COMPUTERS AND BIOMEDICAL RESEARCH **30,** 129–164 (1997)
ARTICLE NO. CO971444

A General Statistical Framework for Frequency-Domain Analysis of EEG Topographic Structure

Edward F. Kelly,* James E. Lenz,† Piotr J. Franaszczuk,‡ and Young K. Truong§

**Department of Diagnostic Sciences and Dental Research Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599; †Department of Radiology, Veterans Administration Medical Center, White River Junction, Vermont, and Dartmouth Hitchcock Medical Center; ‡Laboratory of Medical Physics, Institute of Experimental Physics, Warsaw University, Warsaw, Poland, and Department of Neurology, University of Maryland, School of Medicine, Baltimore, Maryland; and §Department of Biostatistics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599*

Received February 12, 1996

A wide variety of rhythmic electrophysiological phenomena—including driven, induced, and endogenous activities of cortical neuronal masses—lend themselves naturally to analysis using frequency-domain techniques applied to multichannel recordings that discretely sample the overall spatial pattern of the rhythmic activity. For such cases, a large but so far poorly utilized body of statistical theory supports a third major approach to topographic analysis, complementing the more familiar mapping and source-recovery techniques. These methods, many of which have only recently become computationally feasible, collectively provide general solutions to the problem of detecting and characterizing systematic differences that arise—not only in the spatial distribution of the activity, but also in its frequency-dependent between-channel covariance structure—as a function of multiple experimental conditions presented in conformity with any of the conventional experimental designs. This application-oriented tutorial review provides a comprehensive outline of these resources, including: (1) real multivariate analysis of single-channel spectral measures (and measures of between-channel relationships such as coherence and phase), (2) complex multivariate analysis based on multichannel Fourier transforms, and (3) complex multivariate analysis based on multichannel parametric models. Special emphasis is placed on the potential of the multichannel autoregressive model to support EEG (and MEG) studies of perceptual and cognitive processes.

Recent advances in data-collection hardware, together with commensurate advances in methods of data analysis and display, are driving the emergence of high-resolution EEG as an important human functional imaging modality (*1–5*). To date, the improved sampling of the scalp field afforded by modern hardware has been exploited principally in terms of enhancements in the two most commonly applied forms of spatial analysis—namely, topographic mapping and

0010-4809/97 $25.00

source recovery, supported by whatever statistical methods are applied to the outcomes of these operations. In this paper, however, we outline a third major approach to the same class of problems, one based on a large body of existing theory which although highly developed mathematically and statistically has not until recently been computationally practical and is not yet generally as well known to the EEG/MEG research community.

The approach is specifically appropriate to situations in which the activities of interest are naturally measured in the frequency domain. For example, the research situation we usually have implicitly in mind involves the analysis of ''driving'' or frequency-following responses (FFRs) extracted from multichannel scalp EEG recordings obtained in conjunction with exposures to punctate vibrotactile stimuli (*6, 7*). The basic concepts apply equally well, however, to other kinds of electrophysiological data such as multiple spike trains and multichannel recordings from the cortical slice, as well as to a host of other neuropsychological studies of EEG (and MEG) topography. Recent work on induced ''40 Hz'' or gamma-band rhythms, for example, strongly suggests not only that these oscillatory phenomena are engendered in specific cortical regions initially engaged by a task, but also that by virtue of their phase-locked relationship to the spiking behavior of the underlying neuronal population they essentially transcribe themselves into cortical regions with which the zones of origin are functionally connected in the performance of that task (see for example (*8–13*), and note that this argument remains valid even if the oscillatory events are only a sign of the underlying coding activity rather than actually *constituting* it). Thus in a more general setting the measured quantities might reflect ''40 Hz'' responses, activity in one or more of the traditional broad frequency bands (e.g., theta, alpha), etc.

The central questions to be addressed by our analyses of experimental data are then whether—and if so, how—these multichannel batteries of frequency-domain measures change systematically across the conditions of the experiment. Although experimenters are generally accustomed to thinking about such differences primarily in terms of differences between *means*, fundamental biophysical and neurophysiological considerations clearly indicate that the between-channel *covariance* structure of the measures is also likely to undergo systematic changes across experimental conditions, or as a function of time. For example, during adaptation to a prolonged vibratory stimulus, any or all of the following changes may occur in the scalp-recorded image of the activity of the responding cortical neuronal population in areas 3b and 1, as it is progressively reshaped by pericolumnar opponent mechanisms involving a cascade of well-characterized cellular and molecular mechanisms (*6, 14–16*): (1) loss of overall response amplitude, causing low-amplitude regions of the initial response pattern to disappear into the noise background; (2) an apparent shift of mean response location, due to differential modification of neuronal responsiveness over large spatial gradients; and (3) fractionation of the remaining cortical response into multiple, spatially segregated regions of more focal and coherent activity. All of these effects could clearly produce changes in both the mean vectors and covariance structures observed at an overlying array of recording electrodes, the exact mix of effects

depending not only on the neurophysiological changes themselves but also on the geometry of the recording array (*17*). In a more general neuropsychological setting the importance of covariance information again also follows directly from current ''distributed'' models of cognition, which presume that each elementary unit of cortical information processing involves coordination of large and spatially segregated neuronal ensembles, which can be recruited selectively in varying patterns as required by the tasks at hand (*18–20*).

The methods that we will now describe seek to make full use of the information available in the multivariate data structure, extracting quantitative information about both central tendency and covariance properties in relation to the experimental conditions. The subjects to be covered lie at the intersection of two large and difficult technical disciplines—multiple time series analysis and multivariate statistics—to which we can provide here only the barest introduction. (To appreciate the complexities in full, see for example (*21, 22*). Our intent is not to be mathematically complete or rigorous but rather to supply applied research workers with a tutorial overview showing how these disciplines jointly provide a comprehensive collection of practical methods for frequency-domain statistical analysis of topographic structure in multichannel electrophysiological data. To this end, we will present just enough of the mathematical basis to provide insight into the nature of the measures and procedures discussed (together with pointers to the supporting mathematical detail) while emphasizing application-oriented issues and identifying sources of software to perform the relevant calculations. Our goal is to promote this complementary approach to the analysis of functional brain topography, which we believe is poised to become increasingly effective and important given the recent advent of adequate computing resources to support its informed application (mirroring the infusion of multivariate statistical methods into many areas of behavioral research that began in the sixties).

We begin by characterizing more concretely the nature of the data. The entity to be analyzed is an *experiment*, carried out on a single subject,[1] and organized as a sequence of *trials*. We assume that editing for artifact has already taken place (Section 5.2). Each surviving trial involves the sampling of p scalp locations or channels at m equidistant time points, resulting in an $m \times p$ matrix of raw data values. Furthermore, each trial will be associated with one of l experimental levels, or groups, or conditions (for example, vibrotactile stimulus amplitudes, or different stages of adaptation). In principle the experimental conditions might also be cross-classified and/or nested in various ways—i.e., might represent any standard experimental design.

An extremely important and advantageous feature is the availability of re-

[1] The initial restriction to within-subject designs is a considerable departure from the mainstream tradition in EEG research, but (a) this is appropriate, given the strong dependence of scalp EEG patterns on extraneous factors such as skull thickness, cortical folding, etc.; and (b) it involves no loss of generality, as indicated below, since the framework extends readily to between-subject designs. Parenthetically, claims of higher spatial resolution in the early MEG studies—to which most of the material presented in this paper applies equally well—are almost completely confounded experimentally with their more routine use of within-subject designs.

peated observations of the signal expressed in the multiple time series under each of the experimental conditions. In contrast with the situation that dominates most classical discussions of statistical estimation for time series (especially in engineering contexts), in which the analyst has only one realization of the underlying process available for study and the construction of tests and confidence regions depends heavily on stringent assumptions about the properties of that process, we will normally be able to construct estimates of variation and covariation derived from the data itself, and to use these for statistical tests of group differences.[2]

The remainder of this paper consists of a guided tour of Fig. 1, which summarizes the entire collection of methods and their interrelationships. We start at the left with the raw data matrix for a single trial as described above, adding only a notation to the effect that the data matrix may already have been subjected to temporal and/or spatial filtering operations designed to reduce or eliminate competing structure in the raw signal (e.g., background EEG rhythms, line-noise components). From this point the analysis can proceed along three main trajectories, not mutually exclusive, corresponding to (1) real multivariate analysis of vectors of single-channel spectral measures (top section of Fig. 1); (2) complex multivariate analysis based on multichannel Fourier transforms (bottom section); and (3) complex multivariate analysis based on multichannel parametric models (middle section). The sections of Fig. 1 will be covered in this order, together with explanations of their various interconnections.

1. REAL MULTIVARIATE ANALYSIS

This section is for the most part already fairly well represented in the existing EEG literature. Transformation of the raw time data to the frequency domain is first carried out channel by channel, using either nonparametric Fourier transform (FT) or parametric (AR/ARMA) methods. Without going into details which would carry us beyond the scope of this paper, we briefly highlight and contrast some key statistical properties of these very different approaches to spectral analysis.

1.1. FT-Based Spectral Estimation

The dominant FT-based approach capitalizes upon the computational efficiency of fast Fourier transform (FFT) algorithms. Its theory and practice have been thoroughly characterized both in general and in the specific context of EEG analysis (e.g., *23–29*). There are two issues in particular, however, that researchers who wish to apply this very useful tool must bear in mind. First, because the calculated spectrum of the sampled signal involves a convolution of the spectrum of the signal itself with the spectrum of the sampling window

[2] The relatively high ratio of replications to observation points (channels) is also a major advantage of EEG studies with respect to other functional imaging technologies such as 2DG, PET, SPECT, and fMRI in terms of providing access to statistical means for assessment of topographic structure.

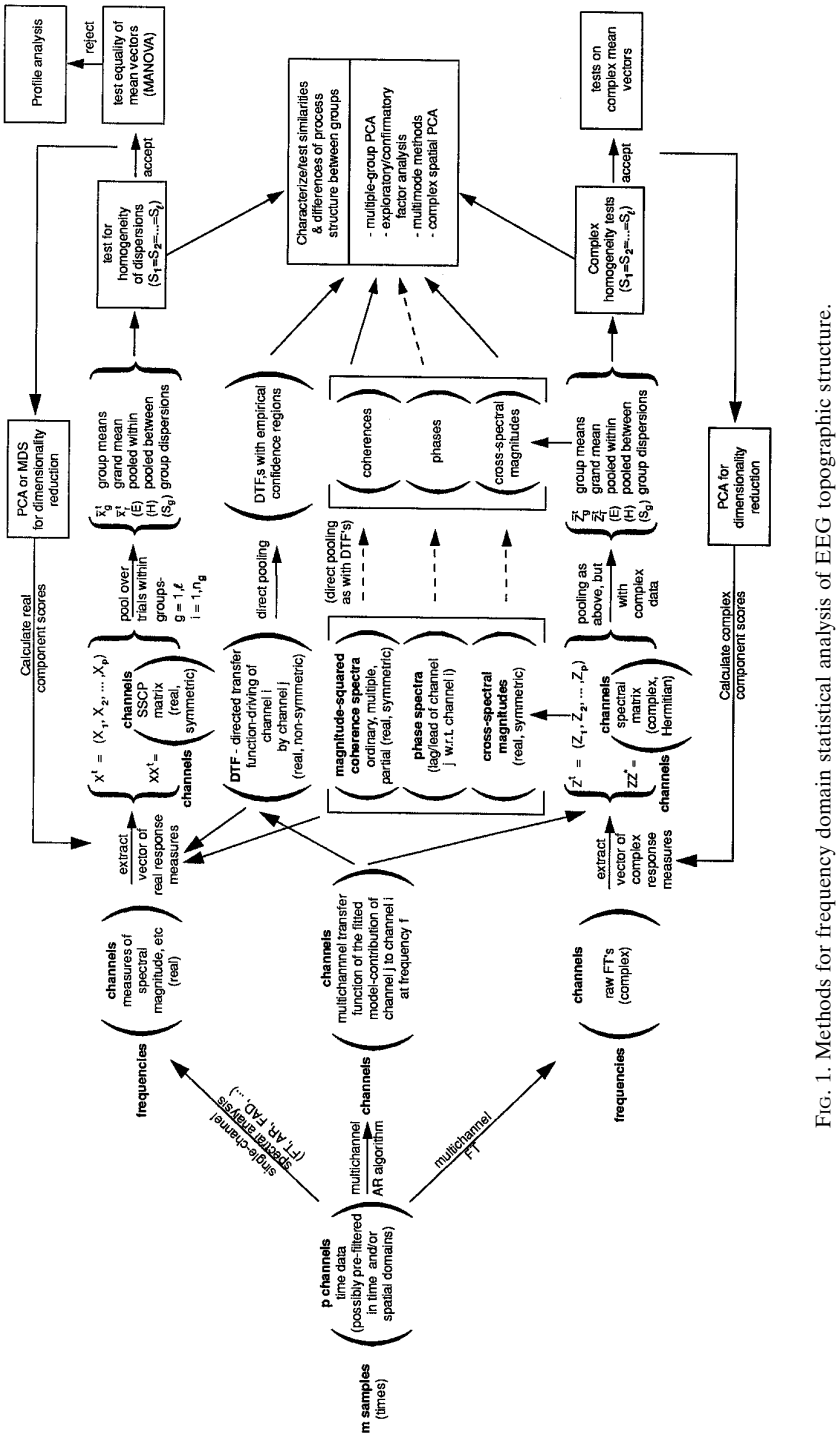

FIG. 1. Methods for frequency domain statistical analysis of EEG topographic structure.

applied to the time data, the window controls "leakage" of EEG power from large peaks into sideband frequencies. A raw EEG time record has in effect been multiplied by a rectangular window of length T sec (the "boxcar" window), which provides maximum frequency resolution ($1/T$ Hz) at the cost of poor sidelobe behavior ($(\sin x)/x$ ripple). The various alternative windowing schemes reduce the leakage (at the cost of discarding data and compromising frequency resolution and in dependence) by smoothly tapering both ends of the raw data record to zero, as discussed in the references cited (see also *30*, which systematically compares the performance of a large number of windows; incidentally, the cosine bell tapers generally used in EEG studies to date fare poorly in these comparisons).

