

## JOHN W. TUKEY'S WORK ON TIME SERIES AND SPECTRUM ANALYSIS<sup>1</sup>

BY DAVID R. BRILLINGER

*University of California, Berkeley*

The contributions of John W. Tukey to time series analysis, particularly spectrum analysis, are reviewed and discussed. The contributions include: methods, their properties, terminology, popularization, philosophy, applications and education. Much of Tukey's early work on spectrum analysis remained unpublished for many years, but the 1959 book by Blackman and Tukey made his approach accessible to a wide audience. In 1965 the Cooley–Tukey paper on the Fast Fourier Transform spurred a rapid change in signal processing. That year serves as a boundary between the two main parts of this article, a chronological review of JWT's contributions, decade by decade. The time series work of Tukey and others led to the appearance of kernel and nonparametric estimation in mainstream statistics and to the recognition of the consequent difficulties arising in naive uses of the techniques.

**1. Introduction.** John W. Tukey (JWT) was one of the pioneers of twentieth century statistics. Near single-handedly he established the practical computation and interpretation of time series spectra among many other contributions.

A univariate time series is a real-valued function of a real-valued variable called time. The scientific analysis of time series has a very long history. Indeed, Tufte [14] presents a purported tenth-century time series plot concerning the rotation of the planets. Spectrum analysis of time series may be thought of as having commenced in 1664, when Isaac Newton decomposed a light signal into frequency components by passing the signal through a glass prism. In 1800 Herschel measured the average energy in various frequency bands of the sunlight's spectrum by placing thermometers along Newton's spectrum. Mathematical foundations for the concept began to be laid in the mid-1800s when Gouy represented white light as a Fourier series. Later Rayleigh replaced the series by an integral. In 1872 Lord Kelvin built a harmonic analyzer and a harmonic synthesizer for use in the analysis and prediction of the series of the height of the tide at a particular location. His devices were mechanical, based on pulleys. During the same time period a variety of workers, for example, Stokes, were carrying out numerical Fourier analyses using computation schedules. In 1898 Michelson and Stratton described a harmonic analyzer (based on springs) and used it to obtain the

---

Received September 2001; revised February 2002.

<sup>1</sup>Supported by NSF Grants DMS-97-04739 and DMS-99-71309.

AMS 2000 subject classifications. 01A60, 01A70, 62M10, 62M15, 65T50.

Key words and phrases. Bispectrum, cross-spectrum, coherence, FFT, history, John W. Tukey, spectrum analysis, time series.

Fourier transform of a function. This Fourier transform provided an estimate of the power spectrum of the signal. Michelson envisaged the signal as a sum of cosines. He saw the estimated spectra as descriptive statistics of the light emitting sources. In a series of papers, written during the years 1894–1898 Schuster proposed and discussed the periodogram statistic based on an observed stretch of a time series. His motivation was a search for “hidden periodicities.”

In the succeeding years many workers computed periodograms and their equivalents for a variety of phenomena. Starting in 1930 Wiener, Cramér, Kolmogorov, Bartlett, and Tukey produced substantial developments in time series analysis. This article reviews some of the contributions of John Tukey. The final part of the References lists the time series papers in chronological order.

Tukey worked in many of the fields where time series data were present. In particular he contributed to the development and popularization of statistical spectrum analysis. For reference his definitions include [54]:

Time series analysis consists of all the techniques that, when applied to time series data, yield, at least sometimes, either insight or knowledge, *and* everything that helps us choose or understand these procedures.

and [3]:

Spectrum analysis is thinking of boxes, inputs and outputs in sinusoidal terms.

A related topic is “frequency analysis” defined as inquiring

how different bands of frequency appear to contribute to the behavior of our data.

He acknowledged that his work in other fields of statistics and data analysis was often driven by his work in time series analysis [56]:

It is now clear to me that spectrum analysis, with its challenging combination of amplified difficulties and forcible attention to reality, has done more than any other area to develop my overall views of data analysis.

A very crude chronology of JWT’s time series work is: the power spectrum and the indirect estimate, the fast computation of the Fourier transform and the direct estimate followed by uses of the Fast Fourier Transform (FFT) and suggestions for robust variants. Around those items he proposed many novel methods for practical implementation.

In this paper Parts I and II correspond to the years before the Cooley–Tukey [37] paper and the years after: the “indirect” and “direct” periods, respectively, from the names of the spectrum estimates generally employed in those periods. The individual sections discuss successive decades in chronological order.

There are two Appendices: (A) A letter from Norbert Wiener, and (B) The doctoral theses on time series that JWT supervised. Nearly all of the time series papers appear in Volumes I and II of *The Collected Works of John W. Tukey* with some discussion by the Editor and Tukey’s rejoinder.

This article focuses on the highlights of successive papers putting things in a historical context. The paper [4] provides some history of time series analysis in the United States up until the mid-seventies.

## PART I: THE "INDIRECT" YEARS

**2. The 1940s.** The accompanying article [5] reviews Tukey's joining the Fire Control Research Office (FRCO) in Princeton, the people he worked with there and some of the problems studied. While at FRCO Tukey became acquainted with Norbert Wiener's seminal memorandum *Extrapolation, Interpolation and Smoothing of Stationary Time Series* [19]. Wiener's early work [18] had already had dramatic effect on the field of time series analysis, but the 1942 memorandum influenced engineering work irreversibly. A letter Wiener wrote to Tukey June 20, 1942 makes it apparent that Tukey was already involved with time series computations. That letter is reproduced in Appendix A. In a review of [18, 23], Tukey's stated intention was to make Wiener's approach accessible to statisticians. The review is noteworthy in contrasting the functional and stochastic approaches to the foundations of time series analysis. Tukey again comments on the functional approach and Wiener's contributions in [29]. (The functional approach envisages single functions, as opposed to an ensemble, for which long term averages exist.) As indicated in [5], during the war years Tukey was also influencing the work of Leon Cohen on computations of an algorithm for a gun's tracking an enemy airplane in preparation for firing.

The article [5] describes some of Tukey's work starting in 1945 at Bell Labs leading to the Nike anti-aircraft missile system. That work led to the paper, "Linearization of solutions in supersonic flow" [20]. It contains the statement:

Compare replacing the equation  $F(x, \lambda) = 0$  by a linear one  $L(x, \lambda) = 0$  with exact solution  $x = g(\lambda)$  to using instead  $x = h(\lambda; A_0, B_0)$  where  $A_0, B_0$  are obtained from  $(x_j, \lambda_j)$ ,  $j = 1, 2$ , satisfying  $F(x_j, \lambda_j) = 0$  and a guessed solution  $x(\lambda) = h(\lambda; A, B)$ .

This statement foreshadows Tukey's oft-quoted aphorism concerning approximations versus exactness [15]:

Far better an approximate answer to the *right* question, which is often vague, than an *exact* answer to the wrong question, which can always be made precise.

On a number of occasions JWT told the story of how he got interested in the spectral analysis of time series per se. In the late 1940s a Bell Telephone Laboratories engineer, H. T. Budenbom, working on tracking radars was heading to a conference and wished to show a slide of an estimated power spectrum. He met with Richard Hamming and JWT. Hamming and Tukey knew the reciprocal Fourier relations between the autocorrelation function and the spectrum of a stationary time series. They wrote it as [22]

$$\rho_p = \text{corr}\{X_i, X_{p+i}\} = \int_0^\infty \cos p\omega dP_0(\omega),$$

with  $P_0$  a “normalized power spectrum” and [22]

$$p_0(\omega) = \frac{dP_0}{d\omega} = \frac{1}{2\pi} \sum_p \rho_p \cos p\omega,$$

the spectral density. Tukey and Hamming computed the empirical Fourier transform of the sample autocorrelation function of the radar data. The resulting estimate was seen to oscillate. This led Hamming to remark that the estimate would look better if were smoothed with the weights  $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ . This was done, and the result was a much improved picture. To quote JWT, “Dick (Hamming) and I then spent a few months finding out why.” And so began JWT’s major involvement in the field of time series analysis, particularly spectrum analysis.

