RESEARCH ARTICLE

Definitions of state variables and state space for brain–computer interface

Part 2. Extraction and classification of feature vectors

Walter J. Freeman

Abstract The hypothesis is proposed that the central dynamics of the action–perception cycle has five steps: emergence from an existing macroscopic brain state of a pattern that predicts a future goal state; selection of a mesoscopic frame for action control; execution of a limb trajectory by microscopic spike activity; modification of microscopic cortical spike activity by sensory inputs; construction of mesoscopic perceptual patterns; and integration of a new macroscopic brain state. The basis is the circular causality between microscopic entities (neurons) and the mesoscopic and macroscopic entities (populations) self-organized by axosynaptic interactions. Self-organization of neural activity is bidirectional in all cortices. Upwardly the organization of mesoscopic percepts from microscopic spike input predominates in primary sensory areas. Downwardly the organization of spike outputs that direct specific limb movements is by mesoscopic fields constituting plans to achieve predicted goals. The mesoscopic fields in sensory and motor cortices emerge as frames within macroscopic activity. Part 1 describes the action–perception cycle and its derivative reflex arc qualitatively. Part 2 describes the perceptual limb of the arc from microscopic MSA to mesoscopic wave packets, and from these to macroscopic EEG and global ECoG fields that express experience-dependent knowledge in successive states. These macroscopic states are conceived to embed and control mesoscopic frames in premotor and motor cortices that are observed in local ECoG and LFP of frontoparietal areas. The fields sampled by ECoG and LFP are conceived as local patterns of neural activity in which trajectories of multiple spike activities (MSA) emerge that control limb movements. Mesoscopic frames are located by use of the analytic signal from the Hilbert transform after band pass filtering. The state variables in frames are measured to construct feature vectors by which to describe and classify frame patterns. Evidence is cited to justify use of linear analysis. The aim of the review is to enable researchers to conceive and identify goal-oriented states in brain activity for use as commands, in order to relegate the details of execution to adaptive control devices outside the brain.

Keywords Beta activity β · Brain–computer interface BCI · Electrocorticogram ECoG · Epsilon activity ε · Gamma activity γ · Hilbert transform · Local field potential LFP · Multiple spike activity MSA · Stationarity

http://sulcus.berkeley.edu

W. J. Freeman (✉)
Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, CA 94720-3206, USA
e-mail: dfreeman@berkeley.edu

Introduction

Fundamental in brain dynamics are the transactions between levels in both directions in all cortices. The upward transaction predominates in sensory cortices from microscopic sensory-driven spike input to mesoscopic wave patterns in perception; the downward transaction predominates in motor cortices from mesoscopic wave patterns constituting goal states to microscopic spike output patterns. In accordance with Haken's (1983) "slaving" principle of circular causality the two transactions are opposite sides of the same

coin; interactions among neurons create patterns of activity that are called ''order parameters'' because they impose structure on the activities of the neurons creating them (Haken 2006). Every area of cortex maintains background activity at all levels by continuous synaptic interaction regulated by brain stem neuromodulators. Input to sensory cortex by sensory-driven activity is identified and measured with spikes; output of cortex is inferred from field potentials, ECoG and LFP. In motor cortex the output microscopic activity is identified with action-related spikes. The input to cortex that guides the order parameters is inferred from the field potentials. These are the main targets for measurement and interpretation. Studies summarized in Part 1 show that macroscopic and mesoscopic patterns exist in time frames in which linearity and stationarity may hold to a good approximation. Already linear analysis (Freeman 1975/2004; Wright and Liley 1996; Basar 1998; Liley et al. 1999; Gordon 2000; Robinson et al. 2001; Haken 2006) has been used as a tool to extract the order parameters.

The correlation between MSA trajectories and limb trajectories is the main focus of current research on BMI (e.g., Chapin et al. 1998; Sanchez et al. 2004; Carmena et al. 2005; Hochberg et al. 2006). However, microscopic spikes, mesoscopic fields and macroscopic EEG differ not only in the forms of state variables but also in the levels of their correlates over the five central steps of the action–perception cycle (Part 1, Section 6). While much is known about neural integrative processes by which percepts emerge from sensory input, little is known about neural differentiating processes by which concepts and goals formulate actions. From considerations of volume conduction and global integration in Part 1, it is clear that every extracellular recorded signal contains contributions from all levels, which can account for the fact that even a single channel can support rudimentary BMI. The problem addressed in Part 2 is how to formulate hypotheses about downward causation in speciation of motor output from goal states. Guidelines in the search for solutions are derived from studies of upward causation in generalization of perceptual states from sensory input.

The most difficult experimental task is to distinguish between mesoscopic and macroscopic components of the field potentials. They mix in two ways, one by summation of their extracellular electric currents in the volume conductor, the other by multitasking of neurons engaged simultaneously in multiple populations. The key difference is in size. Macroscopic fields occupy large areas of cortex and can only be seen in the coordinated activity sampled by large arrays of electrodes that are closely spaced yet widely distributed. Large arrays are necessary. Mesoscopic fields are more local, corresponding to signals from cortical modules. These require smaller arrays at very high density. Macroscopic fields tend to last longer, recur at lower theta rates, and have carrier frequencies in the beta range, compared with mesoscopic fields having carrier frequencies in the gamma range, recurrence at high theta rates, and shorter durations (Freeman 2005). Preprocessing techniques are the same for both mesoscopic and macroscopic signals, requiring low-pass spatial filters and wide-band pass temporal filters, compared with the high-pass spatial filters and narrow-band basis functions commonly used to isolate microscopic components.

Use of the analytic signal to locate frames

The strategy proposed here is linear decomposition of the data from multiple electrodes in arrays, first into broad frequency bands by band pass filtering using the Fourier transform and then into analytic amplitude and phase values by use of the Hilbert transform (Appendix 1). The hypothesis is adopted, based on studies of perception (Freeman 2004a, b, 2005, 2006), that the behavioral correlates of wave activity are carried in frames by spatial patterns of amplitude modulation of self-organized oscillations that serve as carrier waves. The carrier frequencies are predominantly in the beta and gamma ranges, with small drift within frames and large jumps in instantaneous frequency between frames. The frames recur irregularly at rates in the theta and alpha bands. The pass band that is required for use of the Hilbert transform must be broad enough to include the range of variation of the carrier waves over successive frames in motor control but not so broad as to include multiple coexisting carrier waves at different frequencies. Optimization of multiple pass bands for a given data set can be based on maximizing the cross-spectral peak of shared power between the several gating and carrier frequency ranges (Freeman 2004a, 2005). The optimized pass bands in appear to conform to the clinical bands for beta and gamma oscillations, which reflect the value of long empirical experience.