The second issue is more fundamental to the statistical concerns of this paper. The FT is a linear transformation, not a statistical model, of the time record. It amounts to an axis rotation, or equivalently to a projection of the m-point data vector onto an orthonormal basis composed of the cosine and sine functions at $m/2$ equally-spaced Fourier frequencies (see, e.g., *27, 31, 32*). Thus there are as many parameters as data points—i.e., no data compression occurs—and in fact an inverse FT of the complex output exactly reproduces the original time record. The raw spectral estimate or "periodogram" is formed from the complex FT output by conjugation, i.e., by squaring and summing the cosine (real) and sine (imaginary) terms at each frequency. On the assumption that the original time series was zero-mean, Gaussian, and stationary, the linearity of the Fourier transform implies that these estimates are independent and follow a sampling distribution proportional to chi-square with 2 degrees of freedom (df)—i.e., an exponential distribution (these results continue to hold approximately for moderately non-Gaussian and nonstationary data). Since a chi-square with k df has mean k and variance $2k$, the standard deviation of the estimate equals its mean. Extending the time record only generates additional periodogram ordinates and does not change their sampling properties. The estimates can only be improved by "smoothing," which involves averaging the raw periodogram values over frequencies, or over records, or both, each step of the averaging adding another chi-square term with approximately 2 df to the final estimate, depending on the windowing. Chi-square coverges to the normal with increasing df, and the estimates converge toward the true value of the spectrum at each frequency. These improvements are significantly enhanced by applying a log transform, which greatly reduces the right-hand skew of low-df chi-square distributions (*33*).

The choice of a smoothing regime deserves more attention than it customarily receives. For work involving single brief trials as the unit of analysis it is often necessary to strike a compromise by lengthening the record enough to permit some frequency averaging, thus obtaining smoother estimates, but not so much that the biological phenomenon of interest is excessively diluted and/or the record becomes grossly nonstationary. For the analysis of long records a method based on averaging of spectra derived from short, overlapped segments has been shown to have excellent statistical properties (*34, 35*). There is no general solution optimal for all cases.

1.2. Model-Based Spectral Estimation

The model-based or parametric approach to spectral analysis, introduced to EEG studies by Gersch (*36*) and Zetterberg (*37*), is fundamentally different in character. The raw EEG is viewed as a stochastic process, the output of a linear filter excited by white noise (note that this model can be appropriate descriptively whether or not it bears any deep resemblance to the physiological mechanisms that generate the EEG). Spectrum estimation consists of fitting an autoregressive (AR) or autoregressive moving-average (ARMA) model to the observed time data, and then calculating the implied spectrum via FT of the model coefficients (*38, 39*). For the AR case, for example, a family of procedures exist which optimize in a least-squares sense the predictability of the series from its own past and/or future. Thus the basic single-channel AR model is of the form $x(t) = \sum_{i=1}^{k} a_i x(t - i) + e_i$, where $x(t)$ is the observation at time t, the a_i are scalars representing the contributions to $x(t)$ of the observations at earlier times $t - i$, k is the order of the model, and e_i is the residual random error. To fit the model the time series is essentially regressed upon its own prior history (hence the name), using computational methods conceptually analogous to those of classical multiple regression but with algorithmic improvements that capitalize on structural constraints on the autocorrelation or autocovariance matrix of the predictors (see (*38*), and especially (*39*) for algorithms; we omit further discussion here of the computationally and statistically more complex ARMA model).

The crucial properties of these parametric modeling procedures are their capacity to produce smooth, high-resolution spectra even from short data records, provided the signal-to-noise ratio is adequate. The high-resolution property derives from an implied extension of the observed time record beyond its physical boundaries, based upon the fitted model, while the smoothness is associated with the relatively high ratio of data points to estimated parameters. The quantity m/k, where m is again the number of data points and k the model order, is roughly equivalent to the notion of df in the FT case. Model order selection can be accomplished using a variety of automatic or semiautomatic criteria which work by trading off the improved prediction achieved by adding parameters to the model against the increased variability of the resulting estimates. The traditional criteria are known to be asymptotically equivalent, but recent work by Broersen and Wensink (*40*) has identified specific improvements for several of the major AR estimation methods in the finite-sample case (see also *41*).

Based on large-sample results the AR spectral estimates are believed to be asymptotically unbiased, consistent, and normally distributed (*42–44*), and practical experience to date appears to confirm these expectations. Parenthetically, the experience of our lab strongly supports the superior performance on single-channel data of Marple's modified covariance algorithm (*39*) in terms of speed, numerical stability, and resistance to undesirable effects such as peak-splitting and phase-sensitivity. We believe these various desirable properties amply justify the added computational burden of the model-based techniques when short data records (say 2–4 sec or less) are appropriate to the targeted physiological

phenomena. Detailed accounts of the various procedures are provided by (*38, 39, 44*), the latter two including FORTRAN subroutines on diskette. Applications to EEG have been surveyed in (*45*), (*46*).

The measure of signal power at a given frequency in AR spectra is not peak height (as in the FT case) but peak area (*39, 44*). This requires subroutines that can accurately measure areas in the presence of shifting baselines, etc.—a nontrivial problem in itself. Alternative approaches that circumvent this problem are (1) the "FAD" method of Blinowska and Franaszczuk (*47–50*), which represents the EEG as a superposition of a small number of oscillatory modes, and estimates the frequencies, amplitudes, and damping factors of the modes directly from the transfer function associated with the AR model; and (2) the conceptually related model of spectral parameter analysis (*37, 51*).

There is some interest in using the likelihood method to estimate the power spectrum, and it has been shown that this approach is very promising (*52, 53*). More recently, using cubic splines and indicator functions, Kooperberg *et al.* (*54*) developed a fully automatic procedure called LSPEC for estimating both continuous power and line spectra simultaneously. The software is available in `statlib@stat.cmu.edu` and is based on maximum likelihood, with stepwise addition and deletion of basis functions using a *Bayesian information criterion* (BIC). A comparative study is also included in that paper showing that LSPEC performs very favorably relative to other methods including the AR, MA, or ARMA estimates.

1.3. Multivariate Analysis of Variance

Whichever class of spectral analysis methods is applied, the transformation into the frequency domain results for each trial in a frequencies-by-channels real matrix of estimated spectral descriptors. From this matrix we will extract a p-component vector $\mathbf{x}$ of spectral response measures, which in our situation would typically represent the estimated magnitude of the vibrotactile driving response at each of the p channels during this stimulus exposure.[3] This vector times its transpose generates an "SSCP" matrix, containing the associated across-channel sums of squares and crossproducts for the current observation. For the one-way multivariate analysis of variance (MANOVA) situation illustrated in the diagram, these quantities are pooled across observations, yielding mean vectors for each group separately and for the entire sample, the total SSCP matrix, and variance–covariance (or "dispersion") matrices for each group (for computational forms, see for example *55–57*).

The variance–covariance matrix $\mathbf{S}$ is the multivariate analog of the scalar variance associated with univariate ANOVA. When constructed from samples from a multivariate normal distribution it follows the Wishart distribution, which

[3] Other potential sources of real vector-valued observations shown in the diagram will be discussed subsequently. The identification of the number of variables entering the analysis with the number of recording channels is of course not a requirement but a notational convenience appropriate to our experimental situation.

is the matrix analog of chi-square (see, e.g., *56,* Section 3.5, or *58,* Appendix 3A). Its ijth element is $s_{ij} = s_i s_j r_{ij}$, where s_i and s_j are the sample SD's at channels i and j and r_{ij} is their sample linear correlation. From this it can be shown that the determinant of **S**, denoted det(**S**), is a scalar quantity proportional to the squared volume of the p-dimensional solid (the ''parallelotope'') whose principal edges are the vectors representing the variables (channels) in the space of the observations used to form **S** (e.g., *57,* p. 104; *22,* p. 259). This quantity is therefore called the ''generalized variance,'' and it plays a role in multivariate statistics analogous to that of the ordinary variance in univariate statistics.

1.3.1. Testing Homogeneity of Dispersions

One such application is a maximum-likelihood test of homogeneity of dispersions, analogous to the test for equality of variances between groups in the univariate situation. The test is based on the quantity

$$L = N \ln(\det(\mathbf{S})) - \Sigma_g n_g \ln(\det(\mathbf{S}_g)),$$

where $N = \Sigma_g n_g$ and $\mathbf{S}_g$ is the dispersion matrix for group g (*55,* Chap. 4; *56,* Sect. 7.4; *59; 60*). Equality of dispersions is assumed by MANOVA, which for the 1-way case uses the pooled within-group error to test equality of the group mean vectors, and in principle the validity of that assumption should itself be tested. In practice, however—and again as in the univariate case—many workers proceed directly to the MANOVA since the test on dispersions is known to be more sensitive to assumption failures than MANOVA itself. In our scheme, however, it functions not so much as a preliminary to MANOVA as a first step in the analysis of the covariance structure of the topographic data (discussed later). Related maximum-likelihood tests for multiple homogeneity of covariance matrices and homogeneity of populations (simultaneous equality of both the mean vectors and covariance matrices) are discussed in (*59, 60*).

1.3.2. Reduction of Data Dimensionality

As shown in the diagram, the MANOVA procedure can be applied either directly to the original variables (channels) or to a selected smaller set of linear combinations of the channels that preserves most of the original variability. Ratios of sample-size to dimensionality of at least 3 : 1–5 : 1 and preferably 10 : 1 or more are typically recommended for MANOVA (e.g., *61,* p. 232) and it is known that inclusion of variables which do not contribute to discrimination of the experimental groups can severely reduce the power of the multivariate tests (*56,* Sect. 4.8). Thus, with large numbers of intercorrelated channels and small numbers of observations this dimensionality reduction—which can be accomplished using either principal components analysis (PCA) or multidimensional scaling (MDS) techniques—can greatly improve the power and reliability of the multivariate test results. It also directly characterizes structural features of the topographic response pattern common to the experimental groups. We therefore

digress briefly to describe these techniques, which appear in Fig. 1 on the backward-leading arrows at the top (and bottom) of the diagram.

1.3.2.1. Data Reduction by Principal Components Analysis. Principal components analysis (e.g., *56, 57, 62–64*) finds $p \times 1$ linear combinations $\mathbf{v}_i$ of the channels having maximum variance ($\mathbf{v}_i' \mathbf{S}\mathbf{v}_i = \max$) subject to constraints of unit length and mutual orthogonality ($\mathbf{v}_i'\mathbf{v}_i = 1$, $\mathbf{v}_i'\mathbf{v}_j = 0$ for $i \neq j$). The solution corresponds to a rotation of axes into alignment with the principal axes of the original multivariate density, and is obtained by solving the characteristic equation $\mathbf{S}\mathbf{v} = \lambda\mathbf{v}$, i.e., by calculating the eigenvectors and eigenvalues of $\mathbf{S}$. Each eigenvalue represents the variance of the points after projection onto the corresponding axis, or equivalently the amount of the original variance that is accounted for by that linear combination. If the p components are taken in order of decreasing contribution, a small subset $k < p$ will often account for 80–90% or more of the total variance. Criteria for determining the optimal number of components have been reviewed in (*65*). The coefficients or weights express the contribution of each channel to each component, and may help to describe or interpret the associated topographic patterns. It cannot be assumed, however, that the components produced by PCA are physiologically meaningful (e.g., *66; 67*).[4]

If $\mathbf{Y}$ is the original $N \times p$ matrix of raw observations expressed in deviation-score form, $\hat{\mathbf{Y}}$ represents its approximate reconstruction in terms of an $N \times k$ matrix $\mathbf{F}$ of trial-by-trial scores on the k components and the $p \times k$ matrix $\mathbf{V}$ of component weights which minimizes $\mathbf{Y} - \hat{\mathbf{Y}}$ in a least-squares sense. However, for any nonsingular "rotation" matrix $\mathbf{T}$, $\hat{\mathbf{Y}} = \mathbf{F}\mathbf{V}' = \mathbf{F}\mathbf{T}\mathbf{T}^{-1}\mathbf{V}' = \mathbf{F}_1\mathbf{V}_1'$, showing that $\mathbf{F}_1\mathbf{V}_1'$ represents another solution which accounts equally well for $\mathbf{Y}$ (but without possessing the variance-maximizing property that motivated the initial solution). These orthogonal or oblique rotations of the first-stage components according to various criteria of "simplicity" in the loading patterns can sometimes improve interpretability of the component structure, but we will not discuss this large subject here (see refs. cited). An example of physiologically interpretable components extracted by spatial PCA is provided by (*72*).[5]

[4] Silberstein and Cadusch (*68*) have demonstrated that any process, including volume conduction, that causes the covariance between recording channels on a sphere to decrease as a function of their angular separation will lead to the appearance of spherical harmonics as principal components of the covariance matrix—i.e., to components corresponding to the overall size of the signal, anterior vs posterior, left vs right, center vs periphery, etc. In effect this is a spatial analogue on the sphere of the asymptotic equivalence of PCA and Fourier analysis in the time domain (e.g., *27, 31, 34, 69*). However, Valdes *et al.* (*70*) have shown that PCA of spatial structure can be greatly enhanced by performing it on a frequency-specific basis such as we are advocating here, since the spherical-harmonic pattern in real EEG data is principally associated with an overall $1/f$ shape characteristic of the spectral background. In contrast with the spherical harmonics associated with this $1/f$ part of the spectrum, which displayed approximately proportional eigenvalues and zero phase, the principal components associated with specific signal elements such as alpha rhythms displayed clear and appropriate topographic structure, rapidly declining eigenvalues and non-zero phase (see also *71,* p. 268 ff).