Hamming and JWT became convinced of both the computational and statistical advantages of spectrum analysis procedures of the general form: (a) preprocessing (such as trend removal), (b) calculating mean lagged products, (c) cosine transformation and (d) local smoothing of these raw spectrum estimates. Specifically, taking  $R_0, R_1, \dots, R_m$  as estimated autocovariance values, they wrote

$$(1) \quad L_h = \frac{R_0}{m} + \frac{2}{m} \sum_1^{m-1} \cos \frac{ph\pi}{m} R_p + \frac{R_m}{m}$$

for  $h = 1, 2, \dots, m$  with similar definitions for  $L_0, L_m$ . Now the proposed estimate at frequency  $h\pi/m$  is

$$(2) \quad U_h = 0.23L_{h-1} + 0.54L_h + 0.23L_{h+1}.$$

It was found that the weights (0.23, 0.54, 0.23) could provide a less biased estimate than the one based on  $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$  first used. These steps produce a so-called indirect estimate. This estimate became quite standard for the next 15 years. For working scientists the development was a breakthrough, providing explicit computational steps and substantial discussion of practical issues such as the uncertainty of the estimate.

Following an invitation to Tukey generated by Shannon the work was presented at the Symposium on Applications of Autocorrelation Analysis 13–14 June 1949. The principal writings on the topic began in 1949, “The sampling theory of power spectrum estimates” [22], and “Measuring noise color” [21], with Hamming junior author. The second paper, dated 1 December 1949, remained unpublished until *The Collected Works of John W. Tukey* appeared. In the first paper it is remarked that it is hoped that the accompanying derivations will appear in *Biometrika*, while the second paper is titled part 1. Neither a *Biometrika* paper nor a part 2 is known. The papers themselves do not appear to have been easily accessible, and in many cases JWT did not even refer to [21], perhaps for proprietary reasons.

The symposium paper [22] contains substantial discussion of possible kernel functions and it derives variances of the consequent spectrum estimates. The

expected value of estimates of the type considered, at frequency  $h/2m$  cycles/unit time, is shown to have the ubiquitous form

$$\int_0^{\infty} u_m\left(\omega - \frac{h\pi}{m}\right) p(\omega) d\omega,$$

where  $p(\cdot)$  denotes the spectral density of interest and  $u_m(\cdot)$  is a window or kernel function. The kernel is sketched in [22] for one case. Kernel estimates of an unknown function were thus at hand for the statistical community, and their properties were being investigated. Among other things the distributions developed provided methods for use in the planning of experiments for the collection of time series data.

Among other crucial matters discussed in [21] are: aliasing (the memorandum presents the classic diagram), chance fluctuations, the effect of sampling interval, Gibbs phenomenon, the frequency window, quantization error. There is further substantial concern and discussion of computational complexity. Tukey often argued the advantages of the power spectrum over the autocorrelation function. Such an argument is needed because the two are mathematically equivalent in the population. In this paper he says:

...the power spectrum seems inevitably to have a simpler... interpretation. Much practical advice is provided for practitioners, with an emphasis on the researcher's need for indications of sampling variability to restrain his optimism about the reality of apparent results, but ... have a chance to discover something unexpected ...

An effect of his efforts was that both autocorrelation and periodogram analysis were reduced to tools for very specific circumstances.

In the penultimate section of [22] Tukey initiates his argument (continued throughout his career) against the use of few-parameter models for the power spectrum unless there is well-grounded theory for the model. Today that argument seems long won, if for no other reason than the common occurrence of large data sets.

Unfortunately, the mass of the population of practitioners had to wait for the appearance of the papers of Blackman and Tukey [26, 27] and they found these papers hard going. In fact comments of readers motivated Blackman to provide further discussion, in Chapter 10 of [2].

Perhaps unsurprisingly it was an engineer who led JWT to focus attention on the spectrum, as opposed to the autocorrelation function. Engineers had been using cosines and transfer functions for studying systems.

Re concerns of priority we note that JWT often referred to the independent developments of Bartlett [1].

**3. The 1950s.** In this decade JWT wrote a number of papers "selling" spectrum analysis with its many details and provisos. Some of this material surely contributed to the oftmade remark that the subject was an art, not a science. (Again

with the appearance of large data sets, the remark became less appropriate, e.g., bandwidth parameters could be estimated.)

The 1956 tutorial paper, “Power spectral methods of analysis and their application to problems in airplane dynamics” [25], provided a host of examples of the usefulness of spectrum estimates and laid out details of the computations. The important technique of prewhitening was introduced and motivated. This technique made the issue of which window to employ in forming a spectrum estimate a minor one and ended a research effort of Tukey’s. We quote from [32]:

During the early ’50s I spent considerable effort on a variety of ways to improve windows. The results have never been published because it turned out, as will shortly be explained, to be easier to avoid the necessity for their use.

In prewhitening one preprocesses the series to make it more like a sequence of independent identically distributed values, that is, to make the spectrum more nearly constant. The kernel estimate is then less biased. On some occasions one may recolor the prewhitened estimate, in others it may not be felt necessary.

One anecdote, [13], related to this work is the following. The Cornell Aeronautical Laboratory was studying turbulence in the free air by flying a highly instrumented airplane for a few hundred miles over Lake Erie. (The answers were to be relevant to naval antiaircraft fire.) The researchers learned that the visual “averages” obtained in reading strip charts caused substantial problems with cross-spectral analysis. In analytic terms it was found to be preferable to use the values  $x_t$ ,  $t = 0, 1, 2, \dots$ , in the analysis rather than the values  $\int_t^{t+1} x_s ds$ ,  $t = 0, 1, 2, \dots$ .

Another anecdote concerns Hans Panofsky’s use of John von Neumann’s new computer (in whose design JWT had a part), to analyze, in spectrum terms, wind velocity vectors at various heights on the Brookhaven tower. The cospectrum (the real part of the cross-spectrum) was given a physical interpretation as the frequency analysis of the Reynolds stress. The meaningfulness of the quadrature spectrum (the imaginary part) required some “selling.” A result of the research was the discovery of “eddies” rolling along the ground and steadily increasing in size from early morning to late afternoon. JWT liked to provide such “enlightening examples” when speaking and writing about spectrum analysis.

Neither the paper [22] nor the memorandum [21] had discussed the spectrum analysis of bivariate time series, that is cross-spectral analysis. However, Wadsworth, Robinson, Bryan and Hurley [17] indicate that they had no difficulty extending the Tukey–Hamming method and carried out an empirical analysis. There was a parallel need for theory. In a Ph.D. thesis written under Tukey’s supervision, Goodman [6] extended the results of [21, 22] to bivariate stationary series. For example, Goodman derived needed approximations to distributions, such as one for the coherence estimate.

Prewhitening was substantially elaborated upon in the book *The Measurement of Power Spectra* [28]. It was with this book that many engineers and scientists

learned how to compute estimates of power spectra, how to interpret the results of those computations and about difficulties that could arise in practice. The work was intended for four groups of readers: communications engineers, "digital computermen," statisticians and data users. It is structured in a novel fashion; mainly formal sections are paralleled exactly by mainly verbal sections. At the outset a "wonderous result" of Walter Munk's is mentioned. Quoting from a letter of Munk's:

... we were able to discover in the general wave record a very weak low-frequency peak which would surely have escaped our attention without spectral analysis. This peak, it turns out, is almost certainly due to a swell from the Indian Ocean, 10,000 miles distant. Physical dimensions are: 1 mm high, a kilometer long.

