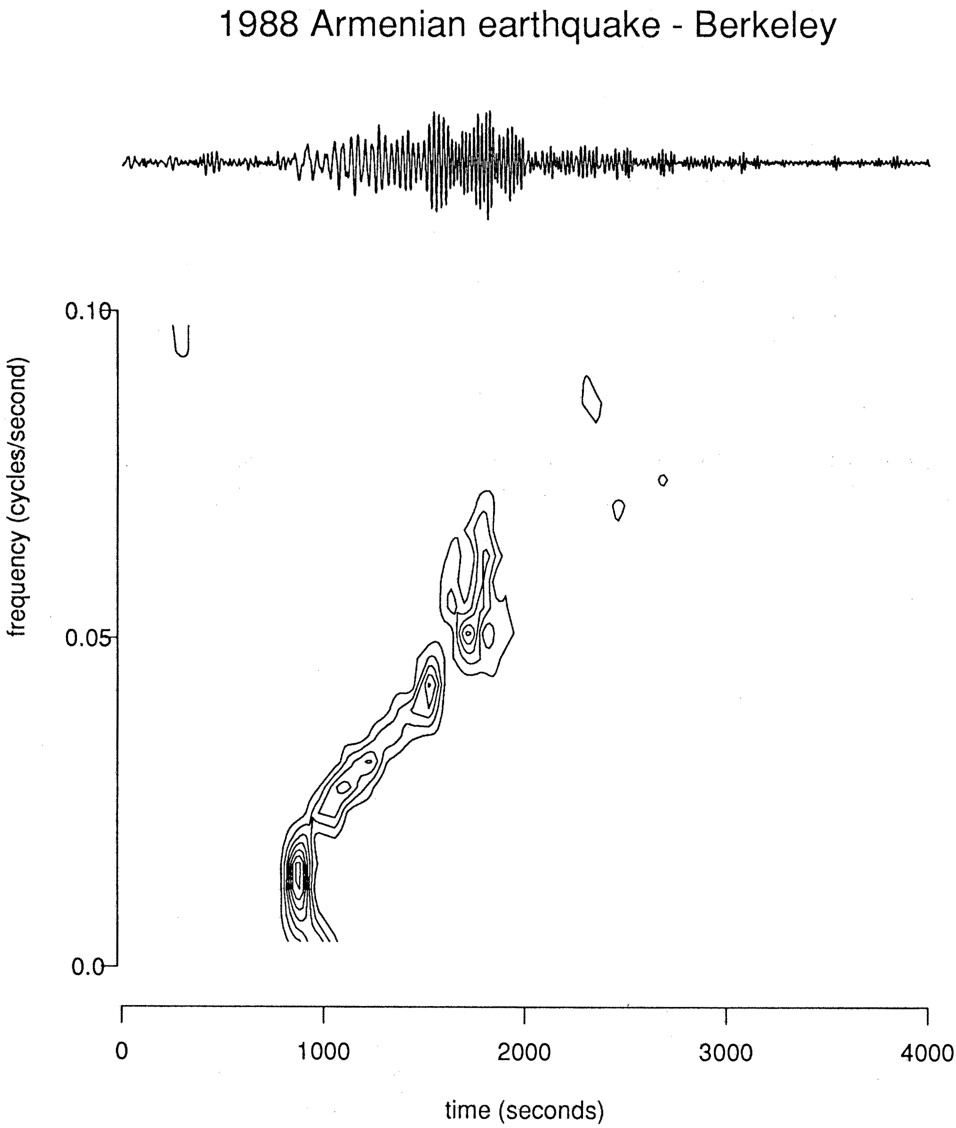


FIGURE 7: The top graph gives the trace of the 7 December 1988 earthquake in northern Armenia. Below, on the same time scale, is a dynamical spectrum expressed as a contour plot.