With the Hilbert transform the analytic amplitude, $A_j(t)$, and the analytic phase, $\phi_j(t)$, are calculated for every channel, $j = 1,..., n$ channels in each selected carrier frequency range from the same time series. The spatial pattern of amplitude at each digitizing time step, t, is expressed by a normalized $n \times 1$ feature vector, $\mathbf{A}(t)$, with mean length, $A(t)$, and by a point in

n-space. $\mathbf{A}(t)$ is adopted as the order parameter (Haken 1983), because it gives direct expression to the high-order textured pattern of amplitude created by interactions from experience stored by synaptic modification in attractor landscapes in the brain. The rate of change in the order parameter, $D_e(t)$ (from Eq. (5) in Appendix 1) is approximated by the Euclidean distance in *n*-space between successive points in *n*-space with digitizing steps after normalization (dividing the values in each feature vector by the standard deviation of the *n* values). Large differences show wide-ranging trajectories in *n*-state space during periods of instability. Small values of $D_e(t)$ show clustering of points that indicate the location in *n*-space of an attractor, which when accessed gives a frame with a stable spatial pattern.

These clusters are one of the two bases for identification of frames (Ohl et al. 2001); the other basis is the analytic phase (Freeman and Rogers 2002). The *n* values of analytic phase from a 2-D array, $n^{.5} \times n^{.5}$, form a phase surface, which is fitted with a 2-D conic spatial basis function at each time step (Freeman 2004b). The location and sign of the apex of the cone vary unpredictably at times of rapid change in spatial amplitude pattern between frames (high values of $D_e(t)$). Within frames the sign (maximum lead or lag) and spatial phase gradient are fixed, and $D_e(t)$ varies less than the distance between electrodes, which is the limit of spatial resolution. Successive values of the slope of the cone are averaged across the frame to give the phase gradient, $\gamma(t_c)$, in rad/mm where t_c is the center time of the frame. The phase difference from the preceding value, $\Delta\phi_j(t) = \phi_j(t) - \phi_j(t-1)$, $j = 1,...,$ *n*, after unwrapping approximates the rate of change in phase at each time step. Averaging over *n* and division by the duration of the digitizing step in s gives the frame frequency, $\omega(t)$ in rad/s; division by 2π gives frequency in Hz. The ratio of analytic frequency in rad/s to phase gradient in rad/m gives phase velocity in m/s for comparison with spike conduction velocities. The spatial standard deviation, $SD_X(t)$, of the phase differences, $\Delta\phi_j(t)$, measures the variance of the spatial phase pattern at each time step.

An example in Fig. 1 shows the analytic phase differences, $\Delta\phi_j(t)$, $j = 1,...,64$ in a time period of 400 ms. The ECoG was recorded from the visual cortex of a rabbit trained to respond to a visual conditioned stimulus and band pass filtered in the gamma range (20–80 Hz) prior to application of the Hilbert transform. The spikes in phase differences tend to occur almost simultaneously over the whole array. Between the spikes the rate of change converges to a constant value with a low spatial variance, indicating
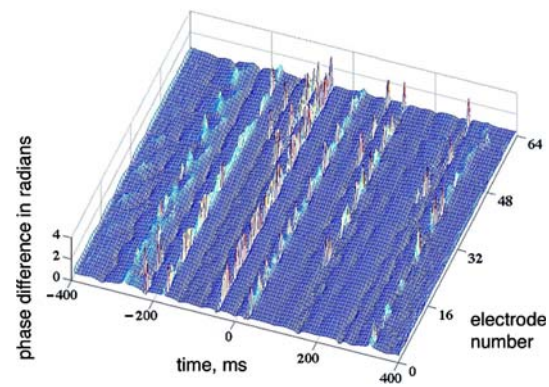

Fig. 1 The raster plot shows the successive differences of the unwrapped analytic phase, $\Delta p_j(t)$, changing with time (left abscissa) and channel (right abscissa). The 8 columns of 8 rows are aligned to show the near-coincidence of the sudden jumps and dips given by forward and backward phase slip (Fig. 6, D in Appendix 2) with near-zero lag across channels. Phase slip shows the incidence of widespread state transitions; plateaus show epochs of near-stationarity of visual cortical dynamics. From Freeman (2004a)

the relative stationarity defined by the constancy of frequency. Nonstationarity appears in the change of frequency from each frame to the next; the average frequency difference between frames was 7 Hz (about ±20% of the peak spectral frequency of the pass band, 33 ± 7 Hz). Figure 2 shows the spatial standard deviation, $SD_X(t)$, giving a sequence of spikes that indicate the time locations of state transitions. During those jumps the analytic amplitude ($A(t)$, gray curve) tends to low values; within the frames $A(t)$ tends to high values that are accompanied by low values of $D_e(t)$ (not shown). This configuration shows that the frame is characterized by high intensity activity with a nearly constant carrier frequency and a stable spatial pattern
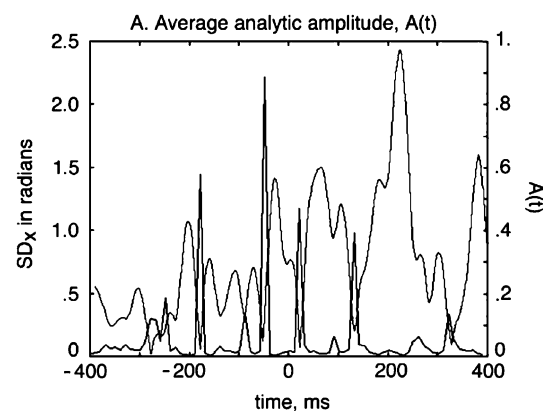

Fig. 2 The gray curve shows the analytic amplitude, $A(t)$. The black spikes show the spatial standard deviation of the analytic phase differences, $SD_X(t)$, in Fig. 1. See rotating vector in the polar plot (Fig. 6, B, Appendix 2). From Freeman (2004a)

of amplitude and phase. It forms the archetype for a linear, stationary mesoscopic pattern of ensemble dynamics, likely with superposition of multiple components (Haken 2006).

In brief, within each frame there may emerge a spatial pattern of amplitude that specifies the normalized feature vector, $\mathbf{A}(t)$, with its point in n-space. A sequence of points gives a trajectory. A cluster of points indicates formation and maintenance by an attractor of either a stable spatial pattern over the duration of a frame or a delimited trajectory within the attractor. The feature vector at the peak amplitude of $A(t)$ in a frame best specifies the spatial pattern of amplitude modulation governed by the attractor.

In extrapolation of these results in sensory cortices to motor cortices, the brief spikes of high $SD_X(t)$ may show the state transitions by which the intentional control systems of the neocortex organize and dissolve the successive neural activity patterns that execute intended actions. The use of the LFP and ECoG for BCI will require determination of the multiple frequency ranges for the carrier waves, whose amplitude modulations give the spatial amplitude patterns of dendritic currents and spike densities by which commands are transmitted to other areas of cortex and to the brain stem. The LFP and ECoG thereby may temporally and spatially localize the stationary frames in different frequency bands prior to classification. Newer evidence (Gonzalez et al. 2006) suggests that epsilon activity can also be identified in scalp EEG, but this range has been explored only in respect to basic properties of stability (Freeman 1974).