The discovery was made by estimating spectra for a succession of segments of the time series and then noticing a small peak sliding along in frequency.

There is extensive discussion of both the continuous- and discrete-time cases. As mentioned above, the importance of prewhitening is stressed, as is the related need for detrending the series prior to computing the spectrum estimate. There is substantial discussion of planning considerations prior to data collection. The book contained a useful Glossary of Terms. The computational schemes presented matched the facilities that were becoming widely available in the United States and Great Britain. Although the preferred computational scheme today is quite different, the Blackman-Tukey advice re planning and interpretation remains apropos.

Next, historically, the papers "The estimation of (power) spectra and related quantities" [29] and "An introduction to the measurement of spectra" [31] extended the introduction of the ideas of spectrum analysis. The first paper was prepared for a mathematical audience and is formal in tone, while the second is meant for a statistical audience. One novelty in these papers is the inclusion of extensive lists of papers applying the methods (the technique had become understood) and remarks like:

We here try to sketch an attitude toward the analysis of data which demonstrably works . . . .  
It may be that digital calculation will become the standard method.

Cross-spectrum analysis is presented in more detail with the remark that the idea is originally Wiener's. One sentence in [29] suggests that John had some pleasant surprises as the years passed:

Few of us expect to ever see a man who has analyzed, or even handled, a sequence of a million numerical values, . . .

The final two papers for this decade are of different character. The second is discussed here out of chronological order. John was involved in the nuclear test ban treaty negotiations taking place in Geneva, Switzerland in the late 1950s. The paper "Equalization and pulse shaping techniques applied to the determination of

initial sense of Rayleigh waves” [30] addresses an important technical problem. If one can measure, at a well-distributed set of observatories, the initial sense of the motion (“up” or “down”) of an arriving wave, then one has an estimate of the event’s radiation pattern. The pattern of an explosion differs from those of an earthquake, and one has a way to infer whether the event was an earthquake or an explosion. The series involved are nonstationary, and a moving-window crosscorrelation analysis is offered. The operation of “tapering” (i.e., multiplying the time series record of interest by a function that is near zero at the beginning and at the end of the record and near one in between) is introduced here. The idea is referred to as a data window in [28].

3.1. *Polyspectra.* Second-order spectrum analysis is particularly appropriate for Gaussian series and linear operations. The 1953 paper “The spectral representation and transformation properties of the higher moments of a stationary time series” [24] breaks away from the second-order circumstance. Tukey begins an extension of spectrum analysis to the higher-order case, laying out the definition and properties of the so-called bispectrum. As suggested, this quantity is useful for non-Gaussian and nonlinear situations. The integrated power spectrum,  $P(\omega)$ , at frequency  $\omega$  of a zero-mean stationary time series may be defined via the Fourier transform of a second-order moment function as in

$$\text{ave}\{x_t x_{t+h}\} = \int e^{i\omega h} dP(\omega)$$

where  $h$  is the time lag. Analogously, the integrated bispectrum  $P(\omega, \nu)$  at bifrequency  $(\omega, \nu)$  may be defined as a two-dimensional Fourier transform of a third-order moment function, specifically via

$$\text{ave}\{x_t x_{t+h} x_{t+k}\} = \iint e^{i(\omega h + \nu k)} dP(\omega, \nu).$$

While this definition directly suggests the possible use of bispectra in handling non-Gaussian data, it perhaps misses their essential use in handling and searching for nonlinear phenomena. In this early paper Tukey introduces the terminology and the algebra of bispectral analysis. He begins the physical interpretation as well. It seems to have taken ten years before any bispectral estimates were actually computed. In some of these developments JWT would possibly have built on the work of Blanc-Lapierre. The paper lists an important (still!) open problem: develop a method of constructing a series with prespecified higher-order spectra or polyspectra. JWT indicates an approximate method for the second- plus third-order case.

The topic of bispectral analysis is returned to in a number of his later writings including: [31, 35, 36, 40, 41, 44, 58]. There now exists a moderately large literature concerned with bispectral analysis. Luckily one can search the word “bispectrum” on the Internet easily. (This is one important advantage of JWT’s providing unusual names for concepts in order to reduce the likelihood of confusions.)



**4. The early 1960s.** At the start of this decade JWT wrote a watershed paper "Discussion, emphasizing the connection between analysis of variance and spectrum analysis" [32]. The strength of connection of the topics that he saw is illustrated by the remark:

... spectrum analysis of a single time series is just a branch of variance component analysis, ...

On reflection JWT's earlier papers could be viewed as the work of a combination applied mathematician-communications engineer. The overriding consideration of the paper [32], though, is that of the statistician doing inference both in the classical and in the post-modern sense. Some new techniques are introduced (complex demodulation, the use of a pair of bounding estimates, the problems of discrimination, and canonical correlation). Some formulas are set down. Some computational considerations are mentioned. However dominant themes are ones of the philosophy of data analysis and of the contributions and limitations of the statistical analysis of time series. This is a paper that bears much rereading.

Numerous sentences in the paper stand out:

... , when examined under a microscope, no known phenomenon is precisely periodic. There is no mixing up of frequency variance components. This is simultaneously true for all black boxes, and is the basic reason why the user, be he physicist, economist, or epidemiologist, almost invariably finds frequency variance components the most satisfactory choice for any time series problem which should be treated in terms of variance components.

... regression goes on separately at each frequency ...

*It may pay not to try to describe in the analysis the complexities that are really present in the situation.*

... I have yet to meet anyone experienced in the analysis of time series data ... who is over-concerned with stationarity.

... it does not pay to try to estimate too much detail, even if the detail is really there.

... it is not uncommon for spectrum estimates based upon different experimental repetitions to differ more than might be expected from their internal behavior.

... at least all the things that are permissible will happen.

The purpose of asymptotic theory in statistics is simple: to provide usable approximations before passage to the limit.

Time series analysis follows its usual pattern, like most statistical areas, only more so.

Most data analysis is going to be done by the unsophisticated.

... try, look, and try something a little different as the typical pattern of data analysis.

Of course such remarks might have been anticipated because this is the period when JWT was working on his landmark paper "The future of data analysis" [15]. The steps of the approach may be found in the box and arrow diagram of Figure 1, taken from *The Collected Works of John W. Tukey I*. This figure bears much reflection. It surely gives some important clues as to how Tukey's mind worked.