[5] In an interesting paper which builds upon earlier uses of PCA, Harner (*73*) has proposed the use of singular value decomposition (SVD) as a general method of analysis for $m \times p$ rectangular

1.3.2.2. Data Reduction by Multidimensional Scaling. Multidimensional scaling refers to a family of data analysis techniques that attempt to represent the structure of the original topographic data in terms of a low-dimensional (usually 2-D or 3-D) space containing one point per variable (channel). The points are arranged so that those that are close together represent channels with similar responses, while those that are far apart represent channels with dissimilar responses (*74–77; 56,* p. 391–398; *57,* p. 572–578). One possibly strong advantage of MDS is its ability to relax assumptions concerning the level of measurement of the responses. Good general-purpose MDS programs are available in SAS, SPSS, and IMSL (see also *76*).

1.3.3. MANOVA

We now proceed to the 1-way MANOVA itself. Matrices **H** and **E** can be constructed that are analogous to the numerator and denominator of the familiar univariate F test, specifically

$$\mathbf{H} = \sum_{g=1}^{l} n_g(\bar{\mathbf{x}}_g - \bar{\mathbf{x}}_T)(\bar{\mathbf{x}}_g - \bar{\mathbf{x}}_T)',$$

$$\mathbf{E} = \sum_{g=1}^{l} \sum_{i=1}^{n_g} (\mathbf{x}_{gi} - \bar{\mathbf{x}}_g)(\mathbf{x}_{gi} - \bar{\mathbf{x}}_g)',$$

where $\bar{\mathbf{x}}_g$ and $\bar{\mathbf{x}}_T$ are the mean vectors for group g and the total sample, respectively, and $\mathbf{x}_{gi}$ is the observation on the ith trial in group g. Parallel to the univariate case the total deviation sum of squares and cross products matrix is thus decomposed exhaustively into parts reflecting the experimental conditions and the within-cells error, and a test is constructed based on the relative magnitudes of these quantities. There are actually a number of test statistics, arising from different test construction methods, all of which are functions of the eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$ (or equivalently of $(\mathbf{E} + \mathbf{H})^{-1}\mathbf{H}$). These same eigenvalues are also solutions of the linear discriminant problem, which is to identify linear combinations of the original variables (channels) that optimally discriminate among the experimental conditions in the metric of the error. Analogously to the PCA situation, this involves finding $p \times 1$ vectors $\mathbf{v}_i$ such that $(\mathbf{v}_i'\mathbf{H}\mathbf{v}_i)/(\mathbf{v}_i'\mathbf{E}\mathbf{v}_i)$ is maximized, subject to the same constraints of unit length and mutual

---

matrices of raw evoked potential data. Basically,

$$\underset{m\times p}{\text{SVD}\,(\mathbf{X})} = \underset{m\times p}{\mathbf{T}}\ \underset{p\times p}{\Lambda}\ \underset{p\times p}{\mathbf{S}'},$$

where Λ is a diagonal matrix containing the p singular values λ_i in order of decreasing size. Rearranging this expression shows that $\mathbf{X} \approx \sum_{i=1}^{k} \lambda_i \mathbf{t}_i \mathbf{s}_i'$ is a weighted sum of the principal ($k < p$) spatiotemporal patterns in the data, where the ith such pattern, described by the outer product of the ith columns of the **T** and **S** matrices, is weighted by the corresponding singular value. SVD thus generalizes previous approaches which have analyzed either the temporal or the spatial structure of the data, but not both simultaneously, using PCA methods. Harner also provides an interesting discussion of properties and applications of the SVD, including references to a rotation method that approximately standardizes the component structure across subjects (see *63*).

orthogonality of the $\mathbf{v}_i$. Successive discriminant functions, up to $l - 1$ in number, are extracted and tested in order of decreasing discriminating power (eigenvalues).

The MANOVA procedure thus exploits the covariance structure of the observations to determine not only *whether* but *how* the experimental groups are statistically distinguished by the across-channel patterns of response. The linear discriminant procedure is known to be optimal for the case of two groups with equal covariance matrices, and in practice usually performs well, particularly when p is large and N small, even if inhomogeneity is present and the theoretically correct rule is quadratic (*62,* p. 163). Furthermore, the procedure is very general. The expressions for **H** and **E** given above, which were chosen to illustrate the general character of the multivariate test in a familiar setting, are specialized forms of much more general expressions that permit testing of arbitrary linear hypotheses of the form $\mathbf{C}\beta\mathbf{U} = \mathbf{A}$ on the parameter matrix β of the multivariate general linear model $\mathrm{Y} = \mathbf{X}\beta + \mathbf{E}$, where **Y** is the $N \times p$ matrix of measured responses for all trials, **X** represents the association between the observations and the cells of the experimental design, and **E** is the error matrix (see *56* or *57* for details). Given an overall significant result, various kinds of followup procedures are also available for determining more precisely the sources and character of the between-group differences. Of particular relevance to topographic studies is profile analysis (loc. cit.), which provides separate tests for differences of shape and elevation of the response patterns between conditions. In combination with a log transform of the channel-by-channel spectral response measures this approach solves in principle the problem identified in (*78*)—i.e., that treatment of topographic effects as channels $\times$ conditions interactions in the context of univariate ANOVA fails to distinguish true changes in the topographic organization of the scalp field from changes in the amplitude of the underlying source configuration. In sum, the MANOVA procedure (which is available in all of the major statistical packages such as SAS, BMD, and SYSTAT, provides a powerful general-purpose vehicle for experimental studies of EEG topography (see also *79*).

We will postpone discussion of further steps in the analysis of between-channel covariance structures until we have worked through the lower and middle sections of Fig. 1, since each of these trajectories generates similar kinds of data structures that can be analyzed in largely parallel ways.

2. COMPLEX MULTIVARIATE ANALYSIS BASED ON MULTICHANNEL FT's

Whereas the top section of Fig. 1 as described above is based upon p-fold repetition of what remains fundamentally a single-channel spectral analysis method, the bottom and middle sections represent analytical approaches that attempt to preserve and utilize more of the information available in the raw multichannel time series. The bottom section, like the top, starts with FT's of all the channels. However, instead of immediately calculating spectral magnitudes,

which destroys phase information, we keep the data in complex form.[6] The extracted p-dimensional vector $\mathbf{z}$ for a particular frequency contains the cosine and sine terms of the response at each channel in its real and imaginary components, and when multiplied by its conjugate transpose $\mathbf{z}^*$ yields an estimate of $\mathbf{S}$, the spectral matrix of the multichannel process at that frequency. $\mathbf{S}$ is complex and asymmetric, but since $\mathbf{S} = \mathbf{S}^*$ it is Hermitian, which implies that its eigenvalues are real. This leads in turn to the interpretation of the product of the eigenvalues, $\det(\mathbf{S})$, as a ''generalized spectrum,'' analogous to the generalized variance of a real multivariate sampling density, which represents the total energy of the multichannel system at that frequency (*80*).

The estimate of $\mathbf{S}$ derived from the raw multichannel FT, like the single-channel periodogram, is statistically unstable, and must be smoothed as usual by averaging over frequencies, observations, or both (*25–28*). With short data records of the sort we have so far been collecting (typically, $T = 2$–4 sec), frequency averaging in steps of $1/T$ would quickly lead to unacceptable loss of frequency resolution, and averaging over observations is therefore the preferred method (but see also below for additional comments on the smoothing regime).

2.1. Complex MANOVA

As shown in the bottom section of Fig. 1, estimates of the mean complex response vectors and spectral matrices for the experimental groups and the entire sample can be developed by pooling trials (observations) in precisely the same fashion described earlier for real multivariate data. Moreover, a substantially parallel body of statistical theory for complex multivariate data has been developed, based on the work of Goodman (*81–83*). Thus if the response vectors $\mathbf{z}_i$ are samples from a complex multivariate normal distribution, the spectral matrix follows a complex Wishart distribution and this in turn permits not only maximum-likelihood tests for homogeneity and multiple homogeneity of the spectral matrices but a complex discriminant-analysis procedure including significance tests on the mean vectors (summarized in (*34, 58, 84–93*)). As in the real case, complex PCA could also be carried out on the pooled spectral martrix to provide dimensionality reduction and analysis/description of the common covariance structure before proceeding to the test of group differences. To our knowledge, these recent statistical developments have so far been utilized in the EEG context only by Rawlings *et al.* (*69, 94*) and by a group of physicists and mathematicians at the Cuban Neuroscience Center in Havana (*70, 80*), although they were also clearly anticipated in the early work of Walter and Adey (*95*). They are not yet available in ''canned'' form in any of the major statistical packages, but can be implemented (and extended where necessary—e.g., (*92*))—using existing

[6] Parenthetically, as pointed out by Ron Katznelson in (*71,* Chap. 6), since the FT is a linear operation, enormous computational savings can be achieved in spatial filtering (Laplacian) procedures by applying them to complex FT output at the frequencies of interest, rather than to the complete raw time data.

libraries of matrix subroutines such as LAPACK or IMSL, or high-level languages such as MATLAB.

2.2. Properties of the Spectral Matrix

Insight into the nature of the FT-based spectral matrix can be obtained by considering more closely the complex operations applied in forming it. If $z_i = (\mathrm{Re}_i + j\,\mathrm{Im}_i)$ is the raw Fourier transform for channel i at frequency f, then the ik element of the spectral matrix zz^* containing cross-periodogram estimates at f is defined as $s_{ik} = z_i z_k^* = (\mathrm{Re}_i\mathrm{Re}_k + \mathrm{Im}_i\mathrm{Im}_k) + j(\mathrm{Re}_k\mathrm{Im}_i - \mathrm{Re}_i\mathrm{Im}_k) = (\mathrm{Re}_{ik} + j\,\mathrm{Im}_{ik}) = \mathrm{CO}_{ik} + j\,\mathrm{QUAD}_{ik}$. If z_i and z_k are thought of as complex vectors describing the activity at channels i and k at this frequency, with lengths $|z_i| = (\mathrm{Re}_i^2 + \mathrm{Im}_i^2)^{1/2}$ and angles $\Theta_i = \tan^{-1}(\mathrm{Im}_i/\mathrm{Re}_i)$, then $|s_{ik}| = |z_i| \cdot |z_k|$ and $\Theta_{ik} = \Theta_i - \Theta_k$. That is, the length of s_{ik} is the product of the lengths of z_i and z_k, and its angle is the difference between their angles, i.e., their relative phase. The real, in-phase, or ''CO''-spectrum part of the cross-products becomes largest if z_i and z_k have the same orientation (positive if they are superimposed, negative if inverted), while the imaginary, 90°-out-of-phase, or ''QUAD''-spectrum part becomes largest if they are at right angles to each other (*96*).

Smoothing the raw cross-periodogram estimates derived from the FT (i.e., the spectral matrix) by vector averaging over frequencies, observations, or both yields an improved estimate of the cross-spectrum with better statistical properties. From this, two important real quantities are derived—i.e., (magnitude-squared) coherence and phase—that more directly capture specific structural features of the multichannel process, i.e., the between-channel patterning of the frequency-domain activity.[7]

2.3. Coherence

The coherence between channels i and j is defined as the normalized cross-spectrum, i.e., $s_{ij}/(s_{ii}s_{jj})^{1/2}$, which is complex and has the same phase as the cross-spectrum itself. For convenience, and to emphasize large values of coherence, most workers prefer to use *magnitude-squared-coherence* (MSC), defined as $\gamma_{ij}^2 = |s_{ij}|^2/s_{ii}s_{jj}$. Since $|s_{ij}|^2 \leq s_{ii}s_{jj}$ (Cauchy–Schwarz), the MSC lies between 0 and 1 (and $\sqrt{\mathrm{MSC}}$, which occasionally appears in the literature, has the same property). $\sqrt{\mathrm{MSC}}$ is analogous to the ordinary bivariate correlation coefficient and measures the linear association between channel i and channel j at frequency f, ignoring a constant phase difference. Similarly, γ_{ij}^2 represents the proportion of power at f common to both channels (*32*).

For pure sinusoids at frequency f observed simultaneously at two channels,

[7] Nunez (*71,* Chap. 7, Sect. 14) also discusses use of the spectral matrix as the point of departure for frequency–wavenumber spectral analysis, a set of techniques commonly applied in the physical sciences for decomposition and measurement of complex wave phenomena such as water waves traveling through the ocean or seismic waves propagating across an array of detectors. See also (*97*).

γ^2 is identically 1. The measured coherence is reduced, however, by a variety of factors including addition of noise or other signal components at one or both channels, instability of phase relations across the frequencies or segments averaged, small shifts in frequency between the sites, and other kinds of non-linear effects (*32, 98–100*). Since $s_{jj} = s_{jj}\gamma_{ij}^2 + s_{jj}(1 - \gamma_{ij}^2)$ the power at channel j can be decomposed into a part linearly predictable from channel i plus an orthogonal residual representing contributions from all other sources, and the quantity $\gamma_{ij}^2/(1 - \gamma_{ij}^2)$ thus becomes a useful measure of the relative size of these contributions. For a purely linear system it measures the signal-to-noise ratio, and for a purely nonlinear system it measures the linear to nonlinear ratio, while for EEG data both noise and nonlinearities presumably contribute to the denominator (*99*). A test of the null hypothesis $H_0: \gamma^2 = 0$ is provided by

$$F = (k - 1)\frac{\hat{\gamma}^2}{1 - \hat{\gamma}^2},$$

which follows an F distribution with 2 and $2(k - 1)$ df if H_0 is true (*58,* p. 73).