Use of the analytic signal to classify frames

The primary aim in BCI is to use n channels of recording at m time steps to acquire $m \times n$-dimensional feature vectors with high information content as control signals. Each $m \times n$ feature vector is derived by decomposition of an array of n electrical signals into a set of n scalar values that specifies a set of m points in $m \times n$-space (Freeman 2005). The set of $m \times n$ concomitant feature vectors can be conceived as time functions that specify m frames of a trajectory in n-space during the performance of an operant. The dimensions of the $m \times n$-space are determined in part by the empirical number of components derived from temporal, spatial and spectral decomposition, and in part by the conceptual framework in which an intentional action is conceived, constructed and described.

An example of mesoscopic frame classification from sensory cortex is shown in Fig. 3A. A rabbit was trained to discriminate each of two CS in a selected modality (here vision), one reinforced (CS+) and the other not (CS–) in a classic aversive conditioning paradigm. There appeared in the 8×8 ECoG from the visual cortex to which each CS was directed a sequence of 3 frames (Fig. 3A); in each frame the carrier frequency approached a constant value in the gamma range and the AM pattern stabilized as shown by a fall in $D_e(t)$. The order parameter, $\mathbf{A}^2(t)$, in each AM pattern gave a feature vector specifying a point in 64-space.

The frames in succession differed from each to the next within each trial but tended to cluster across the
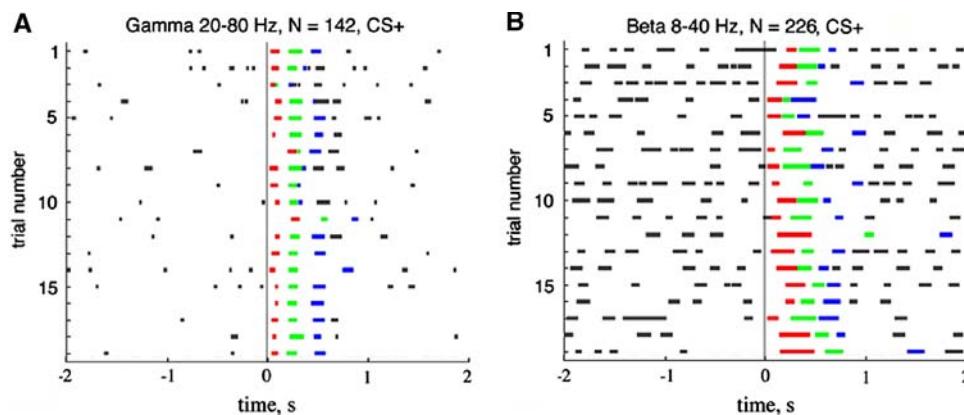

Fig. 3 The latency and duration for each frame is shown by a colored bar as measured by an index for pragmatic information, $H_e(t)$, derived from the analytic amplitudes of the 64 signals at each digitizing time step of 2 ms (see legend of Fig. 4, A). The first 3 frames in a set of 20 trials were labeled by color: first, red; second; green; third, blue. (**A**) Frame latencies from onset of CS+ (with reinforcement) at $t = 0$ and their durations were represented by colored bars for gamma carrier frequencies after classifier-directed optimization of the threshold for the pragmatic information index. A similar plot held for frames after unreinforced CS–. Gamma bursts occurred preferentially in the first half second after CS onset. (**B**) The frames with beta carrier frequencies had longer start latencies, longer durations, and larger diameters than those for gamma frames. At optimized thresholds for the pragmatic information index the frames were equally likely to occur before and after CS onsets. From Freeman (2005)

20 trials. Therefore a 3×64 ($m \times n$) normalized feature vector was constructed to represent each stimulus category in 192-space for each trial. The hyperspace clusters were mapped into 2-space for display of two clusters of points that represented the differing spatial patterns corresponding to the discriminated stimuli (Fig. 4B). Optimal classification (Fig. 4A) was found by dividing the feature vector at each time step by the rate of change, $D_e(t)$, giving the pragmatic information index, $H_e(t)$ (Freeman 2005). Similar results were found in frames with carrier frequencies in the beta range (Fig. 3B) but with significant differences as described below.

The variance of the clusters was measured by the SD in 2-space (circles) around the projections of the two centers of gravity (Fig. 4B), and the goodness of classification was assayed by the shortest Euclidean distance from each point to the nearest center of gravity. The variance in each cluster was attributed to noise in the system, to errors of measurement, and to incremental changes in the dynamics from on-going learning with each new trial. The clusters displayed AM patterns of activity that manifest nonconvergent attractors in a high-dimensional landscape of basins of attraction. The trajectories and clusters appeared to constitute chaotic itinerancy among dynamical states in neocortical systems (Tsuda 2001). Details of the statistical techniques for spatial category classification have been

described (Barrie et al. 1999; Kozma and Freeman 2001; Ohl et al. 2001; Freeman, 2005, 2006).

An important distinction must be drawn between the feature vectors from MSA at the microscopic, sensorimotor level and the LFP and ECoG at the mesoscopic, perceptual level. Owing to their inherently high degree of spatial localization, the feature vectors based in spikes give extra weight in classification to the sites of high spike frequency (locally in time) and to the specific electrodes (locally in space). Deletion of channels yielding few spikes has little effect on the rates of correct classification, so the MSA reveals high temporospatial contrast in the classificatory value of the information. In contrast, the channels of mesoscopic feature vectors contribute equally to correct classification irrespective of their amplitudes or locations. No channel is any more or less important for classification than any other channel, which means that perceptual information is spatially distributed uniformly (Freeman and Viana Di Prisco 1986; Barrie et al. 1996; Ohl et al. 2001).

An example is shown of macroscopic pattern classification in Fig. 5, in which 64 electrodes were placed in groups of up to 16 on the visual, auditory, somatomotor and entorhinal cortices and the olfactory bulb of cats and rabbits (Freeman et al. 2003) trained to discriminate auditory or visual stimuli in an operant appetitive paradigm. An index (Tass et al. 1999) of the
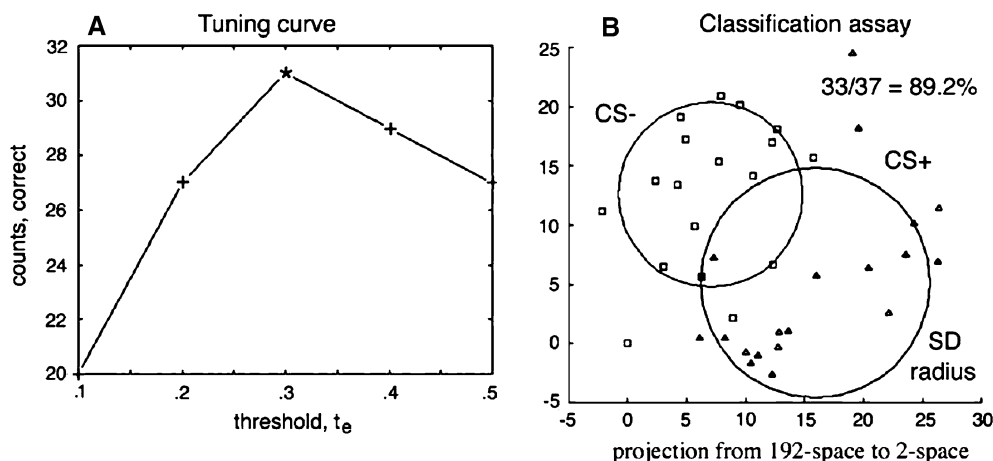