In the paper JWT presents the details of complex demodulation, essentially a way to apply a narrow bandpass filter to a time series and to approximately obtain the corresponding Hilbert transform at the same time. It may be used to

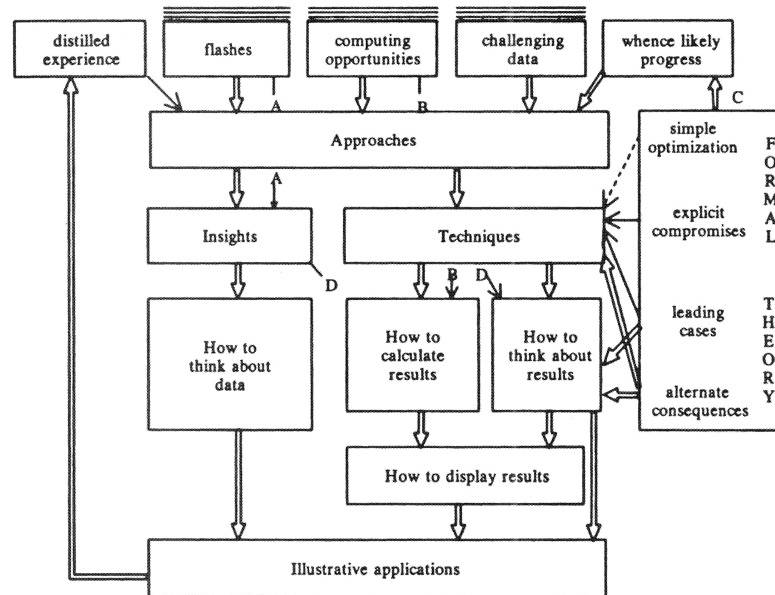


FIG. 1. JWT's structure of any branch of data analysis.

provide local estimates of spectra, of various orders, and other frequency domain parameters in the nonstationary case. This step is important because no real series is purely stationary. To end the discussion of the paper [32] consider the remark:

... I know of no large machine installation whose operations are adapted to the basic step-by-step character of most data analysis, in which most answers coming out of the machine will, after human consideration, return to the machine for further processing.

Here one finds JWT calling for the tools of modern interactive statistical computing.

The paper "Curves as parameters and touch estimation" [33] is rather unusual amongst JWT's papers. It is concerned with the estimation of continuous curves—specifically estimation of regression functions, of probability density functions or of power spectra. The field considering such problems is now called nonparametric estimation. The difficulty is that of trying to estimate an infinite number of parameters, when only a finite amount of data is available. The approach is one of "using the data to determine the size, shape, and location of certain sets . . . , which are regarded as trying to touch (that is, meet, intersect, overlap), not enclose, the curve." In the case of power spectrum analysis one uses two estimates via a pair of window functions. Quoting from [35]:

... choose two windows, one with all positive side lobes and another with all negative side lobes. If both are routinely used in calculation, the difference of the corresponding spectral estimates makes clear where we ought to recalculate, usually after prewhitening or elimination, if we want to avoid difficulty from lobes.

The problem of constructing confidence touch estimates is also addressed.

Tukey wrote a number of papers placing the data analysis of time series into perspective with current research in the physical sciences, in statistics, and in computing. The paper "What can data analysis and statistics offer today?" [35] constituted discussion of papers presented at a conference on oceanography. JWT had remarked that Bill Pierson and his concern with (one- and two-dimensional) spectra of the sea surface had been a stimulus in his work. The paper is notable for its heuristic description and resolution of problems and issues arising in spectrum analysis, including window choice, bispectra and nonstationarity. Tukey's philosophical approach to certain problems arising is well illustrated:

I, for one, place meaningfulness and understandability before efficiency or information content.

Many of the time series techniques that John Tukey has created, have been motivated by problems in geophysics. In the other direction, through his forceful presentation of statistical procedures, John Tukey has motivated geophysicists to novel applications. The paper, "Data analysis and the frontiers of geophysics" [38] is an address presented at the dedication of the La Jolla Laboratories of the Institute of Geophysics and Planetary Geophysics of the University of California. It shows John in the latter role. The topics emphasized include: spectrum analysis, reexpression and robust/resistance (that is, techniques that continue to work well even when there are substantial departures from the assumptions used in their construction.) On page 1286 one finds what might be taken as a definition of spectrum analysis:

... the science and art of frequency analysis . . . .

Also in the paper the advantage of cross-spectrum analysis, over the analysis of a single series, is summed up via:

measuring relationship is almost always more rewarding than measuring relative contribution to variability.

The paper contains substantial discussion of "why the techniques of spectrum analysis are so useful."

The paper "Uses of numerical spectrum analysis in geophysics" [40] is full of striking examples. The paper was prepared for statisticians, but it is of equal importance for geophysicists. It surveys the methods of spectrum analysis including complex demodulation, cepstrum analysis and bispectrum analysis. It addresses the issue of why the frequency-side approach worked in the examples described. There is a section on the Fast Fourier Transform. Concerning learning about spectrum analysis the advice is given to:

read any two different accounts.

Advice is also given concerning the analysis of array data, namely:

It is good to choose coefficients to magnify the signal, but is far better to choose them to cancel out the noise.

The paper concludes with an extensive list of applied papers, classified, and in many cases annotated. The paper was presented at a meeting of the International Statistical Institute.

A further paper in this series is “Spectrum analysis of geophysical data” [41], joint with R. A. Haubrich. It illustrates that spectrum analysis has played an important role in addressing many geophysical problems with examples from seismology, oceanography and astronomy. The uses of cross-spectrum and bispectrum analysis are illustrated with examples. That spectrum analysis can provide *clues to source mechanisms* is shown. All told, this paper provides a gentle introduction to many of the complexities of spectrum analysis.

The English statistician Healy came to Bell Labs, Murray Hill in the early sixties with the intention of learning how statistics was used in an industrial context. Thanks to the International Nuclear Test Ban negotiations, however, he, Bruce Bogert and JWT worked on seismology with emphasis on discriminating between explosions and earthquakes. Out of this came the paper “The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking” [34]. The work was presented at a time series conference at Brown University. The audience must have been absolutely amazed. What appeared for the first time were the idea, the language and the methodology of cepstrum analysis—novel concepts and novel words. The research was motivated by the problem of estimating the depth of the source of a seismic signal. The insight was the recognition that in some circumstances a seismogram would consist of an echo of a signal superposed on the signal itself (with the echo delay proportional to the depth of the source). The empirical approach was to estimate the power spectrum of a smoothed power spectrum estimate of the seismogram. A variant of complex demodulation, “saphe-cracking,” was also proposed. The approach found early use in the problem of pitch-detection—the identification of the time spacing between repetitions of vocal-chord behavior. There is some later material on cepstrum analysis in [58].

## PART II: THE “DIRECT” YEARS

**5. The later 1960s.** In the spring of 1963 John Tukey presented a graduate course, *Mathematics 596*: “An Introduction to the Frequency Analysis of Time Series,” at Princeton. Notes were taken and finally appeared in print in 1984 [36]. The people in the course were principally graduate students in Mathematics, specializing in statistics, and some students from other departments. The course involved a mixture of techniques, philosophy, examples and terminology. There was much interaction with the audience. The notes begin with a discussion of the roles of models and modeling. One notable statement JWT quoted from [11] was:

We must be prepared to use many models, . . .

to which he added:

Models should not be true but it is important that they be applicable.

The material is remarkable for including a prescription for a Fast Fourier Transform (FFT) algorithm (Topic T, Section 3, later formalized in [37]) and for introducing the operation of taking running medians (Topic U, Section 6). One also finds substantial discussions of: the Hilbert transform, complex demodulation, nonlinear operations (particularly polynomial), aliasing, decimation and spectral representation. There are allusions to constructing robust/resistant statistics. The material foreshadowed an abrupt change in the practice of time series analysis. JWT showed that when  $N = GH$  the computation of the empirical Fourier transform

$$\sum_{t=1}^N y_t \exp\{-i2\pi jt/N\}, \quad j = 0, \dots, N-1,$$

requires only  $(H + 2 + G)GH$  multiplications. He further remarked that for  $N = 4^k$  one needs fewer than  $2N + N \log_2 N$ .