Ernst and Anderson (1966) proposed the computation of the periodogram

$$\left(\sum_t Y(t)\cos\lambda t\right)^2 + \left(\sum_t Y(t)\sin\lambda t\right)^2,$$

$0 \leq \lambda \leq \pi$, for such data following a pulse input. An example is given in Figure 10 in the case of a simulated model for the substance 2,3-dibromothiophene (2,3-DBT). The top graph is the response itself, the bottom the absolute value of the Fourier transform of the response. There are two doublets. These result from two hydrogen atoms of the 2,3-DBT. It is much easier to read the lower graph than to puzzle out the structure from the upper one.

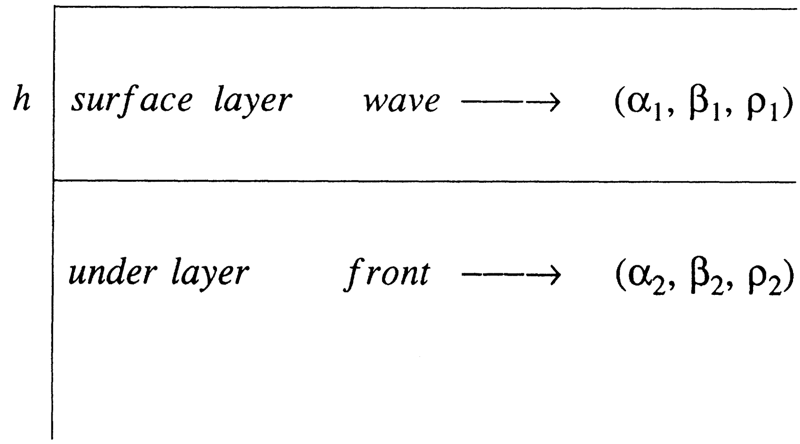


FIGURE 8: Schematic of an earth model of a surface layer of depth h over a half space. The wave propagates from left to right. The parameters are represented by α, β, ρ.

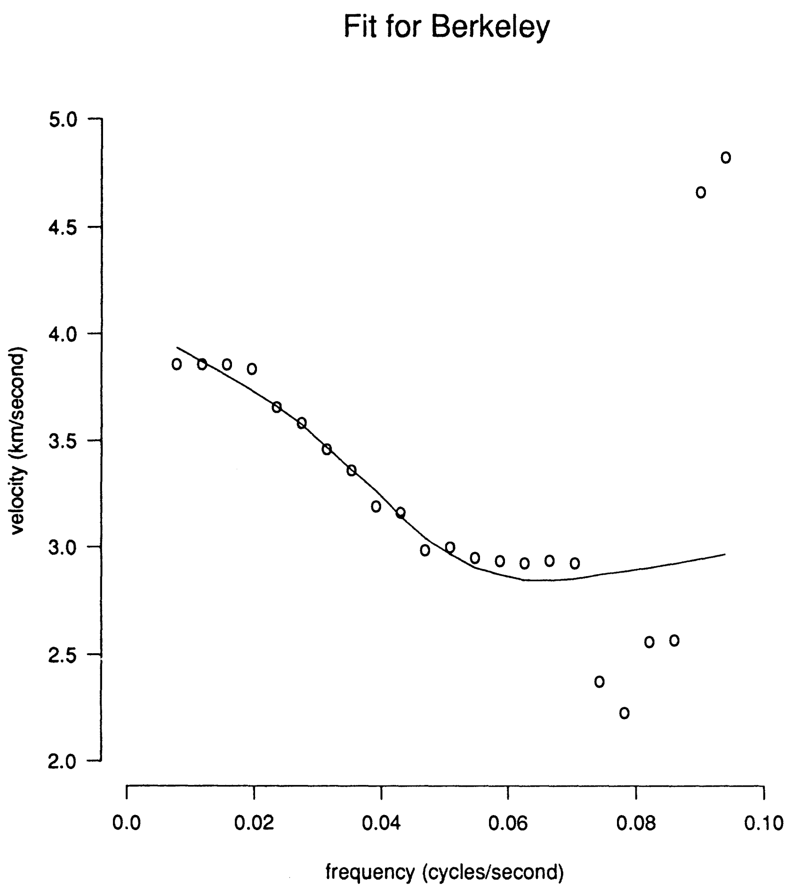


FIGURE 9: For each frequency, the points give the velocity at which the dynamic spectrum is largest. The line is the fitted curve.