Several technical factors are important in FT-based estimation of the MSC. The raw data should be smoothly windowed to suppress leakage, which has been shown to distort coherence estimates severely (*99*). Second, large fluctuations of phase across the vector-averaged frequencies are known to bias the estimate downward and should therefore be avoided if possible (*32*). This again leads to the important issue of the structure of the smoothing regime. Our recommendation is to do whatever minimum of averaging across frequencies is necessary to provide estimates suitable for use by MANOVA (see below; related recommendations are discussed in (*99, 101*)). Third, rapid phase changes between adjacent frequencies can be produced by a sufficiently large time delay between two entirely coherent series, and for this reason some authors recommend *alignment*—i.e., time-shifting so that the maximum cross-correlation occurs at lag zero—as a general practice to eliminate the resulting negative bias (e.g., *27, 96*). However, in the range of frequencies and time delays likely to be encountered in EEG work, this source of bias appears too small to warrant the considerable computational expense of the alignment procedure. Carter (*102*) has derived an expression showing that the downward bias is approximately equal to the true coherence times a factor which is twice the ratio of the time delay to the total duration of the sample.

In parallel with the classical statistical theory of multiple correlation and regression, the analysis of coherence structures can be extended to a variety of conditional forms. If we can appropriately think of the channels as divided into two groups **X** and **Y**, where our interest is in the relationships among the **Y**s with the effects of the **X**s removed, then the spectral matrix can be partitioned accordingly, viz.

$$\begin{pmatrix} \mathbf{S_{yy}} & \mathbf{S_{yx}} \\ \mathbf{S_{xy}} & \mathbf{S_{xx}} \end{pmatrix},$$

and the spectral matrix of the **Y**s conditioned on the **X**s then becomes

$\mathbf{S}_{y|x} = \mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$. From this the conditioned coherences and phases of the **Y**s are calculated as usual (*28, 83;* see also *103*, for an application to EEG). The composite linear predictability of the **Y**s (as outputs) from the **X**s (as inputs) can also be investigated using canonical correlation techniques (*34,* Chap. 10).

Two special cases of the general conditioning theory are usually of particular interest, namely multiple and partial coherences, which are conceptually analogous to the multiple and partial correlations familiar from classical statistics. The multiple coherence associated with a given channel represents its linear dependence on all other channels taken together, while the partial coherence of channels i and j describes their linear dependence with the effects of all other channels removed (*27, 28*). Using expressions derived in (*27*), it can be shown that the squared multiple coherence of channel j is $1 - 1/(s_{jj}s^{jj})$, and the squared partial coherence of channels i and j is $|s^{ij}|^2/(s^{ii}s^{jj})$, where s^{ij} is the ijth element of $\mathbf{S}^{-1}$. Since the ordinary pairwise and partial coherence matrices are each symmetrical—i.e., $\hat{\gamma}_{ij}^2 = \hat{\gamma}_{ji}^2$, etc.—all three types of estimates can be conveniently summarized in a single matrix containing the multiple coherences on the diagonal and the other two types in the upper and lower triangles (e.g., *104*). The conditioning methods outlined here provide in effect a statistical counterpart of physiological interventions such as cooling and ablation, permitting the study of functional interactions between particular brain structures or cortical regions with the linear effects of other structures or regions removed. Lopes da Silva *et al.* (*105*) have applied this approach productively in their studies of thalamocortical relations underlying genesis of alpha rhythms in the dog, and Gersch and colleagues (*106, 107*) have used it to identify the driving sources of human epileptic seizure activity.

Theoretical and empirical studies indicate that the statistical properties of coherence estimates are quite decent. The exact sampling distribution of the MSC depends upon both the population coherence and the df associated with the estimate, and has essentially the same form as the sampling distribution of the squared multiple correlation coefficient, determined originally by R. A. Fisher (*108*). The parallel development for MSCs was carried out by Goodman (*81–83*) and extensive tables have been generated (*109, 110*). The properties of the distribution are summarized in a particularly accessible and useful form by Carter *et al.* (*98*), who provide the density and distribution function, exact and approximate expressions for the bias and variance of the estimator, and illustrations of the density and distribution function under various conditions. The density is unimodal and nearly symmetric for coherences in the range likely to occur in EEG data (roughly, 0.1–0.9). Also, increasing the df of the estimate makes the density taller and narrower without altering its location.

There is, however, a sizeable positive sampling bias which becomes progressively larger as the population coherence grows smaller, and which could therefore confound between-group comparisons if not removed. An approximation to the bias is provided by the quantity $B = k^{-1}(1 - \hat{\gamma}^2)$, where k is the number of frequencies and/or segments used to form the spectral matrix (*111*). Note that

entirely incoherent series will show an average coherence of about $1/k$ (*98*). A quantity suitable for further statistical analysis can be obtained by a simple two-stage procedure which has been shown to provide good estimates even for severely non-Gaussian (e.g., uniform or chi-square) data (*111, 112*): (1) remove the bias—i.e., set $\hat{\gamma}_1^2 = \hat{\gamma}^2 - k^{-1}(1 - \hat{\gamma}^2)$ or zero, whichever is larger; (2) apply Fisher's *r*-to-*z* transform to the square root of the result (which both normalizes the distribution and makes its variance approximately constant and independent of the population coherence)—i.e.

$$\tanh^{-1}(\hat{\gamma}_1) = \frac{1}{2}\ln\frac{1+\hat{\gamma}_1}{1-\hat{\gamma}_1}.$$

More complex procedures which may yield slightly better results, especially near the extremes of population coherence, are presented in (*113, 114*).

Calculations based upon the exact expression for the variance (*98*) indicate that even under worst-case conditions—i.e., using estimates based on averaging over only two frequencies (or segments), with a reference population MSC of 0.33—between-group MSC differences on the order of 0.1 should be resolvable using group sample sizes of about 25–30.

The same procedures can be applied to all other forms of coherence, which as shown by Goodman (*81–83*) follow the same sampling distribution but with df reduced by the number of additional channels involved in the conditioning.

2.4. Phase

The phase relationship between channels i and j at frequency f is calculated from the s_{ij} cell of the smoothed spectral matrix at f as $\Theta = \tan^{-1}(\text{Im}/\text{Re})$. The angles returned (for example by the FORTRAN function ATAN2 (Im, Re)) are in the range $[-\pi, \pi]$, with positive angle implying that channel i leads channel j. Partial phase spectra can also be calculated, using the conditioning operations described above for coherence, which would reflect for example the time relationship between channels i and j with the effects of other channels removed.

The sampling distribution of phase appears to be less fully characterized than that of coherence. Its shape and variance both depend critically on the population coherence. For $\gamma^2 = 0$ the distribution is uniform, while as γ^2 increases it converges toward the normal (*31*, p. 315). Jenkins and Watts (*27*) and Bloomfield (*96*) approach the distribution in terms of a Taylor series expansion of small complex perturbations ($a + jb$) around the true value of the cross-spectrum, and conclude similarly that the estimate is unbiased, is uncorrelated with the coherence estimate, and follows a sampling distribution that is asymptotically normal. Furthermore, the variance decreases (and normality improves) with a factor $\omega = (1 - \gamma^2)/(\gamma^2 \times \text{df})$, where df is the degrees of freedom associated with the estimate. Thus the variability of phase estimates approaches zero as the MSC and/or df increase, as we might expect. There is also a complicated dependence on the magnitude of the cross-spectrum, since a constant amount of scatter in

the z-plane entails progressively less uncertainty about the phase angle as one moves away from the origin. Unpublished theoretical work by one of us (Lenz) revealed that phases sampled from true cross-spectra lying close to the origin of the z-plane can have quite bizarre and non-normal distributions. Although the problem clearly merits further investigation, currently available information suggests that individual phase estimates should have reasonable distributional properties and SD's on the order of 15° or less if $\omega \leq 0.1$, particularly if the coherence is high. Jenkins and Watts (*27,* p. 381) provide a useful chart, from which one can estimate the sample size needed to resolve a given time difference as a function of coherence (see also *115, 116*).

Another possibility meriting further investigation is the use of phase information to improve the structure of the smoothing regime by assuring that we average across frequencies for which the phase is approximately constant. A possible enhancement, if interest centers on a particular small segment of the spectrum, would be to use the chirp-Z transform instead of the FFT (*117,* p. 263), to concentrate the $m/2$ raw cross-spectral estimates around that narrower region. The idea is to find some rational, data-guided basis for grouping raw estimates made at intrinsically higher frequency resolution, so that the final smoothed estimate optimally represents the targeted physiological phenomenon.

A related issue of special significance for topographic studies is the detection and measurement of *time delays* between channels. The *presence* of a time delay between two channels is indicated by a linear relationship between phase and frequency across a band of adjacent frequencies at which they are coherent, and the *magnitude* of the delay is proportional to the slope of that line. Gotman (*118*) applied this idea fruitfully to analysis of the propagation of epileptic seizure activity, and Ktonas and Mallart (*119*) have recently developed an improved estimation method in which the phase at each frequency entering the regression is essentially weighted by the associated coherence. In (*119*) it is also demonstrated via simulations that this estimator and its associated theoretical variance are robust in application to non-Gaussian (e.g., uniform or chi-square), band-passed, and simulated seizure waveform data. A recent comparative study of time-delay estimation methods (*120*) emphasizes the limitations of FT-based (vs model-based) approaches to analysis of rapidly changing dynamic systems such as epileptic seizure EEGs (see below).

Adey and Walter (*121*) provided a prescient early example of the potential value of EEG phase information in behavioral and cognitive studies. One of their many interesting observations involved an apparent reduction in cross-spectral phase scatter within particular frequency bands at high levels of T-maze performance in the cat. A possible alternative to their analysis procedure, which involved construction of "confidence fans" for the magnitudes and phases of the between-channel transfer functions estimated from single trials (*28, 83, 122*), is to utilize a measure of the coincidence of angles sampled from a circular distribution, defined as $C = N^{-1}\{(\sum_{i=1}^{N} \cos \Theta_i)^2 + (\sum_{i=1}^{N} \sin \Theta_i)^2\}^{1/2}$ with standard deviation approximately $(2(1 - C))^{1/2}$ radians (*123,* p. 428–431). Large-sample approximations and a table of critical values appear in (*124,* p. 133–136). A potential

difficulty with this procedure, however, is that all angles are weighted equally, without reference to their associated cross-spectral magnitudes or coherences. Note that in smoothing raw cross-spectra by separately averaging their real and imaginary parts over frequencies this problem does not arise, since phases are implicitly weighted by magnitudes in the vector averaging, and that the same property will apply to complex multivariate analysis as described earlier. Scalar operations on the real angles themselves, however, such as averaging over trials, must be handled extremely carefully to assure meaningful results. In particular, the set of angles averaged must not straddle the discontinuity at 0°/360° (see also *125* for related discussion).

2.5. Real Multivariate Analysis of Coherences and Phases

The distributional information summarized above clearly supports the applicability of real multivariate procedures (top section of Fig. 1) to coherence measurements, and with adequate care to phase measurements as well. Consequently, one general approach to analysis of these matrices is simply to string out a vector of the observed values—coherences, for example–on a trial-by-trial (or subject-by-subject) basis and subject these to MANOVA as described previously. Since there are $p(p-1)/2$ distinct pairings of the p channels, however, some kind of selection or dimensionality reduction must be applied to constrain the number of variables as the number of channels increases, in accordance with the sample size/dimensionality rules quoted earlier. Prior EEG or metabolic imaging studies, for example, might have identified a subset of especially important coherences suitable for further analysis (e.g., *126*). Alternatively, PCA or MDS could again be applied for reduction of dimensionality and identification of key structural components. The p multiple coherences could themselves form an interesting set of variables reflecting the total interrelatedness of the channels. Carrying this compression of data to the limit, the negative logarithm of the determinant of the complex coherence matrix can be interpreted (*80*) as a scalar measure of "generalized coupling," analogous to the "generalized spectrum," $\det(\mathbf{S})$. Similarly, one might calculate the mean and SD of the $p(p-1)/2$ real pairwise coherences.[8]

A significant example of the use of phase information in topographic analysis is the work of Walter Freeman (*11*), who has carried out extremely detailed studies of spatiotemporal patterns of gamma-band ("40 Hz") activity engendered in rabbit olfactory bulbar cortex near the peak of the respiratory cycle. For this purpose, phases determined individually at each channel could be measured (and mapped) relative to that of an arbitrarily selected channel, or to the average phase of all channels. In the context of our own studies (*6*) the phase of the

[8] We also mention in passing an interesting paper (*127*) which develops an overall coherence measure, $C_{ij} = \Sigma_f |s_{ij}(f)| \hat{\gamma}(f) / \Sigma_f |s_{ij}(f)|$ that weights coherences across frequencies by the associated cross-spectral magnitudes. Although the neurophysiological rationale offered for this measure appears flimsy, the author applies it successfully both to narrow-band photic driving responses and to wide-band characterization of cortical functional state prior to stimulus onset.

cortical driving response with respect to that of the vibratory stimulus itself is an important response parameter which is expected to undergo systematic and spatially differentiated changes as a function of experimental conditions such as stimulus magnitude or the stage of adaptation. Consequently, it would be natural for us to extract from the spectral matrix for each trial a vector containing the observed response lag at each channel, and to subject these to MANOVA.