It seems that when one asks scientists and engineers, from no matter what nation, if they have ever heard the name Tukey, they immediately refer to the FFT and perhaps mention the Cooley–Tukey paper “An algorithm for the machine calculation of complex Fourier series” [37]. Before 1965 spectrum estimation had focussed on the indirect method [see expressions (1) and (2)], which avoided the use of Fourier transforms of the data themselves. Following the appearance of [37] signal processing very quickly switched from analog to digital in many important cases. The paper is a citation classic, and the method provided what has been named one of the top 10 algorithms of the 20th century [*SIAM News* (2000) 33(4)]. FORTRAN programs quickly became available for the Fourier transform of the data themselves. This was followed by a quick shift to the so-called direct method of spectrum estimation, (involving smoothing the mod-squared of the FFT of the data) and using the FFT to compute autocovariance estimates and filtered values. The Fourier transform had long been known to have convenient mathematical and statistical properties. This paper made clear that it had convenient computational ones as well. It has since turned out that the ideas had already been available for a substantial time (see [47]), they just had not been noticed much.

JWT liked geometrical arguments. The paper [47] has one. He asks the question: “Why should a one-dimensional DFT become almost a two-dimensional DFT?” His heuristic was the following: suppose one has  $T = r_1 r_2$  observations. Consider a segment of length  $T$  of a cosine wave. Stack the segment in  $r_2$  segments of length  $r_1$  underneath one another. The one-dimension cosinusoid becomes a tilted “plane” cosinusoid. As such, it differs from a product of coordinate-wise waves only by a set of phase factors.

There is a regrettable aspect to the story of the Cooley–Tukey FFT. On numerous occasions JWT made remarks of the type [47]:

Had I thought, I would surely have recalled that Gordon T. Sande ... had purely independently found an algorithm to reach the same effect.

(*Collected Works II*, xlv)

Details of the complementary forms of the general algorithm—in the equal radix case—were developed independently and simultaneously by Gordon Sande (unpublished) and by J. W. Cooley.

Stockham (1966) and Sande (unpublished) independently discovered the use of FFT's to calculate linear convolutions.

He clearly felt that he had contributed to Gordon Sande's not receiving an appropriate share of the credit associated with the FFT.

Another side of Tukey's FFT work was his and his collaborators' finding novel applications of the algorithm(s). The 1966 paper "Fourier methods in the frequency analysis of data" [42] was prepared for a mathematical audience (the Mathematical Association of America) and hence emphasizes particular topics, for example, coordinate systems, transformations, Fourier methods. There is much description, motivation and advice for the general reader, for example, why frequency analysis is important, why frequency analysis is impossible, why frequency analysis *IS* possible, the Fast Fourier Transform and the steps of a practical frequency analysis.

In a special issue of *IEEE Trans. Audio Electroacoustics* the important paper "Modern techniques of power spectrum estimation" [43] written with Bingham and Godfrey, appeared. It laid out the computations of spectrum estimation and complex demodulation via a Fast Fourier Transform. Many of the necessary details of the computations are presented. A variety of concepts, such as signal, noise, spectrum, are given pertinent definitions, for example:

Spectrum as a general concept: An expression of the contribution of frequency ranges to the mean-square size, or to the variance, of a single time function or of an ensemble of time functions.

Essential distinctions are made between signals, where:

a repetition would be an exact copy

and noise, where:

a repetition would have only statistical characteristics in common with the original.

This distinction was further explored by Brillinger and Tukey [58].

There was more statistical work in "An introduction to the calculations of numerical spectrum analysis" [44]. The paper provided historical and expository discussion of the statistical and computational aspects of spectrum analysis with a broad variety of references to geophysical and engineering applications. The indirect method of estimating a spectrum is reviewed. The importance of prewhitening and tapering is emphasized. It is commented that spectrum analysis

is an iterative procedure. Window carpentry is discussed at some length, and one finds the remarks:

We still need inner and outer windows. We shall need them or some modification of them for as long a time as we can dream about.

The history of the move to computation of spectrum estimates via a Fast Fourier Transform algorithm is presented with mention of the work of Danielson and Lanczos, Good, Welch and Sande. In that connection, the paper ends with the remark:

We have gained arithmetic speed and an ability to make more subtle analysis routine, but the name of the game is still the same.

The dangers of looking at autocorrelations are referred to.

**6. The 1970s.** JWT opened this decade by presenting the (Arthur William) Scott Lectures [46, 47], at Cavendish Laboratory, Cambridge University. In deference to the British audience he talks about “goods-wagons” instead of “boxcars.” The lectures are an important event. In the first five years they were given by Bohr (1930), Langmuir (1931), Debye (1932), Geiger (1933) and Heisenberg (1934). In his lectures JWT appears intent upon educating the audience to exploratory data analysis, to the computations of spectrum analysis, to the FFT and to the FFT's history, amongst other things. One reads:

The first task of the analyst of data is quantitative detective work, . . . . His second task is to extract as clearly as possible what the data says about certain specified parameters. His later tasks are to assess the contributions to the uncertainties of these statements from all causes, systematic or whatever . . . . Often the purpose of good analysis is not so much to do well in catching what you want but rather to do well . . . in rejecting what you don't want.

The lecture notes are a tour de force for physical scientists, with many examples and sections: Principles, Time and Frequency, More Parameters than Data, Key Details, Computation, Fast Algorithms, Fast Multiplications, Fast Fourier Transforms, Cosine Polynomials and Complex-Demod(ulation). The attitude that an adaptive approach to learning about frequencies is vital is propounded. There is substantial comparative discussion of analog and digital analysis techniques. The general principle is expounded that how one gets an estimate of the spectrum near one frequency needs to take into account what is known about the activity at other frequencies. This is followed up by “The adaptive approach is vital to learning about frequencies.” There is this summary:

The finiteness of real data forces us to face three things in any attempt to study the frequencies in something given by numbers: 1) aliasing, 2) the need for sensible data windows, 3) the inevitability of frequency leakage.

JWT's interactions with practitioners were referred to earlier. A conference was held in 1977 to assess the field of event-related brain potential (ERP) research. The

structure of the conference involved earlier preparation of state-of-the-art reports, each circulated to a senior scholarly critic. The role of the critic was to assess a particular report and to formally discuss it at the conference. JWT's discussion, titled "A data analyst's comments on a variety of points and issues" [49] was of the paper "Measurement of event-related potentials" prepared by John, Ruchkin and Vidal. Generating event-related potentials (ERP) is a traditional means of studying the nervous system. It involves applying sensor stimuli (the events) to a subject and at the same time examining the electroencephalogram. The researchers in the area have employed a broad range of statistical techniques in the analysis of their data, but statisticians have not been too much concerned with the field. JWT's comments covered a broad range of issues. He suggests the use of cross-validation. He believes that regression should "be more widely used in ERP work." He knows:

of no field where data analysis has progressed even moderately far toward sophistication where only one expression of the data has been found worthy of being looked at.

He suggests that:

the overriding importance of noise "rejection" over signal "enhancement" deserves special attention.

once again. He refers to the problem of multiplicity in testing. He proposes the use of iterative reweighting of curves to obtain a robust/resistant estimate of the ERP signal. He makes some suggestions re design. As in many other cases he prepared quite a Referee's Report!

The paper "Nonlinear (nonsuperposable) methods for smoothing data" [48] was presented at a 1974 EASCOM conference. It has the aggressive remark:

To try to tell modern engineers that optimization is not the wave of the near future may perhaps be a little unpalatable.

The focus is median filtering, that is, the operation of replacing the middle observation, of a sliding segment (or window) of a time series, by the median value of the observations appearing in the window. Once the idea was understood and programs written, the technique of median filtering found many immediate applications in signal processing. The idea was actually introduced in 1963 in [36], but those notes remained unpublished until *Collected Works I* appeared. Median filters have some important characteristics: they reduce spiky noise, and they preserve jump discontinuities (edges). "Salt and pepper" noise is taken right out of images, and straight objects are preserved. The technique was further discussed in Chapters 7 and 16 of the EDA book [16]. The tools of "re-roughing" and "twicing" may also be noted here.