Later Ernst (1970) and Kaiser (1970) proposed taking the input $X(\cdot)$ to be random or pseudorandom and employing a form of cross-correlation analysis. Figure 11 shows 4 seconds of actual response of an experimental sample of 2,3-DBT to binary-noise input.

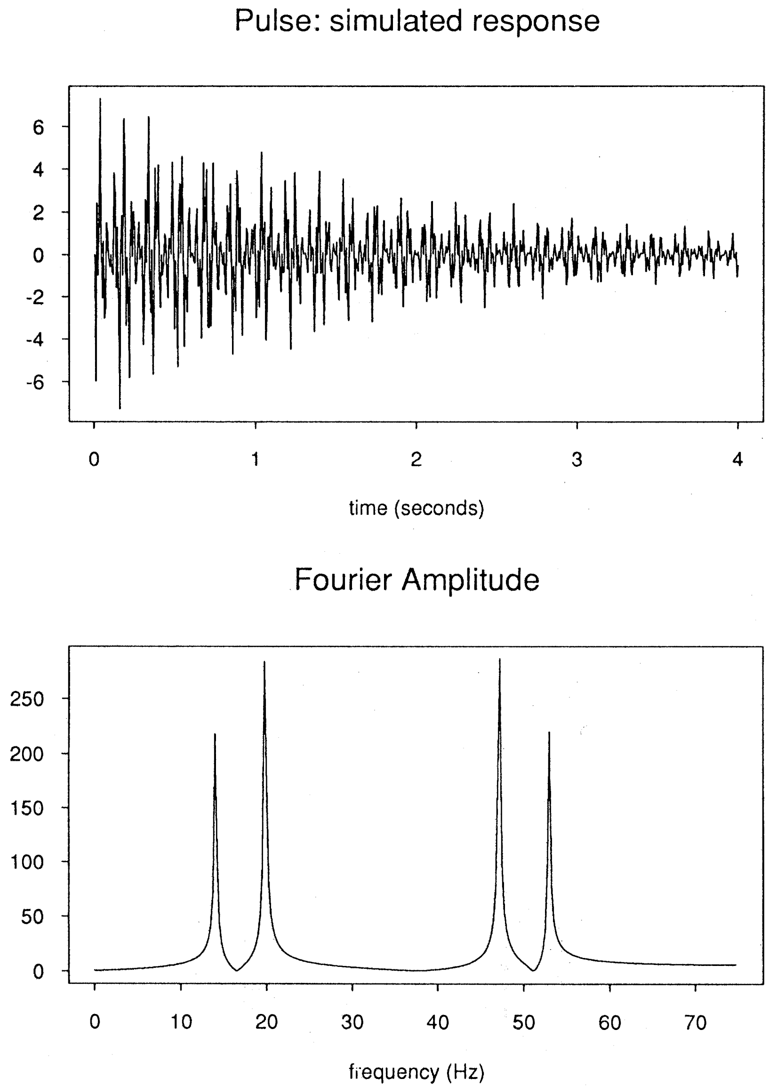


FIGURE 10: Top: plot of response of a simulated 2,3-DBT molecule to a pulse. Bottom: the absolute value Fourier transform of the response.

The lower part of the figure provides the square root of the periodogram. As in Figure 10, two doublets are apparent, but reading their locations is not easy. Since the input $X(\cdot)$ is available, cross-spectral analysis is a more appropriate tool. Figure 12 provides an estimate of the modulus of the transfer function of the system,

$$\frac{|\hat{f}_{YX}(\lambda)|}{\hat{f}_{XX}(\lambda)}, \tag{11}$$

where $\hat{f}_{YX}(\lambda)$ is obtained by smoothing the cross-periodogram

$$(2\pi T)^{-1}\left(\sum_t Y(t)e^{-i\lambda t}\right)\left(\sum_t X(t)e^{i\lambda t}\right)$$