Fig. 4 (**A**) The spatial AM pattern of each frame was expressed by a feature vector given by $m \times n = 3 \times 64$ values of normalized index, $H_e(t)$ = amplitude squared divided by the rate of change in the order parameter, $A_{i,j}^2(t)/D_{e,I,j}(t), 1 = 1.3; j = 1, 64$, at the time point, t, of the maximal mean amplitude, $A(t)$, in each frame. A threshold, t_e, for $H_e(t)$ was set by visual inspection of the time series of $H_e(t)$ to demarcate the beginnings and endings of frames (Fig. 3). The $m \times n$ feature vectors from 20 CS+ trials with reinforcement and 20 CS– trials without reinforcement were classified with respect to CS+/CS–. The classification level was calculated

repeatedly as the threshold was varied in search of an optimal value. (**B**) The multidimensional scaling technique of nonlinear mapping (Sammon 1969) projected clusters from n-space into 2-space, optimizing their separation while preserving the relative distances between all the data points. Two clusters were specified in this example: the 1×192 feature vector from the first three 1×64 feature vectors in the CS+ trials, and the 1×192 feature vector from the first three 1×64 feature vectors in the CS–. The circles representing the standard deviations (SD) of the clusters were calculated in the display plane. From Freeman (2005)
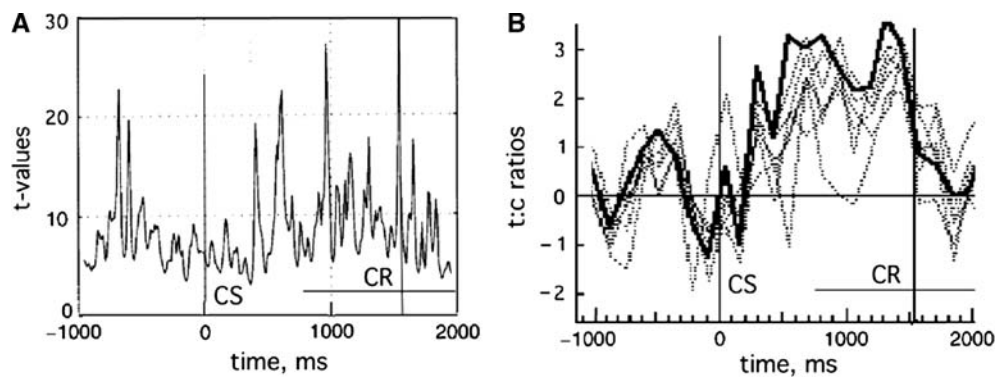
Fig. 5 (**A**) An index of synchrony revealed intermittent global synchronization across multiple cortices between onsets of CS and CR. From Freeman and Rogers (2003). (**B**) Classification of frames was based on Euclidean distances between feature vectors, **A**(*t*), from reinforced CS+ and unreinforced CS–. Deletion of the EEG data from each cortical area reduced the classification assay in the test period but had no significant effect in the control period. The strongest effect was by removal of the olfactory signals, while the least was by deletion of the entorhinal signals. The window was 128 ms stepped at 64 ms. The mean t:c-ratios (test:control) from the control period from 1.6 to 2.4 s and the test period from 3.6 to 4.4 s were derived after the deletions specified as follows. None: 0.34 vs. 2.71. EC: –0.01 vs. 2.36. VC: 0.01 vs. 2.17. SM: 0.00 vs. 2.04. AC: 0.01 vs. 1.60. OB: 0.07 vs. 0.74. From Freeman and Burke (2003)

level of global synchronization was calculated from the analytic phase, $\phi_j(t)$, $j = 1,..., 64$, which revealed episodic increases in phase locking (Fig. 5A) particularly in the interval between CS and CR (Freeman and Rogers 2003).

The frames having carrier frequencies in the gamma range occurred early after CS onset in the primary sensory areas and were modality specific (Fig. 3); they were found only in the sensory cortex to which the CS were directed. The classifiable frames with carrier frequencies in the beta range occurred >400 ms after CS onset, had longer durations, and had greater diameters (Freeman 2005). In subjects with widely dispersed electrodes the 1×64 feature vectors from AM patterns were optimally classified with respect to CS when data from all cortices were used (solid curve, Fig 5B); deletion of data from any of the four sensory areas and the entorhinal cortex diminished the classification (dotted curves). The somatosensory and motor areas in the cat and rabbit are intermingled, so the macroscopic pattern included the motor areas, thereby offering a bridge between sensory input and motor output that included much if not all of each hemisphere.

In summary, linear decomposition with the Hilbert transform of an array of ECoG serves to evaluate five basic state variables at each point in time and space in the array: the order parameter, **A**(*t*); its mean amplitude, $A(t)$; its rate of change with time, $D_e(t)$; the rate of change in phase with time (instantaneous frequency, $\omega(t)$); and rate of change with space (phase gradient, $\gamma(t)$). From these can be calculated the times of start and end of frames, durations, diameters, conduction velocities for initiation, and the times required for state transitions and convergence to attractors. These frames have been observed, measured, and classified with respect to conditioned stimuli both locally in the several sensory cortices and globally across multiple cortices. They have been found to coexist in at least two carrier and gating frequency ranges. Since the beta patterns include the motor cortices, they could provide the macroscopic frames in which motor actions are constructed. Owing to the low rates of change in space-time of the carrier frequency and the order parameter, it can be inferred that the frames reflect brief epochs of stationarity and linearity, from which might be extracted by linear regressors the time-varying structures of MSA and LFP that carry signals from cortex into the brain stem for motor control, if the frames can be found.

However, these frames are by no means obvious in raw data streams. Other techniques are required; remarkably two of these techniques are also linear. One is linear decomposition based on the theory of volume conduction (Sect. "Use of ECoG to decompose LFP by distinguishing open versus closed fields"), and the other is decomposition based in the theory of linear feedback systems (Sect. "Phase relation between state variables imposed by negative feedback"), based on the property that the frames in the beta and gamma bands appear to conform to the principle of superposition. This treatment may not hold for the alpha and theta bands, which manifest the nonlinear mechanisms that determine frame rates.