3. COMPLEX MULTIVARIATE ANALYSIS BASED ON MULTICHANNEL PARAMETRIC MODELS

We turn now to the middle section of Fig. 1, which emphasizes a parametric or model-based approach to multichannel analysis that has only recently become computationally practical for realistically large arrays of channels. Such models have been advanced for EEG especially by W. Gersch and his colleagues (*36, 128–132*) and by K. Blinowska and hers (*104, 133*). As in the case of single-channel analysis, the basic idea is to identify an AR or ARMA model that fits the time-domain data optimally in a statistical sense, and then to calculate the required spectral parameters from the fitted model. The basic multichannel time-domain AR model[9] is of the form $\mathbf{x}(t) = \sum_{i=1}^{k} \mathbf{A}_i \mathbf{x}(t - i) + \mathbf{e}(t)$, where $\mathbf{x}(t)$ is the vector of samples at the p channels at time t, each of the k $\mathbf{A}_i$'s is now a $p \times p$ matrix, each row of which expresses the contributions to the corresponding channel at time t of all the channels at time $t - i$, k is again the order of the model, and $\mathbf{e}(t)$ is the residual error vector. In (*104*) such a model is fitted by efficiently solving a multichannel version of the Yule–Walker equations. (Other algorithmic approaches to fitting multichannel AR models are available, and may as in the single-channel case prove to have better properties (*39, 134*), but this would not affect subsequent discussion.) Model order selection can again be determined in a semi-automated way using multivariate generalizations of the single-channel test criteria (*39, 44, 104*). An FT of the model coefficients produces the estimated transfer function of the underlying multichannel process at each frequency, and this in turn leads to estimates of the spectral matrix and its offspring—i.e., the matrix of spectral and cross-spectral magnitudes, and the matrix of multiple, pairwise, and partial coherences as described previously.

Kaminski and Blinowska (*133*) have developed one further $p \times p$ real matrix displaying a new quantity, the "directed transfer function (DTF)." Building upon earlier attempts to define a directed coherence function, the directed transfer of channel j to channel i is calculated directly from the transfer function of the

[9] At the moment there seems to be little reason to pursue further development of the computationally much more intensive multichannel ARMA model for purposes of EEG analysis. Marple (*39,* p. 397) points out that the multichannel AR model in fact provides for ARMA-type spectra of the individual channels, with corresponding flexibility to fit a variety of spectral shapes. Not surprisingly, the few direct comparisons of multichannel AR and ARMA models currently available in the literature—notably from (*129*)—yielded virtually identical results. The penalty for the slight improvements possibly gained by the ARMA approach is a considerable inflation of the number of estimated parameters, and a corresponding increase in the required length of the time sample.

fitted AR model, essentially by dividing the contribution of channel j to channel i by the sum of all contributions to channel i (see the original paper for details). In effect this calculation utilizes phase information implicit in the transfer function to construct measures of the capacity of each channel to "drive" each other channel in an informational sense. This (non-symmetric) $p \times p$ matrix of DTFs therefore appears particularly useful for tracking the "flow" of frequency-specific information through the recording array. The statistical properties of the DTF have not yet been determined, but are likely to be similar to those of coherence estimates. The properties and power of the DTF method are illustrated in (*135*) using intracranial recordings of epileptic seizure EEGs.

Two general points about the model-based (vs FT-based) approach to multichannel analysis merit special emphasis: First, the AR model is a true multichannel model in the sense that the calculated spectra and cross-spectra reflect the interdependencies of *all* the channels taken simultaneously, in contrast with the FT approach which obtains them as products of the transforms of individual channels taken just one or two at a time. Second, the desirable high-resolution and smoothness properties described earlier for single-channel parametric spectral estimates continue to hold in the multichannel case. That is, relatively good and low-variability estimates of the spectral matrix and associated quantities can be obtained from quite short data records, since most of the noise has already been removed, as it were, by fitting the AR model. For coherence estimates, for example, the bias factor is reduced to k/m, and the variance becomes approximately $(k/m)^2$, where k is again the order of the model and m the number of time samples (*129*). Similarly, multichannel AR algorithms yield superior estimates of time delays in rapidly changing dynamic systems (*120*). This appears to open up significant new possibilities for single-trial correlation of EEG patterns with human psychophysical or cognitive responses (*48, 129, 134*).

One pathway to further analysis of the multichannel AR results again involves extracting vectors of real-valued measures from the various matrices and subjecting these to MANOVA. As described previously, these measures might represent all the cells of a matrix, a selected subset of the cells, or summary measures constructed to reflect overall properties of the system in a more compact way. In the case of DTFs, for example, one such summary vector might contain for each channel j of the recording array $\Sigma_{i \neq j}\, \mathrm{DTF}_{ij}$ which would measure the total capacity of that channel to "drive" the other channels in the array, integrated over a particular range of frequencies. Coherence and phase information can be extracted and processed in similar fashion, as described earlier.

The complex spectral matrices themselves can also be pooled across observations and analyzed using the methods of complex multivariate analysis described previously. However, there is no quantity analogous to the FT-based "observation vector" z at frequency f, the spectral matrix $\mathbf{S}$ being computed for the AR model from the relation $\mathbf{S} = \mathbf{HVH}^*$, where $\mathbf{H}$ is the estimated transfer function and $\mathbf{V}$ is the calculated matrix of residual errors. For this reason, complex discriminant analysis (*58, 93*) is not available in this case as a method for testing and characterizing between-condition differences.

One readily available and practical alternative is to construct empirical confidence regions cell by cell for the matrices of interest by pooling them across observations (trials or subjects) within the experimental groups. Although this is certainly a worthwhile first step, it suffers essentially the same liabilities, in terms of multiple-testing problems and preoccupation with pairwise comparisons of experimental groups, as current topographic mapping methods (*136, 137;* hence the broken lines in Fig. 1).

4. METHODS FOR BETWEEN-GROUP COMPARISON OF MATRICES

For a more general solution, depicted on the right side (center) of Fig. 1 as the final common pathway for all three of its major sections, we turn to the growing family of methods, many quite recent in origin, for comparison of covariance, correlation, and proximity structures among multiple groups. Although with DTFs the final matrix estimate for each group is normally obtained by averaging over the individual DTF matrices themselves (*133*), for the other spectral quantities the preferred method is to average the underlying spectral matrices and then calculate the coherences, phases, etc., from these averages.

Unlike the univariate case, in which the scalar variances of the groups are either homogeneous or inhomogeneous, period, the multivariate situation permits varying degrees or types of similarity between matrices. A useful tool for exploring the degree of similarity is a recent generalization of ordinary principal components analysis to multiple groups. This procedure, simultaneous component analysis (SCA, *138–140*), identifies a single set of components that accounts for as much as possible of the variance in all the groups simultaneously. Comparison of the SCA solution with individual PCAs of the groups in terms of the amount of variance explained and the structure of the weights provides insight into both the extent and the character of between-group differences in covariance structure. Software to perform the analyses, including rotation procedures and interpretational aids, is available from IEC ProGAMMA, c/o Drs. K. P. Koning, P.O. Box 841, NL-9700 Groningen, The Netherlands.[10]

Results generated by SCA, together with any known neuroanatomical/neurophysiological constraints, can be deployed into a more formal, confirmatory,

[10] Flury (*62*) has presented a different approach to simultaneous PCA in multiple groups, based on a novel algorithm for simultaneous diagonalization of matrices. Specifically, he establishes a five-level hierarchy of relationships of increasing complexity in terms of the number of parameters required by an adequate model of the data, together with a sequential maximum-likelihood decomposition/testing procedure for identification and characterization of the best-fitting model for a given dataset. Among the datasets he analyzes in this way, with new and interesting results, is the iris data used by R. A. Fisher in his famous original exposition of the linear-discriminant problem. He also provides in FORTRAN the core algorithm (also available in IMSL as subroutine KPRIN, or double-precision version KDPRIN, which provides it in both maximum-likelihood and generalized least-squares versions). Although further practical experience will be needed to evaluate the utility of these procedures relative to SCA, it appears that Flury's method could, in principle at least, produce solutions that are nicely diagonal but whose common components account for relatively little of the variance in the original data (H. Kiers, personal communication). See also below and note 11.

or hypothesis-testing stage of analysis using existing commercial software for specification and testing of structural equation models. For example, as part of their long-term effort in this area, Jöreskog and Sörbom (*141–145*), have presented a comprehensive maximum-likelihood procedure which applies specifically to real, symmetric matrices of the covariance type—such as the dispersion matrices generated in real multivariate analysis (top section), or matrices of cross-spectral magnitudes at a particular frequency (bottom and middle sections)—and which in principle permits testing of any hypothesis regarding between-group differences in factor structure. To illustrate these fully general estimation and testing procedures, Jöreskog (*142*) also outlines a sequence of progressively restrictive tests that might naturally and routinely be applied to a set of covariance matrices for which the null hypothesis of homogeneity of dispersions (complete invariance of factor structure across groups) has been rejected by the overall test described earlier. For example, if the hypothesis of an equal number of common factors in each group is tenable, one can go on to test equality of the loading matrices and so on, adding further elements of the complete factor structure at each step. Sörbom (*143*) extends Jöreskog's method to test invariance of the factor means, and Byrne *et al.* (*146*) illustrate how to deal with partial measurement invariance and identify sources of between-group differences. These procedures are available, fortified by large amounts of practical experience (particularly in educational and sociological contexts), through the LISREL software package (*144, 145*), which is available in SPSS and as a standalone, and through related systems such as EQS in BMDP, COSAN, and (soon) PROC CALIS in SAS. The potential sensitivity of the maximum-likelihood test procedures to failures of their distributional assumptions, particularly with small sample sizes, is effectively addressed within these packages by providing tests of multivariate normality, alternative model-fitting and test criteria, bootstrap methods for estimating the distributions of model parameters, etc. A structural equations approach to analysis of functional brain topography has recently appeared in the context of metabolic imaging studies (*147*).

Coherence matrices are much like conventional bivariate correlation matrices, which for decades have been the central data objects of the factor-analytic tradition in psychometrics. Tucker and Roth (*148*) were apparently the first EEG researchers to recognize this relationship, and to explore systematically the application of both PCA and factor-analytic methods to coherence matrices as a method of topographic analysis. Factor analysis is a large and complex subject (e.g., *149–151*), but in brief, it differs conceptually from PCA in attempting to fit a statistical model which assumes that the observed variables include measurement error, and which attempts to explain the correlational structure of the original data in terms of a small set of underlying or ''latent'' common variables (the factors). A key step in the iterative computational procedures associated with the principal-factor method applied by Tucker and Roth (*148*) is initial estimation of the ''communalities,'' quantities placed on the diagonal of the correlation matrix which represent the proportion of the variance of each variable (channel) to be explained by the common factor structure, and for this purpose

they sensibly used the squared multiple coherences of the channels. Orthogonal and oblique rotations were applied to the initial solutions to obtain simpler and more interpretable factor structures. As is often the case in practice the PCA and factor-analytic approaches produced generally similar results (*64*), although the latter appeared to be slightly more consistent over repeated measurements. We comment parenthetically that most statisticians strongly prefer recently developed scale-free maximum-likelihood and generalized least-squares factoring methods (*56, 57, 63*); however, the results are again unlikely to differ remarkably from those obtained by the more traditional methods, which are also less computationally demanding.

A secondary goal of Tucker and Roth (*148*) was to compare and contrast the factor structures emerging under two different experimental conditions. They did this in the traditional impressionistic, narrative style, but pointed out correctly that development of more formal and general comparison methods has become an area of intense activity in psychometrics. In particular, the LISREL-type procedures discussed above, although initially developed for analysis of covariance matrices, should be applicable to coherence matrices as well.

Another set of approaches derives from the seminal work of Ledyard Tucker (*152–154*), which has inspired development of a family of ''multimode'' extensions of traditional factor-analytic, PCA, and MDS methods to data matrices of more than two dimensions (see (*155*) for an excellent survey). In the situation at hand, for example, exploration and description of similarities and differences in topographic structure expressed in the covariance and coherence matrices of the experimental groups can appropriately utilize techniques of weighted or individual-differences multidimensional scaling applied to the *channels* × *channels* × *groups* data structure (e.g., *76,* especially Section 4.5; *77; 155*). The general-purpose MDS program ALSCAL, for example—which is available in SPSS, IMSL, and older versions of SAS, and as a FORTRAN standalone—can perform metric and nonmetric individual-differences scaling for both symmetric and nonsymmetric matrices of real data, and would thus support exploratory analysis of phase or DTF matrices as well. Related multimode techniques can in principle also be applied to more complete data structures such as *frequencies* × *channels* × *conditions*, and even to these same data structures with *subjects* added as a fourth mode of classification (*155*).[11]

[11] Field and Graupe (*156, 157*) have recently provided the first application of a multimode model to multichannel evoked-potential data, fitting the direct parallel-factors (PARAFAC) model of Harshman (*155, 158*) to multi-subject visual evoked potential data (i.e., a *timesample* × *channels* × *subjects* design). Importantly, an identical trilinear statistical model had been derived independently by Möcks (*159*) from *a priori* reasoning about the biophysical origins of evoked potentials. The authors explored a number of the many data preprocessing and computational options associated with the nonlinear least-squares minimization procedure and were able to obtain a unique and apparently meaningful (i.e., interpretable) solution that was reproducible across subject groups. The rotational indeterminacy and orthogonality constraints of PCA-type methods, including the SVD method (*73;* see note 5), thus appeared to be overcome by this multimode approach, at the cost of both significantly increased computation and the need to make strong assumptions about the structure of the data—in this case that each temporal component is associated with a specific spatial component,

Principal components analysis of the Hermitian spectral matrix was introduced to EEG topographic studies by Nunez (*71,* pp. 268 ff) who used it to analyze the spatial distribution of certain alpha-band frequency components of the resting EEG in individual subjects. The eigenvectors of this complex matrix are themselves complex, which permits construction of separate topographic maps for the amplitudes and the phases of the extracted components. Although certainly useful, these results remain largely at the level of description of topographic patterns condition by condition, rather than formal analysis of between-condition similarities and differences in topographic organization of the overall activity patterns. A partial solution is provided by the tests on complex covariance matrices discussed in (*92*). However, the LISREL-type and multimode methods referenced above unfortunately apply only to real data at present, and although the extension to the complex case is presumably feasible, to our knowledge this has not yet been attempted.

5. RELATED TOPICS

This completes our outline of the basic framework of statistical resources for frequency-domain analysis of topographic patterns. Before concluding, however, we need to say a bit more about some practical issues in application of these techniques, and about some related techniques that we have consciously slighted.