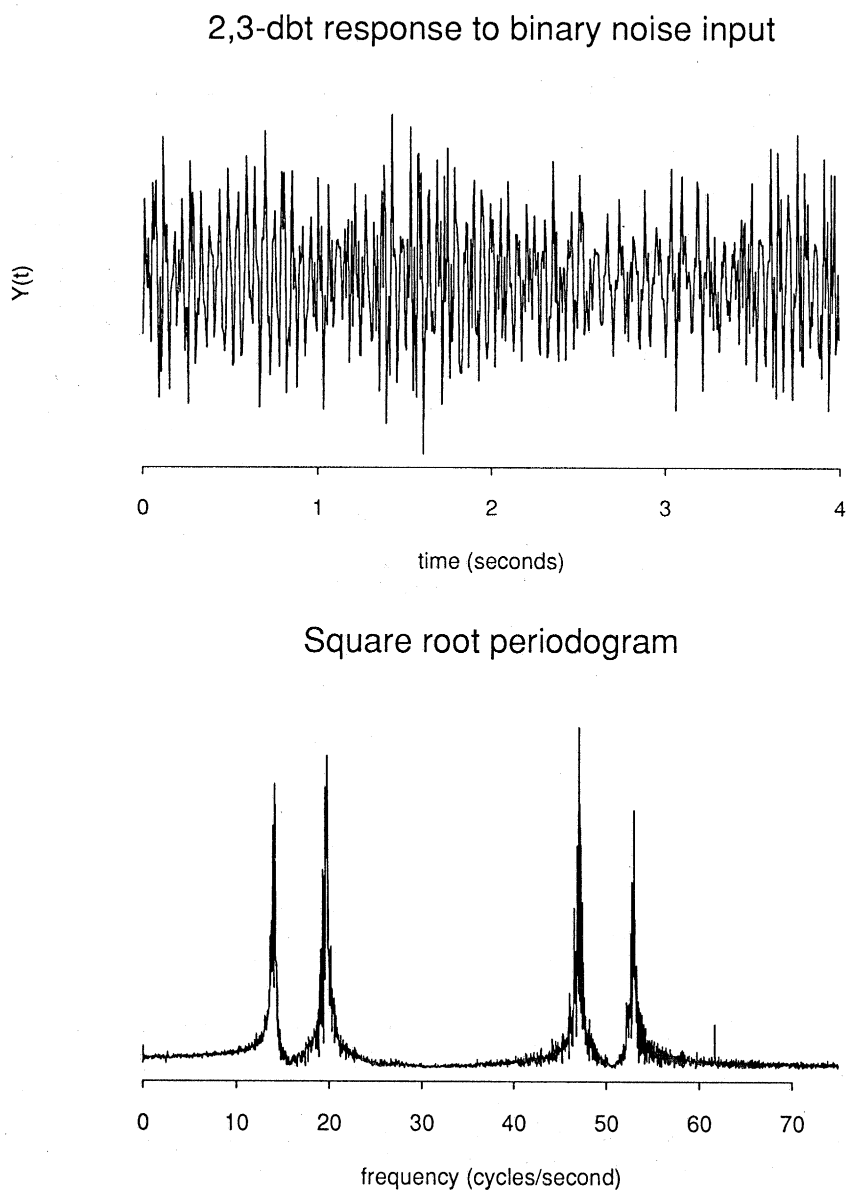


FIGURE 11: Top graph: response of 2,3-DBT to white-noise input. Bottom graph: absolute value of the Fourier transform of the response.

with $\hat{f}_{XX}(\cdot)$ a similar smoothed periodogram of $X(\cdot)$. This figure is much nearer to the ideal of Figure 10.