Use of ECoG to decompose LFP by distinguishing open versus closed fields

The relations between ECoG, LFP and MSA are obscured by the overlap in the volume conductor of the dendritic currents from multiple neuronal populations. The tissue in which neurons are embedded is an electrically conducting medium with specific resistance that is far less than the membrane resistance of the neurons, which explains the low amplitudes of extracellular signals from flows of dendritic current. The electric current fields established by neuron populations result from loop currents that exit neurons at "sources" and re-enter the same neurons at "sinks". The potential differences from the extracellular limbs of the current loops all sum in the volume conductor, with contributions at every point from every neuron but in varying amounts dependent on distance from the sources and sinks in accordance with Coulomb's inverse square law. The alignment and lamination of neurons in cortex is prerequisite for the summation of the potential fields from sources and sinks that are intense enough to measure as population state variables. The summation of potentials from all populations in the volume conductor is linear and instantaneous. This fact is greatly to the advantage of analyzing the contributions of multiple overlapping populations. In principle the contribution of the dendrites of every neuron to the LFP can be distinguished by linear decomposition, provided one knows the waveform of its activity and that one has sufficient recording sites within and around the field of every neuron for sufficiently long time.

However, in practice, the low-frequency dendritic potentials of single neurons can only be extracted by intracellular recording, which reflects intimately the synaptic input from the surround. The LFP is especially complex and is extremely sensitive to the depth of the recording site. Researchers in the field should adopt a convention of subscripting ECoG records by site (e.g., motor cortex, Brodmann Area 4, as in $ECoG_{BA4}$) and LFP records also by depth (for example, $LFP_{BA4,950}$ microns, preferably after post-mortem verification by layer, as in $LFP_{BA4,V}$). The ECoG is far less complex. When it is recorded at points on the surface of the cortex overlying the sites of depth recording with respect to a far distant reference site, it can be used to partition the LFP into local versus remote components. The reason is clear from an understanding of the geometry of the loop currents generated by electrochemical activity of neurons.

The source-sink pair for each neuron (equal but opposite in sign as parts of the same current loop) tends to either of two idealized forms. Pyramidal cells have axial symmetry, so that the source and sink are separated and typically of equal density; they generate a dipole field called "open", which extends instantaneously throughout the cortex to its surface and throughout the brain to the scalp. Each cell has a dipole axis. When the axes of the cells are randomly oriented as in the reticular formation, their dendritic fields sum to zero at the mesoscopic level, and their activity can only be detected from their spikes. High-amplitude open fields appear only when the individual dipoles are aligned parallel to each other and with cell bodies in a tight layer. That is why cortex forms the EEG, and basal ganglia do not. Stellate cells have radial symmetry, so that the source and sink for each cell tend to be concentric; the polarity of the field of potential is dominated by the source or sink with the highest density. The field is called "closed" because it is not detectable outside the anatomical radii of the axons and dendrites of the stellate cell type. Closed fields sum within the layer of cells of the same type, but the sums tend to have low amplitudes, owing to the cancellation of the potential fields of the concentric sources and sinks (Freeman 1975/2004).

The open fields tend to span the cortical depth with one pole in the most superficial "marginal" layer and the other pole in the deeper layers, and with the zero isopotential surface between the poles, at which the dipole field cannot be detected. This surface is observed as the "turn-over" where the field potential reverses polarity when recorded from a penetrating electrode. Closed fields tend to have maximal amplitudes close to the zero isopotential surface of the open field; they have no "turn-over" (p. 246, Fig. 4.44 in Freeman 1975/2004). LFP are mixtures of the closed and open fields, whereas ECoG are solely from open fields that are dendritic in origin in the frequency range <300 Hz and spike in origin for frequencies >300. Therefore, the separation of open field components in LFP from closed field components is facilitated by recording the ECoG with respect to a distant site from an electrode on the cortical surface at the site of entry of a microelectrode into the cortical depth. However, this decomposition does not guarantee that the LFP contains the dendritic correlate of the spike activity; in fact, the optimal recording site is at the trigger zone of a pyramidal cell with axial symmetry, where the dipole field has its maximal spatial rate of change at the zero isopotential, so there is no LFP component at all. Small shifts in position can give dramatic changes in components, hence the label of depth is highly desirable for interpretation and replication of field potential recording.

A common misconception holds that field potentials are the envelopes of spikes. The individual spike has a source-sink-source along an axon with a wavelength that is given by the product of the propagation velocity and the duration (e.g., 10 m/s × 1 ms = 1 cm). When the axon has axial symmetry, the field is open and extends instantaneously throughout the brain. However, most axons are short (<1 mm) with low velocity and with multiple branches tending to radial symmetry. The compound action potential is detectable as a diphasic or triphasic waveform only when (a) many axons in a parallel bundle are excited simultaneously and (b) then only briefly as in the far-field brain stem auditory evoked potential, before the many action potentials disperse owing to differing propagation velocities and to spatial divergence of axonal branches. These two conditions do not hold for the spike correlates of endogenous cortical field potentials. The spikes of individual cortical neurons are statistically related to LFP and ECoG, and their sum is detectable in the noise >300 Hz, but the action currents do not sum to contribute to the LFP and ECoG in the oscillations <300 Hz. As far as contributions by volume conduction are concerned, the MSA and ECoG yield independent state variables respectively from axons (inputs and outputs of cortical neurons) and dendrites (intracortical operators) by which to assess the underlying neural activity patterns.

Phase relation between state variables imposed by negative feedback

However, the optimal correlate of spike firing may be found reliably in the ECoG, and this may support extraction of a time-lagged correlate of the firing in the $LFP_{BAx,V}$. Cortical oscillatory waveforms in the beta and gamma ranges are generated by the interactions among populations of excitatory and inhibitory neurons; the most crucial evidence is that the oscillations of the inhibitory neurons lag those of the excitatory neurons on average by $\pi/2$ rad (90°), not in phase (0°) as predicted by models based on cellular properties (Traub et al. 1996; Kopell et al. 2000; Whittington et al. 2000). This is because the two most important contributions to LFP and ECoG are the dipole fields from excitatory pyramidal cells and the closed fields from inhibitory interneurons. These two laminar populations interact by negative feedback. Owing to the 90° phase lag (on average) of the output of the feedback limb from the output of the forward limb, the contributions of the dipole and closed fields are linearly separable (Freeman 1975/2004), just as the cosine function is

uncorrelated with its derivative, the sine function. Modeling the interaction with differential equations in a KII set supports the experimental proof that the excitatory and inhibitory populations oscillate at the same instantaneous frequency, and that the inhibitory oscillation on average lags the excitatory oscillation by a quarter cycle (Freeman 2000).