5.1. Data-Acquisition Methods

Accurate measurement of topographic patterns clearly places strong demands on the EEG data-acquisition system. Between-channel phase and amplitude differences of instrumental origin should be eliminated by simultaneous sample-and-hold on all amplifier channels, coupled with careful equalization or at least measurement of their transfer functions (*160, 161*). It has also been pointed out many times that the choice of common reference for the initial recording is especially important for EEG topographic studies, since it determines the zero-potential line of the measured scalp field (e.g., *17, 162, 163*). More generally, since the signal recorded is actually the difference between the potentials at the active and reference sites, activity at the reference can influence the amplitude and phase of signals attributed to the targeted sites—and measurements of coherence and phase between them—in a frequency-dependent, significant, and largely unknown manner (*164*). In our experimental situation a reasonably good approximation to the idealized remote or indifferent common-reference point is provided by the earlobe ipsilateral to the stimulated hand (*165*). The balanced noncephalic reference (*166*) would also be a good choice in many applications.

---

and that this spatiotemporal structure is substantially common to all subjects. These assumptions of the model could be roughly tested by its ability to fit the raw data, and appeared tenable for this single-condition experimental design. They would likely hold even better in a single-subject design with replications as the third mode of classification. In its indirect mode PARAFAC would also complement the LISREL-type approach to analysis of multiple covariance matrices, discussed previously.

However, in most settings reference-independent surface-Laplacian methods appear likely to provide the best starting point for topographic analysis (*164*). These also reduce between-channel correlation due to volume-conduction effects (note 4), enhance spatial resolution, and have the great advantage for most neuropsychological purposes of emphasizing local cortical sources at the expense of deeper and/or more distant ones (e.g., *2–4, 167*).[12]

5.2. Data Screening and Distributional Considerations

Second are some issues related to data integrity and data screening. For example, with larger numbers of recording channels it becomes increasingly likely that on any given trial one or more of the channels will be corrupted by some form of artifact. We believe that despite its high cost the safest policy in general is to exclude all the data from such trials from further analysis, in the interest of maximizing fidelity of the observed topographies. Special attention should also be paid to elimination of outliers from the derived spectral quantities. Outliers have bad effects in general, but especially on tests involving determinants. Related to this is the general question of distributional assumptions and the effects of violations. For several of the measured quantities such as FT-based autospectral magnitudes and between-channel coherences, transformations have been recommended that appear to normalize and/or otherwise improve their statistical properties, and estimates related to coherence and phase have specifically been shown in extensive simulation studies to be robust against violations of Gaussianity assumptions (*112, 119*). Although tests for multivariate normality of vector-valued observations are available (*56, 57, 172*), MANOVA appears to enjoy the same robustness against assumption failures as its univariate counterpart, particularly when the group *N*'s are roughly equal, the number of variables is not too large in relation to the sample size, and the dispersion matrices are not grossly inhomogeneous (e.g., *56, 57, 173*). Also, it is significant that many major contributors to time-series theory have specifically encouraged analysis of real, finite data with the expectation that desirable large-sample or theoretical properties will at least approximately hold (*32, 34, 81, 82, 174*). Marginal statistical results should certainly be regarded with suspicion, however, particularly if they cannot be replicated. Alternative approaches worth considering include appropriate normalizing transformations (*57*), replacement of the original interval-scale data by their ranks, and construction of brute-force approximations to

[12] Recent reports by Biggins *et al.* (*168, 169*) have suggested that the spherical spline procedure of Perrin *et al.* may artifactually and significantly inflate measured coherences. However, Perrin (*170*) and Pascual-Marqui (*171*) have strongly disputed this. We agree in principle with the latter authors, while recognizing that in practice small effects of the type identified by Biggins *et al.* may well exist. Any kind of spatial smoothing involves processing against noise, and thus necessarily increases coherence. However, whether this increase is "artifactual" or not depends on the relationship between the smoothing regime and the volume-conductor properties of the head. The various spline methods seek to provide spatial bandpass filters that match these properties as nearly as possible, and to the degree that they succeed in this they would not be expected to produce spurious increases in coherence.

nonparametric randomization tests by repeated calculation of test statistics under random assignment of trials (cases) to groups (*175*). Jackknifing and cross-validation techniques can also be used to explore the generalizability of results, and are especially important in the context of classification problems (*61*).

5.3. PROBLEMS OF NONSTATIONARITY AND NONLINEARITY

The techniques outlined in this paper derive mainly from linear systems theory applied to jointly Gaussian and weakly stationary random processes. This is clearly at best an idealization of the real situation. The EEG is approximately stationary only over short stretches of time and its amplitude distribution is often strongly non-normal (*176, 177*). Nonlinear effects are also certainly abundant in the nervous system. A pertinent example at the population (EEG) level is the strong tendency for the initial cortical driving response set up by a vibrating probe to be displaced slightly in frequency from that of the stimulus (*7*). This feature can make the connection between the stimulus and the response, or between the responses observed at different scalp locations, invisible to the methods we have described—even if the two frequencies fall within the same frequency band at the base resolution of the analysis—if the record is sufficiently long (*100*). Various techniques have been proposed which may be capable of addressing such issues, including nonlinear cross-spectral analysis (*178*), approaches based on the theory of mutual information (*130, 179, 180*), and higher-order spectral analysis (*181, 182*), all of which have been applied in at least exploratory fashion to EEG data. The theory of nonlinear dynamic systems and nonstationary random processes is also currently undergoing vigorous development in various areas of mathematics and engineering. However, our own strong inclination is to avoid or at least postpone dealing with the substantially greater mathematical and computational complexity of these approaches pending a clear demonstration that the already very general and versatile techniques outlined above are incapable of doing the job. That the basic multichannel AR model is appropriate for most EEG data is strongly suggested by the fact that it usually yields a clear minimum in the order-selection function, associated with low residual variances. Moreover, it can be modified in various ways to handle special situations. To deal with nonstationary EEGs, for example, Gersch (*128*) has proposed a time-varying extension of the basic model in which the AR parameters themselves become smoothly varying functions of time. Less computationally intensive approaches to nonstationarity which can also handle larger numbers of channels apply a conventional multichannel AR model to successive small segments of the process, possibly overlapped (e.g., *134, 183*). Similarly, nonlinear effects can often be handled in the context of linear models by incorporating appropriate functions of the variables (e.g., polynomial terms), while leaving the model linear in its parameters. These examples illustrate how the basic approaches we have outlined, which themselves usually provide at least a good first approximation to a correct model of the situation, can also be adapted further to deal with non-stationarity, non-linearity and related assumption failures.

6. CONCLUSION

We believe it is axiomatic that statistical analysis of interactions among cortical areas is the key to realizing the full potential of high-resolution EEG/MEG techniques as a research tool for human functional imaging studies. For neural activities which are appropriately quantified in the frequency domain, the framework we have outlined addresses this objective by identifying a comprehensive set of analytical methods, together with sources of enough low-cost computing power to apply them economically to real data. By assembling a sizeable part of the relevant theory and practice in one place we hope to have contributed to their effective deployment in future research.

APPENDIX: NOMENCLATURE

Notation	Meaning
$\mathbf{v}, \mathbf{x}, \ldots$	vectors
$\bar{\mathbf{x}}$	vector of means of $\mathbf{x}$
f	frequency
γ_{ij}^2	magnitude-squared-coherence of channels i and j
Γ	diagonal matrix
$\mathbf{H}$	ANOVA sum of squares (hypothesis)
$\mathbf{E}$	ANOVA sum of squares (error)
λ	eigenvalues
Re	real part of a complex number
Im	imaginary part of a complex number
CO	co-spectrum
QUAD	quad spectrum
s_{jj}	power spectrum
s_{ij}	cross-power spectrum
$\mathbf{S}$	variance–covariance matrix
SVD($\mathbf{X}$)	singular value decomposition of a matrix $\mathbf{X}$
det($\mathbf{S}$)	determinant of matrix $\mathbf{S}$
$\mathbf{T}^{-1}$	inverse matrix of $\mathbf{T}$
$\mathbf{T}'$	transpose of $\mathbf{T}$
$x(t)$	time series

ACKNOWLEDGMENTS

The authors express their thanks to K. J. Blinowska, Ross Dunseath, Stephen Folger, Maciej Kaminski, Debra McLaughlin, Abigail Panter, Forrest Young, and Barry Whitsel for helpful comments, discussion, and encouragement contributing to the development of this paper, and to Reneé B. Williams for diligent typing of its multiple editions. The work was supported in part by NIDR program project grant (DE07509) and in part by an associated Fogarty Center grant (FIRCA, TW00185).

REFERENCES

1. Gevins, A. High-resolution EEG. *Brain Topogr.* **5,** 321 (1993).
2. Nunez, P. L., Silberstein, R. B., Cadusch, P. J., Wijesinghe, R., Westdorp, A. F., and Srinivasan, R. A theoretical and experimental study of high resolution EEG based on surface Laplacians and cortical imaging. *Electroenceph. Clin. Neurophysiol.* **90,** 40 (1994).

3. Perrin, F., Bertrand, O., and Perner, J. Scalp current density mapping: Value and estimation from potential data. *IEEE. Trans. Biomed. Eng.* **BME-34,** 283 (1987).
4. Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroenceph. Clin. Neurophysiol.* **72,** 184 (1989). Erratum: **76,** 565 (1989).
5. Thatcher, R. W., Hallett, M., Zeffiro, T., John, E. R., and Huerta, M. (Eds.) ''Functional Neuroimaging: Technical Foundations.'' Academic Press, Orlando, FL, 1994.
6. McLaughlin, D. F., and Kelly, E. F. Evoked potentials as indices of adaptation in the somatosensory system in humans: A review and prospectus. *Brain Res. Rev.* **18,** 151 (1993).
7. McLaughlin, D. F., and Kelly, E. F. Peripheral vs. cortical driving responses evoked by brief vibrotactile stimuli in humans. [Abstract]. *Proc. Soc. Neurosci.* (1993).
8. Gray, C. M., and Singer, W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* **86,** 1698 (1989).
9. Gray, C. M., Konig, P., Engel, A. K., and Singer, W. Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature.* **338,** 334 (1989).
10. Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitbock, H. J. Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biol. Cybern.* **60,** 121 (1988).
11. Freeman, W. J. Analytic techniques used in the search for the physiological basis of the EEG. *In* ''Methods of Analysis of Brain Electrical and Magnetic Signals'' (A. S. Gevins and A. Remond, Eds.), Handbook of EEG, Revised Series, Vol. I, pp. 583–664. Elsevier, Amsterdam, 1987.
12. Singer, W. Search for coherence: A basic principle of cortical self-organization. *Concepts Neurosci.* **1,** 1 (1990).
13. Tononi, G., Sporns, O., and Edelman, G. M. Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex* **2,** 310 (1992).
14. Favorov, O. V., and Kelly, D. G. Minicolumnar organization within somatosensory cortical segregates. I. Development of afferent connections. *Cerebral Cortex* **4,** 408 (1994).
15. Favorov, O. V., and Kelly, D.G. Minicolumnar organization within somatosensory cortical segregate. II. Emergent functional properties. *Cerebral Cortex* **4,** 428 (1994).
16. Whitsel, B. L., Favorov, O., Kelly, D. G., and Tommerdahl, M. Mechanisms of dynamic peri- and inter-columnar interactions in somatosensory cortex: Stimulus-specific contrast enhancement by NMDA receptor activation. *In* ''Information Processing in the Somatosensory System'' (O. Franzen and J. Westman, Eds.), pp. 353–370. Stockton, New York, 1991.
17. Nunez, P. L. Spatial filtering and experimental strategies in EEG. *In* ''Statistics and Topography in Quantitative EEG'' (D. Samson-Dollfus, Ed.), International Workshop, Paris, pp. 196–209. Elsevier, Amsterdam, 1988.
18. Goldman-Rakic, P. S. Changing concepts of cortical connectivity: Parallel distributed cortical networks. *In* ''Neurobiology of Neocortex'' (P. Rakic, and W. Singer, Eds.), pp. 177–202. Wiley–Interscience, Chichester, 1988.
19. Shaw, J. C., O'Connor, K., and Ongley, C. EEG coherence as a measure of cerebral functional organization. *In* ''Architectonics of the Cerebral Cortex'' (M. A. B. Brazier and H. Petsche, Eds.), IBRO Monograph No. 3, pp. 245–255. Raven Press, New York, 1978.
20. Rappelsberger, P., and Petsche, H. Probability mapping: Power and coherence analyses of cognitive processes. *Brain Topogr.* **1,** 46 (1988).
21. Hannan, E. J. ''Multiple Time Series.'' Wiley, New York, 1970.
22. Anderson, T. W. ''An Introduction to Multivariate Statistical Analysis.'' Wiley, New York, 1958. 2nd ed., 1984.
23. Bendat, J. S., and Piersol, A. G. ''Random Data. Analysis and Measurement Procedures,'' 2nd ed. Wiley, New York, 1986.
24. Cooper, R., Osselton, J. W., and Shaw, J. C. ''EEG Technology,'' 3rd ed. Butterworths, London, 1980.
25. Dumermuth, G., and Molinari, L. Spectral analysis of EEG background activity. *In* ''Methods