In fact there exists substantial theory concerning the NMR phenomenon. In particular it may be described by the Bloch equations, which take the form

$$\frac{d\mathbf{S}(t)}{dt} = \mathbf{a} + \mathbf{A}\mathbf{S}(t) + \mathbf{B}\mathbf{S}(t)X(t), \tag{12}$$

$$Y(t) = \mathrm{Re}\{\mathbf{c}^{\mathsf{T}}\mathbf{S}(t)\} \tag{13}$$

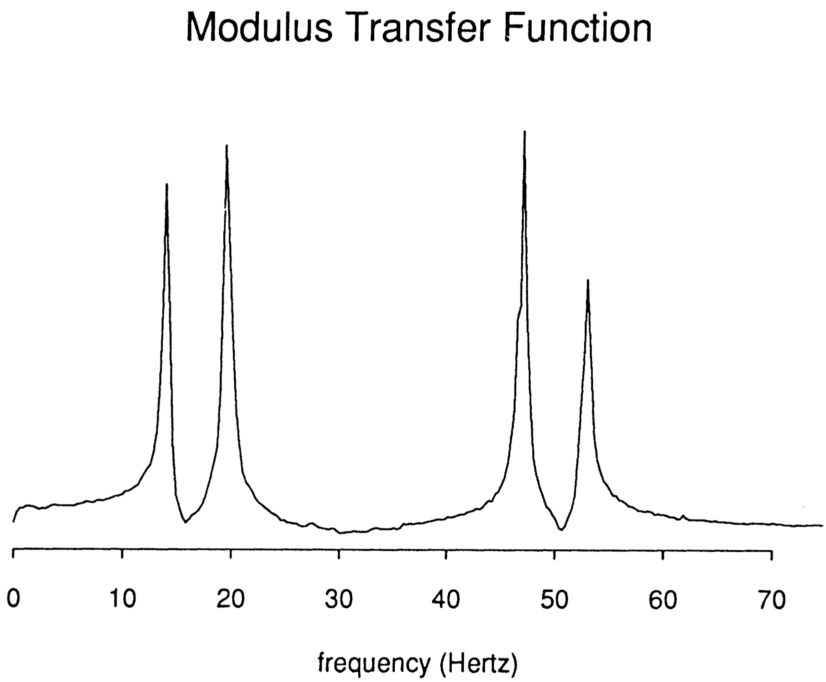


FIGURE 12: Graph of the statistic (11).

for $\mathbf{S}(\cdot)$ a 16-dimensional state vector with complex entries, and $X(\cdot)$, $Y(\cdot)$ corresponding input and output of the system. The parameters of interest are entries of **A**, **B**. The vector **c** is given by the experimental setup. The parametrization of **A**, **B** is provided by the chemist.

The estimation problem may now be approached by nonlinear regression. For given **A**, **B** and initial state $\mathbf{S}(0)$, the solution of (12) may be determined numerically in the case that the input $X(\cdot)$ is piecewise constant, as it is here. Then the parameters may be estimated by putting the data $Y(\cdot)$, $t = 0, \ldots, T-1$, up against the value determined from (13) and the numerical solution. In this way one obtains estimates and associated standard errors. Further details on this work may be found in Brillinger and Kaiser (1992).

3.4. Discussion.

These examples each involve a progression from a naive (spectral) analysis to a likelihood analysis founded on substantive subject matter. This last leads to efficiency, uncertainty estimation, and the ability to make general statistical inferences.

Because of the complication of the basic circumstances, it seems necessary to collaborate with substantive scientists on these problems. In these cases it was not the data analysis that suggested the model, rather it came from the scientific background.

4. FUTURE PROSPECTS

So what is ahead for spectrum analysis in the pot of gold at the end of the rainbow? Spectrum analysis has already found many uses, and there are quite a number of questions it can address. John Tukey, for example, has emphasized its role in the discovery of phenomena; see Tukey (1966, 1980). Tukey (1980) contains other speculations on the future of time series. Brillinger (1987) mentions a variety of specific problems related to Fourier inference. They are repeated below and some additional ones added.

Turning to future possibilities, there are near-immediate generalizations of the vast majority of time-series techniques to other types of "function" data, such as images, point processes, moving surfaces, and tesselations.