The paper "When should which spectrum approach be used?" [53] was presented in 1976 at a conference on time series analysis and forecasting. The conclusion of the paper is:

Spectrum analysis is to be entered upon advisedly and with care.



The paper “Can we predict where ‘time series’ should go next?” [54] constitutes the keynote address at the Institute of Mathematical Statistics Special Topics Meeting on Time Series Analysis held in Ames, Iowa in 1978. JWT’s answer seems to be “No, but...” JWT does set out “a large array of tasks whose satisfactory completion would be of value.” The paper is further notable for laying out what he sees as the three main branches of time series analysis: (i) “where the spectrum itself is of interest,” (ii) “an ‘equivalent record’ would differ from the one before us only by having a different sequence of ‘measurement noise,’” (iii) “when series are not long, models are not subject-matter given, and equivalent records do not look alike.” It further provides a short catalog of aims (discovery of phenomena, modeling, preparation for further inquiry, reaching conclusions, assessment of predictability, description of variability), and the definition of time series analysis referred to earlier. A variety of robust/resistant techniques are suggested with iterative and graphical aspects of the process emphasized. In connection with naive frequency analysis we find the wonderful remark:

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

In counterpoint to the emphasis in [32] one finds the remark:

regression is always likely to be more helpful than variance components.

This leads to an extensive discussion of cross-spectral analysis. Robustness/resistance and leakage are concerns. There is extensive discussion of work of Thomson on spectral analysis of waveguide roughness data [13]. (A waveguide is a physical device within which an electromagnetic signal moves.) It was found that one dust speck could flatten the low portion of a naive spectrum drowning out important information. There are also discussions of cepstrum analysis and of complex demodulation.

**7. The 1980s.** The 1982 paper “Spectrum analysis in the presence of noise: Some issues and examples” [58] written by JWT and myself, was invited by the *Proceedings of the IEEE* and rejected. It was also rejected by *J. Time Series Analysis*. There was a lot of material in the paper. JWT’s comment over the first rejection was that the writing had taken away a lot of his time that he could have spent elsewhere and that I shouldn’t worry, for the paper could appear in *The Collected Works of John W. Tukey*. One of the copies of the submission that came back from a referee had a swear word written on it. John’s Bell Labs secretary was quite concerned that John not read the word.

A talk, “Styles of spectrum analysis,” was presented at a 1982 conference to honor the 65th birthday of the renowned oceanographer Walter H. Munk. JWT starts his paper [55] with:

Walter Munk may well be the most effective practitioner of spectrum analysis the world has seen.

and in the Foreword to *The Collected Works of John W. Tukey I*, JWT remarks:

Through the years, my strongest source of catalysis has been Walter Munk.

Part of the paper is a description of phenomena that Munk discovered via spectrum analysis. However, the greater part of the paper is taken up with a discussion of the relative merits of “overt” and “covert” analyses of (time series) data. He says that both are needed. Drawing an analogy with the alternation of the processes of discovery and refinement characteristic of the progress of science, he suggests the following work sequence for spectrum analysis: (i) initial overt analysis, (ii) repeated covert analysis, (iii) another spell of overt analysis, and so on. He remarks:

Failure to use spectrum analysis can cost us lack of new phenomena, lack of insight, lack of gaining an understanding of where the current model seems to fail most seriously . . . .

while:

Failure to use covert spectrum analysis can cost us in efficiency.

Another concern of the paper is robust/resistant techniques. This leads to the idea that, where a covert approach is adopted, it should be robust/resistant.

JWT’s last paper on time series analysis that I know about, appeared in 1990, “Reflections” [60]. He begins by joking about reflections in a mirror (as always, looking for a physical analogy). Of looking forward to the future he says,

My feeling . . . is that our current frequency/time techniques are quite well developed . . . , so that the most difficult questions are not “how to solve it” but rather either “how to formulate it,” or how do we extend applicability to less comfortable conditions.

During this period he continued to help researchers with their problems in time series analysis and filtering.

**8. Some JWT neologisms.** Over a period of many years JWT introduced a multitude of techniques and terms that have become standard to the practice of data analysis generally and of time series specifically. Various of these have already been mentioned in this article. In particular we note: *prewhitening*, *alias*, *smoothing and decimation*, *taper*, *bispectrum*, *complex demodulation*, *cepstrum*, *saphe cracking*, *quefreny*, *alanysis*, *rahmonic*, *liftering*, *hamming*, *hanning*, *window carpentry* and *polyspectrum*.

Bruce Bogert, one of John’s collaborators in the cepstral paper [34], once described an incident in a restaurant where the paper’s authors were eating. The customer at an adjoining table eventually came over and asked what language the three were speaking.

About his unusual use of words and the creation of new ones JWT said that he was hoping to reduce the confusions that arose when words had a variety of other meanings.

**9. Some stories of JWT and time series analysis.** There are many amusing JWT stories.

1. Of the paper on cepstral analysis [34], Dick Hamming remarked to John that “from now on you will be known as J. W. Cutie.”
2. One of John’s favorite success stories of spectrum analysis concerned the free oscillations that arise in consequence of great earthquakes. When I told him that a seismologist collaborator, Bruce Bolt, and I had a novel method to estimate the parameters of the Earth and their standard errors based on free oscillations and that we would be in the papers with the results the day after the next great earthquake, John’s remark was: “What if the earthquake is in San Francisco?” (Berkeley is across the Bay from San Francisco.)
3. JWT gave a course on time series analysis my first year at Princeton. In the first or second lecture JWT used the word “spectrum.” Later in the class a student asked “What is a spectrum?” JWT replied, drawing a picture “Suppose there is an airplane and a radar . . . that is the spectrum.” The next lecture a similar thing happened. Student: “What is the spectrum?” JWT, drawing another picture, “Suppose there is a submarine sending out a sonar signal and . . .that is the spectrum.” The student never came back to the course. He did go on to a highly successful career, often asking questions of this same type.
4. When sitting next to me once at a seminar, JWT passed over a piece of paper on which he had written:

Measure (?) of “peakiness” of periodogram: 1) take cepstrum (based on log periodogram), 2) average in octaves, laid  $1/2$  to the weather.

The idea had probably occurred to him just then. This idea has not been further developed so far as I know.

5. When JWT read one researcher’s description of the bispectrum, he said it occurred “At page 1426, which the journal has appropriately numbered 4126 . . .” [40].

**10. Discussion.** To begin, we note that some of the other volumes of the *Collected Works* contain papers of time series interest. In particular, there are [15] “The future of data analysis” and [9] “An overview of techniques of data analysis, emphasizing its exploratory aspects.” So too, much of the material of the time series volumes has much relevance to other topics, for example the material on nonparametric density and regression estimation in [33] “Curves as parameters and touch estimation” and Figure 1. In particular, virtually all of the statistical philosophy applies elsewhere. The time series volumes of *Collected Works* contain some noteworthy previously unpublished material. We mention: [21], “Measuring noise color”; [36], “An introduction to the frequency analysis of time series”; and [48], “Nonlinear (nonsuperposable) methods for smoothing data.”