This relation underlies the utility of the Hilbert transform, because the quadrature of the recorded pyramidal cell dendritic output predicts the interneuronal dendritic output. When the pyramidal cell output can be approximated by a cosine, the inhibitory output can be predicted by a negative sine wave at the same frequency. The crosscorrelation between the two waveforms tends toward zero despite the functional interdependence of the two populations, so they are easily separated by PCA or ICA. The phase relation often deviates from $\pi/2$ owing to co-existing feedback by mutual excitation among pyramidal cells and mutual inhibition among interneurons. There is another reason for deviation. Phase differences between signals at differing locations commonly deviate from zero or $\pi/2$ lag in the same interactive populations owing to conduction delays from propagation velocities averaging 2.24 ± 1.18 m/s in rabbit neocortex and 2.62 ± 1.16 m/s in human neocortex, giving phase gradients of 13.2 ± 4.1 rad/mm in rabbit ECoG (Freeman 2004b) and 7.9 ± 2.4 rad/mm in human ECoG (Freeman 2006). The range of phase differences with phase dispersion, when expressed as a fraction of the cycle duration of the carrier wave, seldom exceeds ±1/8 of a cycle ($\pi/4$), which is the half-power level ($\cos^2 \pm 45° = 0.5$). These phase lags do not manifest true traveling waves, but instead they reveal the delays in cortical activation imposed by communication using propagated action potentials. The standing wave property is important, because traveling waves could not sustain the amplitude patterns illustrated for feature vectors in Figs. 3B and 4B. Within the 1/8 cycle dispersion limit the deviations from the assumption of time-invariance on which use of PCA and ICA is based appear to be negligible.

Conclusions

The major unsolved problem in BCI/BMI is to model the descending macro-meso-micro limb of the reflex-arc/action-perception-cycle, where decisions are made on what actions to take and how to take them. Neither MSA recording, nor empirical engineering, nor neuropsychological theorizing is likely to solve it; systematic acquisition and analysis of data from records of

dendritic potentials are essential. Investigation of the ascending micro-meso-macro limb indicates that the operations and transformations of neural activity are conducted in episodic, ''cinematographic'' sequences of frames. Remarkably, it turns out that, despite the obvious nonlinearities in neocortical dynamics, the neural activities within frames and their observable manifestations in electric fields are readily susceptible to linear analysis (Freeman 1975/2004; Basar 1998; Haken 2006). In this respect the temporal coherence of phase-locked waveforms is just as important as the laminar alignment of sources and sinks for the formation of LFP and ECoG. In order to extract useful information from the underlying dendritic fields of potential it is necessary to decompose the recordings from electrode arrays into the distinctive waveforms from identified populations in the mix. Within frames linear decomposition appears to be the method of choice.

The major steps in linear decomposition are provided by selective recording of open and closed fields in the volume conductor; time-lagged correlation to distinguish contributions from phase-locked oscillations of excitatory and inhibitory populations having identical frequencies with on average 90° lag between them; multiple pass bands using spatial and temporal filtering using Fourier techniques; and the identification using the Hilbert transform of the sequence of state transitions in the beta and gamma ranges of the brain activity. The value of this step inheres in the finding that the rates of change in frequency and in AM pattern are often relatively constant within the epochs bracketed by high rates of change. Not only are the frequencies of oscillation relatively constant; the maximal phase dispersion across multiple recordings of the oscillatory waveform is under a quarter cycle, the half-power soft boundary condition (Freeman 2004b), which facilitates transmission of cortical output through divergent-convergent pathways. Those pathways perform spatiotemporal integration that reduces noise in the transmission of spatially coherent signals. Additionally, use of the rate of change, $D_e(t)$, in the order parameter, $\mathbf{A}(t)$, given by Eq. (5) in Appendix 1 shows that the spatial AM patterns are stable under temporal integration. Therefore, linearity and spatiotemporal stationarity appear to hold to a good approximation within such segments, and the classic techniques for decomposition that assume linearity, stationarity, and statistical independence may be applicable to signals within these segments, most importantly FFT for spectral decomposition and PCA for spatial decomposition. ICA will be inapplicable to individual frames, if the number of digitizing steps in

the duration of a typical segment is less than the number of electrodes, as in domains of stationarity. However, ICA applied to ECoG and LFP segments including the CS/CR time interval and decomposed into the temporal frequency pass bands of interest might facilitate localization of stationary frames.

Warning should be given that typically sparse electrode arrays that under-sample the spatial textures of ECoG and LFP will not reveal clear-cut inverse relations between the two components of the analytic signal, $A(t)$ and $\phi(t)$, and the AM patterns, because multiple frames commonly overlap with differing carrier frequencies. Large high-density arrays and multiple spectral pass-bands in both temporal and spatial frequency domains will be needed to bridge from expressions of desired goals in neocortical activity to control of neuroprostheses. The real mystery is how neural masses speciate action patterns from generic states. Solution of that mystery may come through the next stage of BMI/BCI, which is the extraction from subjects of usable feature vectors, and the implementation of that information to train neural networks as universal function generators to drive adaptive control devices in desired tasks toward stated ends.

Appendix 1 The Hilbert Transform

Brain waves are commonly treated as if they were the sum of the outputs of a set of neural oscillators, each of which has a constant frequency and variable amplitude. This treatment is based on the assumption that brain dynamics is linear and time-invariant, which is clearly not the case. The advantage conveyed by this assumption is the ease with which Fourier analysis can be applied to brain waves using the Fast Fourier Transform (FFT) to decompose segments of brain waves into frequency components. The disadvantage is the inability of linear analysis to capture and display the nonlinear state transitions by which brains operate. An alternative linear transform is the Hilbert transform, which when applied to a brain wave recording in effect calculates the rate of change in the amplitude at each time step of the digitized signal. This operation effectively re-expresses an oscillation as a vector that rotates counterclockwise in the complex plane. The amplitude is expressed by the length of the vector, $\mathbf{A}(t)$, and the rate of change is expressed by the angular velocity of the rate of rotation of the vector about the origin of the complex plane. The rate of rotation is expressed as a rate of change in phase in degrees/second, radians/second (rad/s), or cycles/second (Hz). The immediate advantage is that the Hilbert transform

decomposes a brain wave into an analytic amplitude, $A(t)$, and an analytic phase, $\phi(t)$. The change in phase in rad with each time step divided by the digitizing interval in s approximates an instantaneous frequency that can vary, unlike the frequencies that are extracted by Fourier decomposition. The disadvantage is that the Hilbert transform is very sensitive to noise of many kinds; it only works well after band pass filtering of a brain wave. Criteria for optimal band pass filtering have been described elsewhere (Freeman 2004a, b; 2005; 2006).