One can speculate on various aspects of contemporary and future research on the spectrum analysis of time series. Many times it has been said that statistical spectrum analysis is part art and part science. The need to be an artist is likely to diminish as automatic bandwidth selection algorithms are developed, analagous to those coming into use in density and nonparametric regression estimation. For making inferences it is necessary to have some indication of the sampling variability of the statistics computed. Various jackknife, perturbation and bootstrap techniques are under development. It will be some time before that research is finished, for procedures will be required for infinite-dimensional parameters and nonstationary series. In connection with nonstationarity, it may be remarked that effective procedures for estimating spectra from short stretches of data are needed, because many series appear to be at most locally stationary. Perhaps those estimates will come from semiparametric approaches involing both finite- and infinite-dimensional parameters. An examination of the contemporary literature shows much work being carried out on non-Gaussian series, on long memory series, on nonlinear models, and on self-similar processes. Finally, new scientific devices, such as lasers and nonlinear crystals, are leading to high-quality data sets with unusual inference problems.

Some particular research problems related to the topics of the paper are:

1. Diagnostics, influence, robust/resistant procedures.
2. Missing values, quantization, jitter.
3. Estimation of dimension, e.g., by AIC.
4. Inverse-problem formulations, e.g., ridge regression.
5. Local asymptotic normality, contiguity.
6. Adaptive procedures.
7. The absorption model.
8. Signal-dependent noise.
9. Law of the iterated logarithm, large deviations, rates of convergence for the estimates.
10. Random-effects models.
11. Vector-valued cases.
12. Partially parametric formulations, e.g., the periodic case.
13. Models for point-process and telegraph-signal cases.
14. Expansions for distributions.
15. Distributions of test statistics, e.g., of

$$\sup_{\rho,\lambda} \frac{\left|\sum_{t=0}^{T-1} \rho^t e^{-i\lambda t} Y(t)\right|}{\sum_{t=0}^{T-1} \rho^{2t}}$$

or of

$$\sup_{\lambda,\mu} \min \{I^T(\lambda),\ I^T(\mu),\ I^T(\lambda+\mu)\}.$$

16. Properties of the estimates when the model is untrue.
17. The broadband-signal case.
18. Parametric analysis of the frequency case.
19. Approximate distribution of the likelihood-ratio test statistic for the presence of a plane wave crossing an array.
20. Sampling properties of the NMR estimates.
21. Techniques for handling small amounts of nonstationary data and massive amounts of regular data.
22. Uncertainty evaluation for the non-i.i.d. case.
23. Errors in variables.
24. Irregularly observed values.
25. Parametric models for nonstandard processes.
26. Asymptotics when the parameter dimension tends to infinity.
27. Characterization of covariance and spectrum functions for 0-1-valued series.
28. Characterization of the spectrum of a stationary point process.
29. Properties of i.i.d.-motivated techniques in dependent cases.
30. Aliasing regions for higher-order spectra of stationary spatial processes.
31. Simulation of processes with given higher-order spectra.
32. Combination of groups of experiments.
33. Populations that are mixtures.
34. Design, e.g., of NMR input.
35. Study of causal networks of general processes.
36. Less-biased estimates of coherence.
37. Detection of long lags between series.
38. Properties of empirical Fourier transforms of unusual processes.
39. Stable algorithms for fitting ARMAs.
40. Employment of contemporary optimization methods.
41. Special models for abstract-valued processes.
42. Automatic determination of bandwidth parameters.
43. Modelling of wildly nonstationary values, e.g., TV signals or computer tasks.
44. Analysis of qualitative-valued processes.
45. How to define "trend" and "period".
46. Fast computation of Fourier transforms for irregularly placed data.
47. Assessing goodness of fit for process models.
48. Measuring association of abstract-valued processes.
49. Strong approximations for statistics based on process values.
50. Detecting change in, or the presence of, a signal.
51. Incorporation of symmetries and invariance.
52. Analysis of censored data.
53. Improved estimates.
54. Relating processes observed at different locations.
55. Study of multivalued random functions.
56. Statistical aspects of determining $g(\cdot)$ from $V(\cdot)$ in (1), including the case when $V(0)$ is unestimated.

5. CONCLUSIONS

Statistics has long had an intimate connection with the physical sciences. Some statistical concepts, like the spectrum, have been motivated by physical considerations; others have arisen abstractly and gone on to influence practice. (The cross-spectrum and bispectrum might be mentioned as examples of the latter.) Statistics researchers benefit from looking towards science for suggestions of new concepts and techniques. There is often a further bonus: the scientific circumstance suggests a path towards a solution. Students are often surprised that statistical concepts have direct physical interpretations.