JWT’s work has fueled the advance of time series analysis for over fifty years. The work is continually described as seminal, breakthrough, insightful, essential,

and with a host of other superlatives. It directed the greater part of both the theory and practice of time series analysis for many years. One can speculate on why his work has been so dominant. There is no doubt that his writings have played an important role. As any reader can now see, they are full of constructive procedures, practical advice, necessary warnings, heuristics and physical motivation. Equally influential, however, has been John's involving himself with working substantive scientists through seminars, conferences, committee meetings, and in organized and chance encounters. The literature of applied time series analysis contains numerous acknowledgments of his suggestions. For example: "Fortunately Tukey took an interest in the seismic project and conveyed his research ideas by mail." (E. Robinson [12].) "The Project . . . was also particularly fortunate in being advised from its earliest days by Professor John Tukey who made available to us many of his unpublished methods of analysis." (C. Granger [7].)

Re other techniques for time series analysis one should note how flexible and important the state space approach is proving nowadays. I don't remember that JWT ever commented on it.

JWT often referred to his mentors C. Winsor and E. Anderson as stimulants to his work in other branches of statistics. He does not appear to have had a mentor for his work in time series analysis, but it is hard to know.

Anyone who has been involved with John has indeed been fortunate. They have seen his rapid domination of the situation at hand, his extensive knowledge of pertinent physical background, his leaps in unimagined directions to concrete procedures, his vocabulary and his humor.

John Tukey leaves the scientific world with a legacy. It consists of methods, words, warnings, heuristics and discoveries. His name will live on.

**Acknowledgments.** I particularly thank F. R. Anscombe, R. Gnanadesikan, M. D. Godfrey, D. Hoaglin, D. Martin, E. Parzen and the referees for their suggestions and anecdotes. I thank John himself for introducing me to the topic of time series analysis in spectacular fashion, through his Princeton course in 1959–60 and through his having me analyze earthquake and explosion records when I was a graduate student. I thank him for all the enjoyment he brought into my life over a forty-year period.

Of course there were other researchers such as Bartlett, Grenander, Hannan, Jenkins, Parzen, Priestley and Rosenblatt, who were making important contributions to the field of spectrum analysis during much of the same period. I did not see how to separate out their work usefully when writing this piece.

#### APPENDIX A

The letter reproduced here may be found in the John W. Tukey Archive at the American Philosophical Society.

June 20, 1942

Dr. J.W. Tukey  
Princeton University  
Princeton, New Jersey

Dear Tukey:

I looked over your report and intended to answer it soon but you know how busy we are preparing for our big inspection.

The chief thing that I want to say is that I do not believe that the correction by a Gaussian factor after the autocorrelation coefficient has been taken is a good way. To take the autocorrelation coefficient one asks primarily for a quantity whose Fourier transform is essentially positive. In order to do this, one must use a cesaro waving factor in getting the average of a finite integral. This introduces a bad behavior at 0 as well as the extreme frequency, and this bad behavior in the center is not touched by any weighting factor. I therefore much prefer our method of weighting the data with the Gaussian factor before obtaining the autocorrelation coefficient.

I shall write to you more in detail later. We enjoyed your visit very much and hope to keep in touch with you. We have had very good success on our own show.

Very sincerely yours,  
Norbert Wiener

## APPENDIX B

### Doctoral theses on time series that JWT supervised.

GOODMAN, N. R. (1957). On the joint estimation of the spectra, cospectrum and quadrature spectrum of a two-dimensional stationary Gaussian process. Ph.D. dissertation, Princeton University.

VELLEMAN, P. F. (1975). Robust nonlinear data smoothers—theory, definitions and applications. Ph.D. dissertation, Princeton University.

SCHWARZSCHILD, M. (1979). New observation-outlier-resistant methods of spectrum estimation. Ph.D. dissertation, Princeton University.

HURVICH, C. M. (1985). A unified approach to spectrum estimation: Objective choice and generalized spectral windows. Ph.D. dissertation, Princeton University.

## REFERENCES

(CWJWT refers to *The Collected Works of John W. Tukey*)

- [1] BARTLETT, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika* **37** 1–16.
- [2] BLACKMAN, R. B. (1965). *Linear Data-Smoothing and Prediction in Theory and Practice*. Addison-Wesley, Reading, MA.
- [3] BLOOMFIELD, P., BRILLINGER, D. R., CLEVELAND, W. S. and TUKEY, J. W. (1979). *The Practice of Spectrum Analysis*. Short Course offered by University Associates, Princeton.
- [4] BRILLINGER, D. R. (1976). Some history of statistics in the United States. In *On the History of Statistics and Probability: Proceedings of a Symposium on the American Mathematical Heritage* (D. B. Owen, ed.) 267–280. Dekker, New York.

- [5] BRILLINGER, D. R. (2002). The life and professional contributions of John W. Tukey. *Ann. Statist.* **30** 1535–1575.
- [6] GOODMAN, N. R. (1957). On the joint estimation of the spectra, cospectrum and quadrature spectrum of a two-dimensional stationary Gaussian process. Ph.D. dissertation, Princeton Univ.
- [7] GRANGER, C. W. J. and HATANAKA, M. (1964). *Spectral Analysis of Economic Time Series*. Princeton Univ. Press.
- [8] KENDALL, M. G. and STUART, A. (1969). *The Advanced Theory of Statistics 1*. Hafner, New York.
- [9] MALLOWS, C. M. and TUKEY, J. W. (1982). An overview of the techniques of data analysis, emphasizing its exploratory aspects. In *Some Recent Advances in Statistics* (J. Tiago de Oliveira et al., eds.) 111–172. Academic Press, London. [Also in *CWJWT IV* (1986).]
- [10] NOLL, A. M. (1964). Short time spectrum and “cepstrum” techniques for vocal-pitch detection. *Journal of the Acoustical Society of America* **36** 296–302.
- [11] RASCH, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Nielsen and Lydiche, Copenhagen.
- [12] ROBINSON, E. A. (1982). A historical perspective of spectrum estimation. *Proc. IEEE* **70** 885–907.
- [13] THOMSON, D. J. (1977). Spectrum estimation techniques for characterization and development of the WT4 waveguide I, II. *Bell System Tech. J.* **56** 1769–1815, 1983–2005.
- [14] TUFTE, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.
- [15] TUKEY, J. W. (1962). The future of data analysis. *Ann. Math. Statist.* **33** 1–67.
- [16] TUKEY, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- [17] WADSWORTH, G. P., ROBINSON, E. A., BRYAN, J. G. and HURLEY, P. M. (1953). Detection of reflections on seismic records by linear operators. *Geophysics* **18** 539–586.
- [18] WIENER, N. (1930). Generalized harmonic analysis. *Acta Math.* **55** 117–258.
- [19] WIENER, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series, with Engineering Applications*. MIT Press, Cambridge, MA.

### THE TIME SERIES PAPERS

(The listings below are in approximate chronological order of research, rather than of publication.)

- [20] (1947). Linearization of solutions in supersonic flow. *Quart. Appl. Math.* **5** 361–365. [Also in *CWJWT VI* (1990) 29–34.]
- [21] TUKEY, J. W. and HAMMING, R. H. (1949). Measuring noise color. [Also in *CWJWT I* (1984) 1–127.]
- [22] (1950). The sampling theory of power spectrum estimates. In *Symposium on Applications of Autocorrelation Analysis to Physical Problems* 47–67. (NAVEXOS P-735) Office of Naval Research, Washington, DC. [Also in *CWJWT I* (1984) 129–160.]
- [23] (1952). Review of “*Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*,” by Norbert Wiener. *J. Amer. Statist. Assoc.* **47** 319–321. [Also in *CWJWT I* (1984) 161–164.]
- [24] (1953). The spectral representation and transformation properties of the higher moments of stationary time series. [Also in *CWJWT I* (1984) 165–184.]
- [25] PRESS, H. and TUKEY, J. W. (1956). Power spectral methods of analysis and their application to problems in airplane dynamics. In *Flight Test Manual, NATO Advisory Group for Aeronautical Research and Development* 1–41. [Also in *CWJWT I* (1984) 185–255.]