The application of the Hilbert transform to each intracranial recording from an array of microelectrodes is a multi-step procedure. First, a high pass filter set at ~400 Hz extracts the MSA, and a low pass filter set at ~400 Hz extracts the LFP from the same n microelectrode recordings. Second, the low pass data are downsampled from ~40,000/s to 200/s and normalized to zero

mean for every channel and unit standard deviation (SD) for all channels, trials and data sets to give the normalized LFP. Third, the demeaned, normalized LFP are band pass filtered in the classic empirical ranges: theta (3–7 Hz), alpha (7–12 Hz), beta (12–30 Hz), gamma (30–60 Hz), and high gamma (60–200 Hz), and the n channels in each pass band are segmented to save the data from each trial with ~3 s preceding and ~3 s following each CS onset (Fig. 6A). Fourth, the Hilbert transform is applied to get the analytic signal, $V_j(t)$, with a real part (blue curve), the filtered LFP, $v_j(t)$, and an imaginary part (red curve), $i\,u_j(t)$, the output of the Hilbert transform:

$$V_j(t) = v_j(t) + i\,u_j(t), \quad j = 1, \ldots, 64, \qquad (1)$$

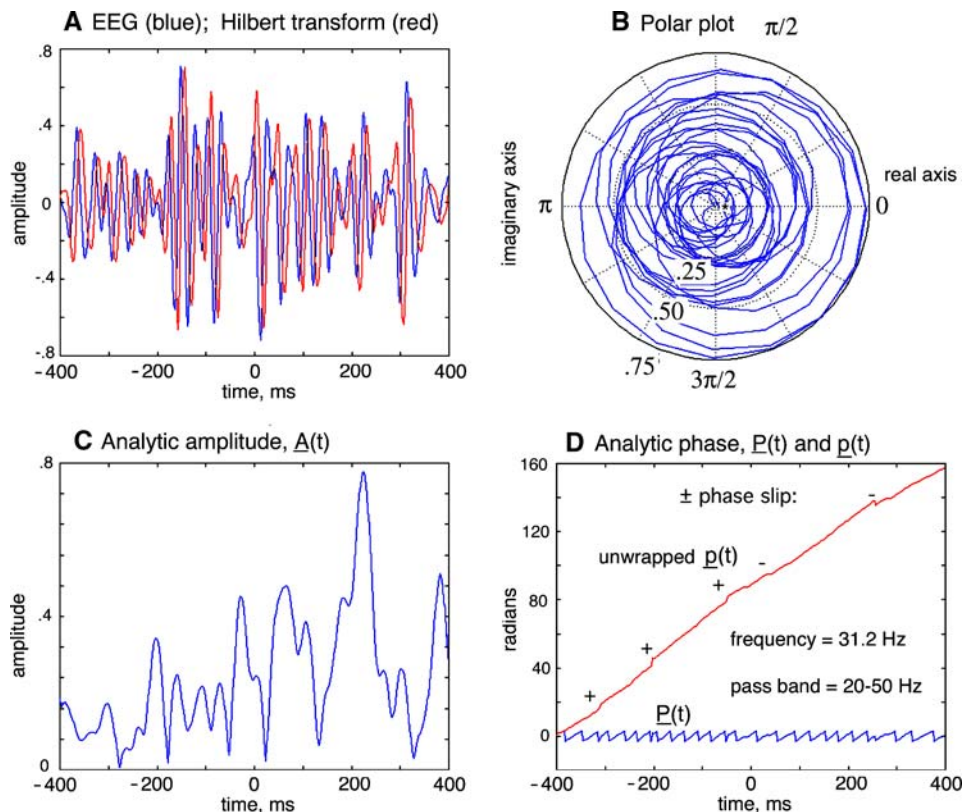where the Hilbert transform of $v_j(t)$ in the time segment, t',

Fig. 6 (**A**) A representative segment on a typical channel was selected from visual cortical beta EEG after band pass filtering (20–50 Hz). The blue curve shows the spatial ensemble average of the real part representing the excitatory neuronal output ($v(t)$ in Eq. (1)). The red curve shows the imaginary part representing the inhibitory neuronal output ($u(t)$). (**B**) The real part of the analytic signal (abscissa) is plotted against the imaginary part (ordinate) as a vector. Time is implicit in counterclockwise rotation of the vector tip starting from the asterix just to the right of the origin where the axes cross. (**C**) The blue curve shows the average analytic amplitude, $A(t)$, in Eq. (3) giving the length of the vector). (**D**) The blue sawtooth curve shows the average analytic phase, $\phi(t)$, given by Eq. (4). The red curve shows the average unwrapped phase, $p(t)$. The analytic frequency ω is taken from the slope in rad/s. The deviations from the average slope show "phase slip" which is due to repeated state transitions. A reinforced conditioned stimulus (CS+, full field dim light flash) was delivered at 0 ms. From Freeman (2004a)

$$u_j(t) = 1/\pi \mathrm{PV} \int_{-\infty}^{+\infty} v_j(t')/(t-t')\mathrm{d}t', \qquad (2)$$

where PV signifies the Cauchy Principal Value. The imaginary part is also known as the quadrature of the signal, because each cosine component in the recorded signal is transformed to a sine component; taking the derivative by the transform is equivalent to shifting the phase of $v(t)$ by 90° ($\pi/2$ rad) to get $u(t)$.

Fifth, the square root of the sum of squares of the real and imaginary parts gives the analytic amplitude, $A_j(t)$, for each channel, $j = 1,...,n$,

$$A_j(t) = (v_j^2(t) + u_j^2(t))^{.5}, \qquad (3)$$

and the arctangent of the ratio of the imaginary part divided by the real part gives the analytic phase, $\phi_j(t)$ (Fig. 6B):

$$\phi_j(t) = a\tan(u_j(t)/v_j(t)), j = 1, 64. \qquad (4)$$

The mean of the square of amplitude, $A_j^2(t)$ over n gives the mean power, $A^2(t)$ (Fig. 6D), and the set of n scalar values of $A_j(t)$ divided by $A(t)$ gives the normalized feature vector at each time step, $\mathbf{A}(t)$. The feature vector provides a measure of the order parameter of the ensemble of cortical neurons that is under observation. $\mathbf{A}(t)$ specifies the normalized spatial pattern formed in the pass band by the signals from the n channels, and it designates a point in n-space that is occupied by the tip of the feature vector as it describes a trajectory through infinite brain state space that is projected into n-space by measurement.