This paper is part historical and part contemporary collaborative research material. One learns from history. In the case of the present topic one sees the important ideas of spectrum, direct spectrum estimate, and FFT missed for a number of years. [For a discussion of the FFT case see Heideman *et al.* (1984).] It is important to know the past; otherwise gems are missed. Further, history can be a stimulating pedagogic tool for introducing the important concepts and ideas, and often the independent reader will find the original writers much clearer than the later ones.

The applications had in common that a spectral analysis started the investigation, then substantive subject matter was employed to develop a conceptual model and a likelihood analysis carried out to complete the work.

ACKNOWLEDGEMENTS

The author is grateful to Bruce Bolt for various discussions on the material concerning free oscillations and surface waves. For similar help with the material on nuclear magnetic resonance, he is grateful to Reinhold Kaiser. Bob Shumway helped the author in obtaining an old reference, Alan Izenman helped find an important picture of a Spectrum, and Stu McCormick provided some pertinent material.

REFERENCES

Aki, K., and Richards, P.G. (1980). *Quantitative Seismology. Volumes I and II.* Freeman, San Francisco.

Bartlett, M.S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, 37, 1–16.

Bath, M. (1974). *Spectral Analysis in Geophysics.* Elsevier, Amsterdam.

Becker, E.D., and Farrar, T.C. (1972). Fourier transform spectroscopy. *Science*, 178, 361–368.

Beveridge, W.H. (1922). Wheat prices and rainfall in western Europe. *J. Royal Statist. Soc.*, 85, 412–459.

Bolt, B.A., and Butcher, J. (1960). Rayleigh wave dispersion for a single layer on an elastic half space. *Australian J. Physics*, 13, 498–504.

Bolt, B.A., and Brillinger, D.R. (1979). Estimation of uncertainties in eigenspectral estimates from decaying geophysical time series. *Geophys. J. Roy. Astronom. Soc.*, 59, 593–603.

Bolt, B.A., and Marusi, A. (1962). Eigenvibrations of the earth observed at Trieste. *Geophys. J. Roy. Astronom. Soc.*, 6, 299–311.

Brillinger, D.R. (1987). Fitting cosines: Some procedures and some physical examples. *Applied Probability, Stochastic Processes and Sampling Theory* (I.B. MacNeill and G.J. Umphrey, *eds.*), Reidel, Boston, 75–100.

Brillinger, D.R. (1993). An application of statistics to seismology: dispersion and modes. Developments in Time Series Analysis. Chapman and Hall, London.

Brillinger, D.R., and Kaiser, R. (1992). Fourier and likelihood analysis in NMR spectroscopy. *New Directions in Time Series Analysis, Part I* (D.R. Brillinger, E. Parzen, and M. Rosenblatt, *eds.*), Springer, New York, 41–64.

Brownlee, J. (1917). An investigation into the periodicity of measles epidemics in London from 1703 to the present day by the method of the periodogram. *Philos. Trans. Roy. Soc. London Ser. B*, 208, 225–250.

Daniell, P.J. (1946). Discussion of paper by M.S. Bartlett. *J. Roy. Statist. Soc. Suppl.*, 8, 88–90.

Dziewonski, A.M., and Hales, A.L. (1972). Numerical analysis of dispersed seismic waves. *Seismology: Surface Waves and Earth Oscillations* (B.A. Bolt, *ed.*), Academic Press, New York, 39–85.

Einstein, A. (1914). Méthode pour la détermination des valeurs statistiques d'observations concernant des grandeurs soumises à des fluctuations irrégulières. *Arch. Sci. Phys. et Naturelles Ser. 4*, 37, 254–255.

Ernst, R.R. (1970). Magnetic resonance with stochastic excitation. *J. Magnetic Resonance*, 3, 10–25.

Ernst, R.R., and Anderson, W.A. (1966). Application of Fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instrum.*, 37, 93–102.

Fisher, R.A. (1929). Tests of significance in harmonic analysis. *Proc. Royal Soc. London Ser. A*, 125, 54–59.