of Analysis of Brain Electrical and Magnetic Signals'' (A. S. Gevins and A. Remond, Eds.), Handbook of EEG, rev. ser., Vol. I, pp. 85–130. Elsevier, Amsterdam, 1987.
26. Hinich, M. J., and Clay, C. S. The application of the discrete Fourier transform in the estimation of power spectra, coherence, and bispectra of geophysical data. *Rev. Geophys.* **6,** 347 (1968).
27. Jenkins, G. M., and Watts, D. G. ''Spectral Analysis and Its Applications.'' Holden–Day, San Francisco, 1968.
28. Otnes, R. K., and Enochson, L. ''Applied Time Series Analysis.'' Wiley–Interscience, New York, 1978.
29. Regan, D. ''Human Brain Electrophysiology. Evoked Potentials and Evoked Magnetic Fields in Science and Medicine.'' Elsevier, New York, 1989.
30. Harris, F. J. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* **66,** 51 (1978).
31. Fuller, W. A. ''Introduction to Statistical Time Series.'' Wiley, New York, 1976.
32. Koopmans, L. H. ''The Spectral Analysis of Time Series.'' Academic Press, New York, 1974.
33. Gasser, T., Bacher, P., and Möcks, J. Transformations towards the normal distribution of broad band spectral parameters of the EEG. *Electroenceph. Clin. Neurophysiol.* **53,** 119 (1982).
34. Brillinger, D. R. ''Time Series: Data Analysis and Theory,'' Holt, Rinehart, & Winston, New York, 1975.
35. Nuttall, A. H., and Carter, G. C. A generalized framework for power spectral estimation. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-28,** 334 (1980).
36. Gersch, W. Spectral analysis of EEGs by autoregressive decomposition of time series. *Math. Biosci.* **7,** 205 (1970).
37. Zetterberg, L. H. Estimation of parameters for a linear difference equation with application to EEG analysis. *Math. Biosci.* **5,** 227 (1969).
38. Kay, S. M., and Marple, S. L., Jr. Spectrum analysis: A modern perspective. *Proc. IEEE* **69,** 1380 (1981).
39. Marple, S. L., Jr. ''Digital Spectral Analysis with Applications.'' Prentice–Hall, Englewood Cliffs, NJ 1987.
40. Broersen, P. M. T., and Wensink, H. E. On finite sample theory for autoregressive model order selection. *IEEE Trans. Signal Process.* **41,** 194 (1993).
41. Hurvich, C. M., and Tsai, C. L. Regression and time series model selection in small samples. *Biometrika* **76,** 297 (1989).
42. Berk, K. N. Consistent autoregressive spectral estimates. *Ann. Statist.* **2,** 489 (1974).
43. Jennrich, R. I. Asymptotic properties of non-linear least squares estimators. *Ann. Math. Stat.* **40,** 633 (1969).
44. Kay, S. M. ''Modern Spectral Estimation: Theory and Application.'' Prentice–Hall, Englewood Cliffs, 1988.
45. Jansen, B. H. Analysis of biomedical signals by means of linear modeling. *CRC Crit. Rev. Biomed. Engrg.* **12,** 343 (1985).
46. Lopes DA Silva, F. H., and Mars, N. J. I. Parametric methods in EEG analysis. *In* ''Methods of Analysis of Brain Electrical and Magnetic Signals'' (A. S. Gevins and A. Remond, Eds.), Handbook of EEG, rev. ser., vol. 1, pp. 243–260. Elsevier, Amsterdam, 1987.
47. Blinowska, K. J., and Franaszczuk, P. J. A model of the generation of electrocortical rhythms. *In* ''Brain Dynamics 2'' (E. Basar and T. H. Bullock, Eds.), pp. 1992–201. Springer-Verlag, Berlin, 1989.
48. Franaszczuk, P. J., and Blinowska, K. J. Linear model of brain electrical activity–EEG as a superposition of damped oscillatory modes. *Biol. Cybern.* **53,** 19 (1985).
49. Franaszczuk, P. J., Mitraszewski, P., and Blinowska, K. J. FAD-parametric description of EEG time series. *Acta Physiol. Pol.* **40,** 418 (1989).
50. Blinowska, K. J., Franaszczuk, P. J., and Mitraszewski, P. A new method of presentation of the average spectral properties of the EEG time series. *Int. J. Biomed. Comput.* **22,** 97 (1988).
51. Wennberg, A., and Zetterberg, L. H. Application of a computer-based model for EEG analysis. *Electroenceph. Clin. Neurophysiol.* **31,** 457 (1971).
52. Wahba, G. Automatic smoothing of the log periodogram. *J. Am. Statist. Assoc.* **75,** 122 (1980).

53. Pawitan, Y., and O'Sullivan, F. Nonparametric spectral density estimation using penalized Whittle likelihood. *J. Am. Statist. Assoc.* **89,** 600 (1994).
54. Kooperberg, C., Stone, C. J., and Truong, Y. Logspline estimation of a possibly mixed spectral distribution. *J. Time Ser. Anal.* **16,** 359 (1995).
55. Cooley, W. W., and Lohnes, P. R. "Multivariate Procedures for the Behavioral Sciences." Wiley, New York, 1962.
56. Morrison, D. F. "Multivariate Statistical Methods," 3rd ed. McGraw–Hill, New York, 1990.
57. Johnson, R. A., and Wichern, D. W. "Applied Multivariate Statistical Analysis," 3rd ed. Prentice–Hall, Englewood Cliffs, 1992.
58. Shumway, R. H. "Applied Statistical Time Series Analysis." Prentice–Hall, Englewood Cliffs, 1988.
59. Chang, T. C., Krishnaiah, P. R., and Lee, C. J. Approximations to the distribution of the likelihood ratio statistics for testing the hypotheses on covariance matrices and mean vectors simultaneously. *In* "Applications of Statistics" (P. R. Krishnaiah, Ed.), pp. 97–108. North-Holland, Amsterdam, 1977.
60. Krishnaiah, P. R., and Lee, J. C. Likelihood ratio tests for mean vectors and covariance matrices. *In* "Handbook of Statistics" (P. R. Krishnaiah, Ed.), Vol. 1, pp. 513–570. North-Holland, Amsterdam, 1980.
61. Bartels, P. H., and Bartels, H. G. Classification strategies for topographic mapping data. *In* "Topographic Mapping of Brain Electrical Activity" (F. H. Duffy, Ed.), pp. 225–253. Butterworths, Boston, 1986.
62. Flury, B. "Common Principal Components and Related Multivariate Models." Wiley, New York, 1988.
63. Reyment, R., and Jöreskog, K. G. "Applied Factor Analysis in the Natural Sciences," 2nd ed. Cambridge Univ. Press, Cambridge, UK, 1993.
64. Velicer, W. F., and Jackson, D. N. Component analysis vs. common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behav. Res.* **25,** 1 (1990).
65. Zwick, W. R., and Velicer, W. F. A comparison of five rules for determining the number of components to retain. *Psychol. Bull.* **99,** 432 (1986).
66. Van Rotterdam, A. Limitations and difficulties in signal processing by means of the principal components analysis. *IEEE Trans. Biomed. Eng.* **BME-17,** 268 (1970).
67. Wastell, D. On the correlated nature of evoked brain activity: Biophysical and statistical considerations. *Biol. Psychol.* **13,** 51 (1981).
68. Silberstein, R. B., and Cadusch, P. J. Measurement processes and spatial principal components. *Brain Topogr.* **4,** 267 (1992).
69. Rawlings, R. R., Rohrbaugh, J. W., Begleiter, H., and Eckardt, M. J. Spectral methods for principal components analysis of event-related brain potentials. *Comput. Biomed. Res.* **19,** 497 (1986).
70. Valdes, P., Bosch, J., Grave, R., Hernandez, J., Riera, J., Pascual, R., and Biscay, R. Frequency domain models of the EEG. *Brain Topog.* **4,** 309 (1992).
71. Nunez, P. L. "Electric Fields of the Brain: The Neurophysics of EEG." Oxford Univ. Press, New York, 1981.
72. Skrandies, W., and Lehmann, D. Spatial principal components of multichannel maps evoked by lateral visual half-field stimuli. *Electroenceph. Clin. Neurophysiol.* **54,** 662 (1982).
73. Harner, R. N. Singular value decomposition—A general linear model for analysis of multivariate structure in the electroencephalogram. *Brain Topogr.* **3,** 43 (1990).
74. Shepard, R. A., Romney, A. K., and Nerlove, S. (Eds.), "Multidimensional Scaling: Theory and Applications in the Behavioral Sciences," Vol. I, "Theory." Seminar Press, New York, 1972.
75. Kruskal, J. B. and Wish, M. "Multidimensional Scaling." Sage, Beverly Hills, CA, 1978.
76. Schiffman, S. S., Reynolds, M. L., and Young, F. W. "Introduction to Multidimensional Scaling." Academic Press, New York, 1981.
77. Wish, M., and Carroll, J. D. Multidimensional scaling and its applications. *In* "Handbook of Statistics" (P. R. Krishnaiah and L. N. Kanal, Eds.), Vol. 2, pp. 317–345. North-Holland, Amsterdam, 1982.

78. McCarthy, G., and Wood, C. C. Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroenceph. Clin. Neurophysiol.* **62,** 203 (1985).
79. Faux, S. F., and McCarley, R. W. Analysis of scalp voltage asymmetries using Hotelling's T^2 methodology. *Brain Topogr.* **2,** 237 (1990).
80. Pascual-Marqui, R. D., Valdes-Sosa, P. A., and Alvarez-Amador, A. A parametric model for multichannel EEG spectra. *Int. J. Neurosci.* **40,** 89 (1988).
81. Goodman, N. R. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Stat.* **34,** 152 (1963).
82. Goodman, N. R. The distribution of the determinant of a complex Wishart distributed matrix. *Ann. Math. Stat.* **34,** 178 (1963).
83. Dubman, M. R., and Goodman, N. R. "Spectral Analysis of Multiple Time Series," North American Rockwell/Rocketdyne Technical Report R-8002. 1969.
84. Brillinger, D. R. The analysis of time series collected in an experimental design. *In* "Multivariate Analysis, III" (P. R. Krishnaiah, Ed.), pp. 241–256. Academic Press, New York, 1973.
85. Brillinger, D. R. Analysis of variance and problems under time series models, *In* "Handbook of Statistics" (P. R. Krishnaiah, Ed.), Vol. 1, pp. 237–278. Elsevier, Amsterdam, 1980.
86. Brillinger, D. R. The finite Fourier transform of a stationary process. *In* "Handbook of Statistics" (D. R. Brillinger and P. R. Krishnaiah, Eds.), Vol. 3, pp. 21–37. Elsevier, Amsterdam, 1983.
87. Brillinger, D. R., and Krishnaiah, P. R. (Eds.) "Time Series in the Frequency Domain," Handbook of Statistics, Vol. 3. North-Holland, Amsterdam, 1983.
88. Giri, N. On the complex analogues of T^2 and R^2 tests. *Ann. Math. Stat.* **36,** 664 (1965).
89. Khatri, C. G. Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Stat.* **36,** 98 (1965).
90. Krishnaiah, P. R. Some recent developments on complex multivariate distributions. *J. Multivariate Anal.* **6,** 1 (1976).
91. Krishnaiah, P. R., Lee, J. C., and Chang, T. C. The distributions of the likelihood ratio statistics for tests of certain covariance structures of complex multivariate normal populations. *Biometrika* **63,** 543 (1976).
92. Krishnaiah, P. R., Lee, J. C., and Chang, T. C. Likelihood ratio tests on covariance matrices and mean vectors of complex multivariate normal populations and their applications in time series. *In* "Handbook of Statistics" (D. R. Brillinger and P. R. Krishnaiah, Eds.), Vol. 3, pp. 439–476. Elsevier, Amsterdam, 1983.
93. Shumway, R. H. Discriminant analysis for time series. *In* "Handbook of Statistics" (P. R. Krishnaiah and L. N. Kanal, Eds.), Vol. 2, pp. 1–46. North Holland, Amsterdam, 1982.
94. Rawlings, R., Eckardt, M., and Begleiter, H. Multivariate time series discrimination in the spectral domain. *Comput. Biomed. Res.* **17,** 352 (1984).
95. Walter, D. O., and Adey, W. R. Analysis of brain-wave generators as multiple statistical time-series. *IEEE Trans. Biomed. Eng.* **BME-12,** 8 (1965).
96. Bloomfield, P. "Fourier Analysis of Time Series: An Introduction." Wiley–Interscience, New York, 1976.
97. Childers, D. G. Evoked responses: Electrogenesis, models, methodology, and wavefront reconstruction and tracking analysis. *Proc. IEEE* **65,** 611 (1977).
98. Carter. G. C., Knapp, C. H., and Nuttall, A. H. Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *IEEE Trans. Audio Electroacoust.* **AU-21,** 337 (1973).
99. Carter, G. C., and Knapp, C. H. Coherence and its estimation via the partitioned modified Chirp-Z transform. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-23,** 257 (1975).
100. Levine, P. H., Herbert, J. R., Haynes, C. A., and Strobel, C. EEG coherence during the Transcendental Meditation technique. *In* "Scientific Research on the Transcendental Meditation Program: Collected Papers," Vol. 1. 1976.
101. Tick, L. J. Estimation of coherency. In "Spectral Analysis of Time Series" (B. Harris, Ed.), pp. 133–152. Wiley, New York, 1967.
102. Carter, G. C. Bias in magnitude-squared coherence estimation due to misalignment. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-28,** 97 (1980).