- [26] BLACKMAN, R. B. and TUKEY, J. W. (1958). The measurement of power spectra from the point of view of communications engineering, Part I. *Bell System Tech. J.* **37** 185–282.
- [27] BLACKMAN, R. B. and TUKEY, J. W. (1958). The measurement of power spectra from the point of view of communications engineering, Part II. *Bell System Tech. J.* **37** 485–569.
- [28] BLACKMAN, R. B. and TUKEY, J. W. (1959). *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover, New York.
- [29] (1959). The estimation of (power) spectra and related quantities. In *On Numerical Approximation* (R. E. Langer, ed.) 389–411. Univ. Wisconsin Press, Madison. [Also in *CWJWT I* (1984) 279–307.]
- [30] (1959). Equalization and pulse shaping techniques applied to the determination of initial sense of Rayleigh waves. In *The Need for Fundamental Research in Seismology* 60–129. Appendix 9. Report of the Panel on Seismic Improvement, U.S. State Department, Washington, DC. [Also in *CWJWT I* (1984) 309–357.]
- [31] (1959). An introduction to the measurement of spectra. In *Probability and Statistics: The Harald Cramér Volume* (U. Grenander, ed.) 300–330. Almqvist and Wiksell, Stockholm. [Also in *CWJWT I* (1984) 359–395.]
- [32] (1961). Discussion, emphasizing the connection between analysis of variance and spectrum analysis, *Technometrics* **3** 191–219. [Also in *CWJWT I* (1984) 397–435.]
- [33] (1961). Curves as parameters, and touch estimation. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 681–694. Univ. California Press. [Also in *CWJWT I* (1984) 437–454.]
- [34] BOGERT, B. P., HEALY, M. J. R. and TUKEY, J. W. (1963). The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe-cracking. In *Proceedings of the Symposium on Time Series Analysis* (M. Rosenblatt, ed.) 209–243. Wiley, New York. [Also in *CWJWT I* (1984) 455–493.]
- [35] (1963). What can data analysis and statistics offer today? In *Ocean Wave Spectra: Proceedings of a Conference* 347–350. Prentice-Hall, Englewood Cliffs, NJ. [Also in *CWJWT I* (1984) 495–502.]
- [36] (1963). Mathematics 596: An introduction to the frequency analysis of time series. [Also in *CWJWT I* (1984) 503–650.]
- [37] COOLEY, J. W. and TUKEY, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comp.* **19** 297–301. [Also in *CWJWT II* (1985) 651–658.]
- [38] (1965). Data analysis and the frontiers of geophysics. *Science* **148** 1283–1289. [Also in *CWJWT II* (1985) 659–675.]
- [39] (1966). A practicing statistician looks at the transactions. *IEEE Trans. Inform. Theory* **IT-12** 87–91.
- [40] (1966). Uses of numerical spectrum analysis in geophysics. *Bull. Internat. Statist. Inst.* **41** 267–307. [Also in *CWJWT II* (1985) 677–738.]
- [41] HAUBRICH, R. A. and TUKEY, J. W. (1966). Spectrum analysis of geophysical data. In *Proceedings of the IBM Scientific Computing Symposium on Environmental Sciences* 115–128. IBM, Armonk, NY. [Also in *CWJWT II* (1985) 739–754.]
- [42] BINGHAM, C. and TUKEY, J. W. (1966). Fourier methods in the frequency analysis of data. [Also in *CWJWT II* (1985) 755–780.]
- [43] BINGHAM, C., GODFREY, M. D. and TUKEY, J. W. (1967). Modern techniques of power spectrum estimation. *IEEE Transactions on Audio and Electroacoustics* **AU-15** 56–66. [Also in *CWJWT II* (1985) 781–810.]
- [44] (1968). An introduction to the calculations of numerical spectrum analysis. In *Spectral Analysis of Time Series* (B. Harris, ed.) 25–46. Wiley, New York. [Also in *CWJWT II* (1985) 811–835.]

- [45] (1968). Proceedings of the workshop: Practical applications of the frequency approach to EEG analysis. In *Advances in EEG Analysis* Suppl. 27 to *Electroencephalography and Clinical Neurophysiology* 10–11. North-Holland, Amsterdam.
- [46] (1970). First 1970 Scott lecture. [Also in *CWJWT II* (1985) 857–884.]
- [47] (1970). Second 1970 Scott lecture. [Also in *CWJWT II* (1985) 885–914.]
- [48] (1974). Nonlinear (nonsuperposable) methods for smoothing data. [Also in *CWJWT II* (1985) 837–855.]
- [49] (1978). A data analyst's comments on a variety of points and issues. In *Event-Related Brain Potentials in Man* (E. Callaway, P. Tueting and S. H. Koslow, eds.) 139–154. Academic Press, New York. [Also in *CWJWT II* (1985) 915–934.]
- [50] (1978). Comment on "A new approach to ARMA modeling," by H. L. Gray et al. *Comm. Statist. B—Simulation Comput.* 7 79–84. [Also in *CWJWT II* (1985) 935–937.]
- [51] (1978). Discussion of "Evidence for possible acute health effects of ambient air pollution from time series analysis: Methodological questions and some new results based on New York City daily mortality, 1963–1976," by H. Schimmel. *Bulletin of the New York Academy of Medicine* 54 1111–1112.
- [52] (1978). Comments on "Seasonality: Causation, interpretation, and implications," by C. W. J. Granger. In *Seasonal Analysis of Economic Time Series* (A. Zellner, ed.) 50–53. U.S. Government Printing Office, Washington, DC. [Also in *CWJWT II* (1985) 939.]
- [53] (1979). When should which spectrum approach be used? In *Forecasting: Proceedings of the Institute of Statisticians Conference* (O. D. Anderson, ed.). North-Holland, Amsterdam. [Also in *CWJWT II* (1985) 981–1000.]
- [54] (1980). Can we predict where 'time series' should go next? In *Directions in Time Series* (D. R. Brillinger and G. C. Tiao, eds.) 1–31. IMS, Hayward, CA. [Also in *CWJWT II* (1985) 941–980.]
- [55] (1984). Styles of spectrum analysis. In *A Celebration in Geophysics and Oceanography—1982, in Honor of Walter Munk* 100–103. Scripps Institute of Oceanography, Reference Series 84-5, La Jolla, CA. [Also in *CWJWT II* (1984) 1143–1153.]
- [56] (1984). *The Collected Works of John W. Tukey I. Time Series: 1949–1964*. Wadsworth, Monterey, CA.
- [57] (1985). *The Collected Works of John W. Tukey II. Time Series: 1965–1984*. Wadsworth, Monterey, CA.
- [58] BRILLINGER, D. R. and TUKEY, J. W. (1985). Spectrum analysis in the presence of noise: Some issues and examples. [Also in *CWJWT II* (1985) 1001–1141.]
- [59] (1986). Sunset salvo. *Amer. Statist.* 40 72–76. [Also in *CWJWT IV* (1986) 1003–1016.]
- [60] (1992). Reflections. In *New Directions in Time Series Analysis, Part I* (D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt and M. Taqqu, eds.) 387–389. Springer, New York.

DEPARTMENT OF STATISTICS  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CALIFORNIA 94720-3860  
E-MAIL: brill@stat.berkeley.edu