Gouy, G.L. (1886). Sur le mouvement lumineux. *J. Physique*, 5, 354–362.

Heideman, M.T., Johnson, D.H., and Burrus, C.S. (1984). Gauss and the history of the fast Fourier transform. *IEEE Acoust. Speech Signal Process Magazine.*, 1, 14–21.

Hochstadt, H. (1975). *Differential Equations*. Dover, New York.

Kaiser, R. (1970). Coherent spectrometry with noise signals. *J. Magnetic Resonance*, 3, 28–34.

Kimura, H. (1913). On the harmonic analysis of sun-spot relative numbers. *Monthly Notices Roy. Astronom. Soc.*, 73, 543–548.

Masani, P.R. (1986). Einstein's contribution to generalized harmonic analysis and his intellectual kinship with Norbert Wiener. *Jahrb. Überblicke Math.*, 191–209.

Michelson, A.A. (1891). On the application of interference methods to spectroscopic measurements — I. *Philos. Mag.*, 33, 338–346.

Michelson, A.A. (1892). On the application of interference models to spectroscopic measurements — II. *Philos. Mag.*, 34, 280–299.

Michelson, A.A. (1902). *Light Waves and Their Uses*. Univ. Of Chicago Press, Chicago.

Michelson, A.A. (1913). Determination of periodicities by the harmonic analyzer with an application to the sun-spot cycle. *Astrophys. J.*, 38, 268–274.

Michelson, A.A., and Stratton, S.W. (1898). A new harmonic analyzer. *Amer. J. Sci.*, 5, 1–13.

Press, F. (1965). Resonant vibrations of the earth. *Sci. Amer.*, 213, 28–38.

Robinson, E.A. (1982). A historical perspective of spectrum estimation. *Proc. IEEE*, 70, 885–907.

Rayleigh, L. (1892). On the interference bands of approximately homogeneous light. *Philos. Mag.*, 34, 407–411.

Schuster, A. (1898). On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terr. Mag.*, 3, 13–41.

Sobel, M.I. (1989). *Light*. Univ. of Chicago Press, Chicago.

Topper, D. (1990). Newton on the number of colours in the spectrum. *Stud. Hist. Philo. Sci.*, 21, 269–279.

Tukey, J.W. (1950). The sampling theory of power spectrum estimates. *Symposium on Applications of Autocorrelation Analysis to Physical Problems*, Office of Naval Research, Washington, 47–67. *The Collected Works of John W. Tukey. Volume 1: Time Series, 1949–1964* (D.R. Brillinger, *ed.*), Wadsworth, Belmont, Calif., 129–160.

Tukey, J.W. (1966). Uses of numerical spectrum analysis in geophysics. *Bull. Internat. Statist. Inst.*, 41, 267–307.

Tukey, J.W. (1980). Can we predict where "time series" should go next? *Directions in Times Series* (D.R. Brillinger and G.C. Tiao, *eds.*), IMS, Hayward, Calif., 1–31.

Tukey, J.W., and Hamming, R.W. (1949). Measuring noise color. Bell Telephone Laboratories Technical Memorandom. *The Collected Works of John W. Tukey, Volume 1: Time Series, 1949–1964* (D.R. Brillinger, *ed.*), Wadsworth, Belmont, Calif., 1–127.

Wiener, N. (1930). Generalized harmonic analysis. *Acta Math.*, 55, 117–258.

Whittle, P. (1952). The simultaneous estimation of a time series harmonic components and covariance structure. *Trab. Estad.*, 3, 43–57.

Yaglom, A.M. (1987a). Einstein's 1914 paper on the theory of irregularly fluctuating series of observations. *IEEE Acoust. Speech Signal Process Magazine*, 4, 7–11.

Yaglom, A.M. (1987b). *Correlation Theory of Stationary and Related Random Functions. Volumes I, II.* Springer-Verlag, New York.

Received 21 May 1992
Revised 17 July 1992
Accepted 21 August 1992

Department of Statistics
University of California
Berkeley, California
U.S.A. 94720