Sixth, the rate of change in the normalized order parameter, $D_e(t)$, is calculated from the analytic power by calculating the Euclidean distance between the tips of the feature vectors in n-space at each successive digitizing step:

$$D_e(t) = |\mathbf{A}^2(t)| - |\mathbf{A}^2(t-1)|. \qquad (5)$$

$D_e(t)$ is a measure of the stability and stationarity of the normalized spatial pattern. Successive points in time specified by $\mathbf{A}(t)$ form clusters, whereas epochs of rapid change are manifested by a wide trajectory through n-space. The ratio of the rate of energy dissipation estimated by mean analytic power, $A(t)$, divided by the rate of change in the order parameter, $D_e(t)$, gives a quantity called the "pragmatic information", $H_e(t)$, which is maximal when the LFP amplitude peaks and when concomitantly the spatial pattern of the LFP is optimally stabilized.

$$H_e(t) = A^2(t)/D_e(t). \qquad (6)$$

Seventh, the analytic phase, $\phi_j(t)$. is unwrapped by adding π radians at each break point where the arctangent goes to infinity (Fig. 6D), and the analytic frequency, $\omega_j(t)$, is estimated by calculating the phase difference between successive digitizing steps in the unwrapped analytic phase, $\Delta\phi_j(t)$, time series and dividing that difference by the duration of the digitizing step, Δt. The mean analytic frequency, $\omega(t)$, and its spatial standard deviation, SD_X, are calculated over the n channels at each time step. Typically in neocortical data the values of $\omega(t)$ and SD_X are nearly constant for time periods of 60–120 ms indicating stationarity, and they fluctuate over the n channels in brief time periods that demarcate sudden transitions in analytic frequency, power, and spatial pattern. The implication is that areas of neocortex function in near-linear, stationary dynamics most of the waking state, but undergo brief state transitions 3–10 times each second. During the transitions the analytic amplitude, $A(t)$, drops to a low level, and the variances of the analytic frequency, $\omega(t)$, given by $\mathrm{SD}_X(t)$ increases briefly but dramatically (Fig. 1) in what is known as "phase slip" (Pikovsky et al. 2001). A state transition appears to be required to initiate the formation of a new spatial pattern, $\mathbf{A}(t)$, which is the order parameter manifesting a nonconvergent "chaotic" attractor in the landscape of basins of attraction sustained by an area of cortex.

References

Barrie JM, Freeman WJ, Lenhart M (1996) Modulation by discriminative training of spatial patterns of gamma EEG amplitude and phase in neocortex of rabbits. J Neurophysiol 76:520–539

Barrie JM, Holcman D, Freeman WJ (1999) Statistical evaluation of clusters derived by nonlinear mapping of EEG spatial patterns. J Neurosci Meth 90:87–95

Basar E (1998) Brain function and oscillations. 1. Principles and approaches. Springer-Verlag, Berlin

Carmena JM, Lebedev MA, Henriquez CA, Nicolelis MAL (2005) Stable ensemble performance with single-neuron variability during reaching movements in primates. J Neurosci 25(46):10712–10716

Chapin JK, Moxon KA, Markowitz RS, Nicolelis MAL (1999) Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat Neurosci 2:664–670

Freeman WJ (1974) Relation of glomerular neuronal activity to glomerular transmission attenuation. Brain Res 65:91–107

Freeman WJ (1975) Mass action in the nervous system. Academic, New York. Electronic 2004: http://sulcus.berkeley.edu/MANSWWW/MANSWWW.html

Freeman WJ (2000) Neurodynamics: an exploration of meso-scopic brain dynamics. Springer, London

Freeman WJ (2004a) Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin Neurophysiol 115:2077–2088

Freeman WJ (2004b) Origin, structure, and role of background EEG activity. Part 2. Analytic phase. Clin Neurophysiol 115:2089–2107

Freeman WJ (2005) Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. Clin Neurophysiol 116(5):1118–1129

Freeman WJ (2006) Origin, structure, and role of background EEG activity. Part 4. Neural frame simulation. Clin Neurophysiol 117:572–589

Freeman WJ, Burke BC (2003) A neurobiological theory of meaning in perception. Part 4. Multicortical patterns of amplitude modulation in gamma EEG. Int J Bifurc Chaos 13:2857–2866

Freeman WJ, Gaál G, Jornten R (2003) A neurobiological theory of meaning in perception. Part 3. Multiple cortical areas synchronize without loss of local autonomy. Int J Bifurc Chaos 13:2845–2856

Freeman WJ, Rogers LJ (2002) Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEG. J Neurophysiol 87:937–945

Freeman WJ, Rogers LJ (2003) A neurobiological theory of meaning in perception. Part 5. Multicortical patterns of phase modulation in gamma EEG. Int J Bifurc Chaos 13:2867–2887

Freeman WJ, Viana Di Prisco G (1986) Relation of olfactory EEG to behavior: time series analysis. Behav Neurosci 100:753–763

Gonzalez SL, Grave de Peralta R, Thut G, Millán J del R, Morier P, Landis T (2006) Very high frequency oscillations (VHFO) as a predictor of movement intentions. NeuroImage, Short communication 32:170–179

Gordon E (2000) Integrative neuroscience. Harwood Academic, Sydney

Haken H (1983) Synergetics: an introduction. Springer-Verlag, Berlin

Haken H (2006) Synergetics of brain function. Int J Psychophysiol 60(2):110–124

Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442:164–172

Kopell N, Ermentrout GB, Whittington MA, Traub RD (2000) Gamma rhythms and beta rhythms have different synchronization properties. Proc Natl Acad Sci USA 97:1867–1872

Kozma R, Freeman WJ (2001) Chaotic resonance: methods and applications for robust classification of noisy and variable patterns. Int J Bifurc Chaos 10:2307–2322

Liley DTJ, Alexander DM, Wright JJ, Aldous MD (1999) Alpha rhythm emerges from large-scale networks of realistically coupled multicompartmental model cortical neurons. Network-Comput Neural Syst 10:79–92

Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. Nature 412:733–736

Pikovsky A, Rosenblum M, Kurths J (2001) Synchronization—a universal concept in non-linear sciences. Cambridge University Press, Cambridge UK

Robinson PA, Loxley PN, O'Connor SC, Rennie CJ (2001) Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. #041909, Physical Ref E6304: N4PT1:1090. US15–US26

Sammon JW (1969) A nonlinear mapping for data structure analysis. IEEE Trans Comput C-18:401–409

Sanchez JS, Carmena JM, Lebede MA, Nicolelis MAL, Harris JG, Principe JC (2004) Ascertaining the importance of neurons to develop better brain–computer interfaces. IEEE Trans Biomed Eng 51:943–953

Tass P, Kurths J, Rosenblum M, Weule J, Pikovsky A, Volkmann J, Schnitzler H, Freund H (1999) Complex phase synchronization in neurophysiological data. In: Uhl C. (ed) Analysis of neurophysiological brain functioning. Springer-Verlag, Berlin, pp 252–273

Traub RD, Whittington MA, Stanford IM, Jefferys JGR (1996) A mechanism for generation of long-range synchronous fast oscillations in the cortex. Nature 383:421–424

Tsuda I (2001) Toward an interpretation of dynamics neural activity in terms of chaotic dynamical systems. Behav Brain Sci 24:793–847

Whittington MA, Faulkner HJ, Doheny HC, Traub RD (2000) Neuronal fast oscillations as a target site for psychoactive drugs. Pharmacol Therap 86:171–190

Wright JJ, Liley DTJ (1996) Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. Behav Brain Sci 19:285–295