103. Tucker, D. M., Roth, D. L., and Bair, T. B. Functional connections among cortical regions: Topography of EEG coherence. *Electroenceph. Clin. Neurophysiol.* **63,** 242 (1986).
104. Franaszczuk, P. J., Blinowska, K. H., and Kowalczyk, M. The application of parametric multichannel spectral estimates in the study of electrical brain activity. *Biol. Cybern.* **51,** 239 (1985).
105. Lopes da Silva, F. H., Vos, J. E., Mooibroek, J., and Van Rotterdam, A. Relative contribution of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analyses. *Electroenceph. Clin. Neurophysiol.* **50,** 449 (1980).
106. Gersch, W. Causality or driving in electrophysiological signal analysis. *Math. Biosci.* **14,** 177 (1972).
107. Gersch, W., and Tharp, B. R. Spectral regression—Amount of information analysis of seizures in humans. *In* "Quantitative Analytic Studies in Epilepsy" (P. Kellaway and I. Petersen, Eds.), pp. 509–532. Raven Press, New York, 1976.
108. Fisher, R. A. The general sampling distribution of the multiple correlation coefficient. *Proc. R. Soc. Ser. A.* **121,** 654 (1928).
109. Amos, D. E., and Koopmans, L. H. "Tables of the Distribution of the Coefficient of Coherence for Stationary Bivariate Gaussian Processes." Washington, D.C.: Office of Technical Services, Dept. of Commerce, 1963.
110. Alexander, M. J., and Vok, C. A. "Tables of the Cumulative Distribution of Sample Multiple Coherence." Rocketdyne Research Report RR63-37, 1963.
111. Benignus, V. A. Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform. *IEEE Trans. Audio Electroacoust.* **AU-17,** 145 (1969).
112. Benignus, V. A. Estimation of coherence spectrum of non-Gaussian time series populations. *IEEE Trans. Audio Electroacoust.* **AU-17,** 198 (1969).
113. Nuttall H., AND CARTER, G. C. Bias of the estimate of magnitude-squared coherence. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-24,** 582 (1976).
114. Lopes Silva, F. H., Hoeks, A., Lierop, T. H. M. T. van, Schrijer, C. F., and Storm van Leeuwen, W. Confidence intervals of spectra and coherence functions—Their relevance for quantifying thalamo-cortical relationships. *In* "Quantitative Analysis of the EEG" (M. Matejcek and G. K. Schenk, Eds.), pp. 437–449. AEG Telefunken, Constanz, 1975.
115. Cooper, R. Measurement of time and phase relationships of the EEG. *In* "CEAN: Computerized EEG Analysis" (G. Dolce and H. Künkel, Eds.), pp. 85–97. Fischer, Stuttgart, 1975.
116. Vos, J. E., Scholten, C. A., and van Woerden, H. H. Phase spectra: Estimation and interpretation. *In* "Quantitative Analysis of the EEG" (M. Matejcek and G. K. Schenk, Eds.), pp. 409–419. AEG Telefunken, Konstanz, 1975.
117. Oppenheim, A. V., and Shafer, R. W. "Discrete-Time Signal Processing." Prentice–Hall, Englewood Cliffs, 1989.
118. Gotman, J. Measurement of small time differences between EEG channels: Method and application to epileptic seizure propagation. *Electroenceph. Clin. Neurophysiol.* **56,** 501 (1983).
119. Ktonas, P. Y., and Mallart, P. Y. Estimation of time delay between EEG signals for epileptic focus localization: Statistical error considerations. *Electroenceph. Clin. Neurophysiol.* **78,** 105 (1991).
120. Harris, B., Gath, I., Rondouin, G., and Feuerstein, C. On time delay estimation of epiletic EEG. *IEEE Trans. Biomed. Eng.* **41,** 820 (1994).
121. Adey, W. R., and Walter, D. O. Application of phase detection and averaging techniques in computer analysis of EEG records in the cat. *Exp. Neurology.* **7,** 186 (1963).
122. Goodman, N. R. Measuring amplitude and phase. *J. Franklin Inst.* **270,** 437 (1960).
123. Zar, J. H. "Biostatistical Analysis," 2nd ed. Prentice–Hall, Englewood Cliffs, NJ, 1984.
124. Mardia, K. V. "Statistics of Directional Data." Academic Press, New York, 1972.
125. Strasburger, H. The analysis of steady state evoked potentials revisited. *Clin. Vis. Sci.* **1,** 245 (1987).
126. Gasser, T. A quantitative topographic component analysis for the EEG at rest. *In* "Statistics and Topography in Quantitative EEG" (D. Sampson-Dollfus, Ed.), pp. 139–149. Elsevier, Paris, 1988.
127. Galbraith, G. C. The effect of prior EEG "coupling" upon the visual evoked response. *IEEE Trans. Biomed. Eng.* **BME-14,** 223 (1967).

128. Gersch, W. Non-stationary multichannel time series analysis. *In* "Methods of Analysis of Brain Electrical and Magnetic Signals" (A. S. Gevins and A. Remond, Eds.), Handbook of EEG, rev. ser., Volume I, pp. 261–296. Elsevier, Amsterdam, 1987.
129. Gersch, W., and Yonemoto, J. Parametric times series models for multivariate EEG analysis. *Comput. Biomed. Res.* **10,** 113 (1977).
130. Gersch, W., Yonemoto, J., and Naitoh, P. Automatic classification of multivariate EEG's using an amount of information measure and the eigenvalues of parametric time series model features. *Comput. Biomed. Res.* **10,** 297 (1977).
131. Akaike, H. Recent development of statistical methods for spectrum estimation. *In* "Recent Advances in EEG and EMG Data Processing" (N. Yamaguchi and K. Fujisawa, Eds.), pp. 68–78. Elsevier/North-Holland, Amsterdam, 1981.
132. Jones, R. H. Identification and autoregressive spectrum estimation. *IEEE Trans. Automat. Control* **AC-19,** 894 (1974).
133. Kaminski, M. J., and Blinowska, K. J. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* **65,** 203 (1991).
134. Gath, I., Feuerstein, C., Pham, D. T., and Rondouin, G. On the tracking of rapid dynamic changes in seizure EEG. *IEEE Trans. Biomed. Eng.* **39,** 952 (1992).
135. Franaszczuk, P. J., Bergey, G. K., and Kaminiski, M. J. Analysis of mesial temporal seizure onset and propagation using the directed transfer function method. *Electroenceph. Clin. Neurophysiol.* **91,** 413 (1994).
136. Oken, B. S., and Chiappa, K. H. Statistical issues concerning computerized analysis of brainwave topography. *Ann. Neurol.* **19,** 493 (1986).
137. Abt, K. Statistical aspects of neurophysiologic topography. *J. Clin. Neurophysiol.* **7,** 519 (1990).
138. Kiers, H. A. L., and ten Berge, J. M. F. Alternating least square algorithms for simultaneous components analysis with equal component weight matrices in two or more populations. *Psychometrika* **54,** 467 (1989).
139. Kiers, H. A. L., and ten Berge, J. M. F. Hierarchical relations between methods for simultaneous component analysis and a technique for rotation to a simple simultaneous structure. *Brit. J. Math. Stat. Psychol.* **47,** 109 (1994).
140. Millsap, R. E., and Meredith, W. Component analysis in cross-sectional and longitudinal data. *Psychometrika* **53,** 123 (1988).
141. Jöreskog, K. G. A general method of analysis for covariance structures. *Biometrika* **57,** 239 (1970).
142. Jöreskog, K. G. Simultaneous factor analysis in several populations. *Psychometrika* **36,** 409 (1971).
143. Sörbom, D. A general method for studying differences in factor means and factor structures between groups. *Brit. J. Math. Stat. Psychol.* **27,** 229 (1974).
144. Jöreskog, K. G., and Sörbom, D. "LISREL 7—A Guide to the Program and Applications, 2nd ed. SPSS Publications, Chicago, 1989.
145. Jöreskog, K. G., and Sörbom, D. "LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language." Erlbaum, Hillsdale, NJ, 1993.
146. Byrne, B. M., Shavelson, R. J., and Muthern, B. Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychol. Bull.* **105,** 456 (1989).
147. McIntosh, A. R., and Gonzalez-Lima, F. Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping* **2,** 2 (1994).
148. Tucker, D. M., and Roth, D. L. Factoring the coherence matrix: Patterning of the frequency-specific covariance in a multichannel EEG. *Psychophysiology* **21,** 228 (1984).
149. Harman, H.H. "Modern Factor Analysis," 3rd ed. Univ. of Chicago Press, Chicago, 1976.
150. Gorsuch, R. L. "Factor Analysis," 2nd ed. Erlbaum, Hillsdale, NJ, 1983.
151. Mulaik, S. A. "The Foundations of Factor Analysis." McGraw–Hill, New York, 1972.
152. Tucker, L. R. The extension of factor analysis to three-dimensional matrices. *In* "Contributions to Mathematical Psychology" (N. Frederiksen and H. Gulliksen, Eds.), Holt Rinehart & Winston, New York, 1964.

153. Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika* **31,** 279 (1966).
154. Tucker, L. R. Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika* **37,** 3 (1972).
155. Law, H. G., Snyder, C. W., Hattie, J. A., and McDonald, R. P. ''Research Methods for Multimode Data Analysis.'' Praeger, New York, 1984.
156. Field, A. S., and Graupe, D. Topographic components analysis of evoked potentials: estimation of model parameters and evaluation of parameter uniqueness. *J. Biomed. Eng.* **12,** 287 (1990).
157. Field, A. S., and Graupe, D. Topographic component (parallel factor) analysis of multichannel evoked potentials: Practical issues in trilinear spatiotemporal decomposition. *Brain Topogr.* **3,** 407 (1991).
158. Harshman, R. A., and Lundy, M. E. PARAFAC: Parallel factor analysis. *Comput. Stat. Data Anal.* **18,** 39 (1994).
159. Möcks, J. Topographic components model for event-related potentials and some biophysical considerations. *IEEE Trans. Biomed. Eng.* **BME-35,** 482 (1988).
160. Dunseath, W. J. R., and Kelly, E. F. PC-based data-acquisition system for high-resolution EEG. *IEEE Trans. Biomed. Eng.* **42,** 1212 (1995).
161. Lenz, J. E. Jr., and Kelly, E. F. Computer-based calibration and measurement of an EEG data-acquisition system. *IEEE Trans. Biomed. Eng.* **BME-28,** 396 (1981).
162. Lehmann, D. Principles of spatial analysis. *In* ''Methods of Analysis of Brain Electrical and Magnetic Signals'' (A. S. Gevins and R. Remond, Eds.), Handbook of EEG, rev. ser., Vol. I, pp. 309–354. Elsevier, Amsterdam, 1987.
163. Walter, D. O., Etevenon, P., Pidoux, B., Tortrat, D., and Guillou, S. Computerized topo-EEG spectral maps: Difficulties and perspectives. *Neuropsychobiology* **11,** 264 (1984).
164. Fein, G., Raz, J., Brown, F. F., and Merrin, E. L. Common reference coherence data are confounded by power and phase effects. *Electroenceph. Clin. Neurophysiol.* **69,** 581 (1988).
165. Tomberg, C., Desmedt, J. E., and Ozaki, I. Right or left ear reference changes the voltage of frontal and parietal somatosensory evoked potentials. *Electroenceph. Clin. Neurophysiol.* **80,** 504 (1991).
166. Stephenson, W. A., and Gibbs, F. A. A balanced noncephalic reference electrode. *Electroenceph. Clin. Neurophysiol.* **3,** 237 (1951).
167. Pernier, J., Perrin, F., and Bertrand, O. Scalp current density fields: Concept and properties. *Electroenceph. Clin. Neurophysiol.* **69,** 385 (1988).
168. Biggins, C. A., Fein, G., Raz, J., and Amira, A. Artifactually high coherences result from using spherical spline computation of scalp current density. *Electroenceph. Clin. Neurophysiol.* **79,** 413 (1991).
169. Biggins, C. A., Ezekiel, F., and Fein, G. Spline computation of scalp current density and coherence: A reply to Perrin. *Electronceph. Clin. Neurophysiol.* **83,** 172 (1992).
170. Perrin, F. Comments on article by Biggins *et al. Electroenceph. Clin. Neurophysiol.* **83,** 171 (1992).
171. Pascual-Marqui, R. D. The spherical spline Laplacian does not produce artifactually high coherences: Comments on two articles by Biggins *et al. Electroenceph. Clin. Neurophysiol.* **87,** 62 (1993).
172. Mardia, K. V. Tests of univariate and multivariate normality. *In* ''Handbook of Statistics'' (P. R. Krishnaiah, Ed.), Vol. 1, pp. 279–320. North-Holland, Amsterdam, 1980.
173. Ito, P. K. Robustness of ANOVA and MANOVA test procedures. *In* ''Handbook of Statistics'' (P. R. Krishnaiah, Ed.), Vol. 1, pp. 199–236. North-Holland, Amsterdam, 1980.
174. Akaike, H. On the statistical estimation of the frequency response function of a system having multiple input. *Ann. Inst. Stat. Math.* **17,** 185 (1965).
175. Blair, R. C., and Karniski, W. Distribution-free statistical analyses of surface and volumetric maps. *In* ''Functional Neuroimaging: Technical Foundations'' (R. W. Thatcher, M. Hallett, T. Zeffiro, E. R. John, and M. Huerta, Eds.), pp. 19–28. Academic Press, Orlando, FL, 1994.
176. Barlow, J. S. Methods of analysis of nonstationary EEGs with emphasis on segmentation techniques. A comparative review, *J. Clin. Neurophysiol.* **2,** 267 (1985).
177. Elul, R. The genesis of the EEG. *Int. J. Neurobiol.* **15,** 227 (1972).

178. Dewan, E. M. ''Non-linear Cross-Spectral Analysis and Pattern Recognition,'' Physical and Mathematical Sciences Research Papers, No. 367, AFCRL-69-0026. Office of Aerospace Research, USAF, 1969.
179. Kullback, S. ''Information Theory and Statistics.'' Wiley, New York, 1969.
180. Mars, N. J. I., and Lopes da Saliva, F. H. EEG analysis methods based on information theory. *In* ''Methods of Analysis of Brain Electrical and Magnetic Signals'' (A. S. Gevins and R. Remond, Eds.), Handbook of EEG, rev. ser., Vol. I, pp. 297–307. Elsevier, Amsterdam, 1987.
181. Nikias, C. L., and Petropulu, A. P. ''Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework.'' PTR Prentice–Hall, Englewood Cliffs, 1993.
182. Saltzberg, B., Burton, W. D., Burch, N. R., Fletcher, J., and Michaels, R. Electrophysiological measures of regional neural interactive coupling. Linear and non-linear dependence relationships among multiple channel electroencephalographic recordings. *Int. J. Biomed. Comput.* **18,** 77 (1986).
183. Kitagawa, G., and Akaike, H. A procedure for modeling of non-stationary time series. *Ann. Inst. Stat. Math. B.* **30,** 351 (1978).