Title
A pseudo-equilibrium thermodynamic model of information processing in nonlinear brain dynamics

Permalink
https://escholarship.org/uc/item/6qb362tx

Journal
Neural Networks, 21

Author
Freeman, Walter J, III

Publication Date
2008

Peer reviewed
A pseudo-equilibrium thermodynamic model of information processing in nonlinear brain dynamics

Walter J Freeman

Department of Molecular & Cell Biology
University of California at Berkeley
Berkeley CA 94720-3206 USA

drwjii@berkeley.edu

Neural Networks 21: 257-265, 2008

Computational models of brain dynamics fall short of performance in speed and robustness of pattern recognition in detecting minute but highly significant pattern fragments. A novel model employs the properties of thermodynamic systems operating far from equilibrium, which is analyzed by linearization near adaptive operating points using root locus techniques. Such systems construct order by dissipating energy. Reinforcement learning of conditioned stimuli creates a landscape of attractors and their basins in each sensory cortex by forming nerve cell assemblies in cortical connectivity. Retrieval of a selected category of stored knowledge is by a phase transition that is induced by a conditioned stimulus, and that leads to pattern self-organization. Near self-regulated criticality the cortical background activity displays aperiodic null spikes at which analytic amplitude nears zero, and which constitute a form of Rayleigh noise. Phase transitions in recognition and recall are initiated at null spikes in the presence of an input signal, owing to the high signal to noise ratio that facilitates capture of cortex by an attractor, even by very weak activity that is typically evoked by a conditioned stimulus.

Key words: action-perception cycle; dissipative structures; ECoG electrocorticogram; non-equilibrium thermodynamics; phase transition; reinforcement learning; self-organized criticality

1. Introduction

Cognitive neurodynamics describes the process by which brains direct the body into the world and learn by assimilation from the sensory consequences of the brain-directed actions. Repetition of the process constitutes the action-perception cycle by which knowledge is accumulated in small increments. Each new step yields a freshly constructed frame that is updated by input to each of the sensory cortices (Freeman, 2004a,b). The continually expanding knowledge base is expressed in attractor landscapes in each of the cortices. The global memory store is based on a rich hierarchy of landscapes of increasingly abstract generalizations (Freeman, 2005, 2006b). The first step of the acquisition of new knowledge is the selection by a stimulus of an attractor among landscape of attractors for the primary categories of sensory stimuli in each modality, for example, the repertoire of odorant substances that an animal can seek, identify, and respond to at any one stage of its lifelong experience. Each attractor is based in a nerve cell assembly of cortical neurons that have been pair-wise co-activated in prior Hebbian association and sculpted by habituation and normalization (Kozma and Freeman, 2001). Its basin of attraction is determined by the total subset of receptors that has been accessed during learning. Convergence in the basin to the attractor gives the process of abstraction and generalization to the category of the stimulus. This categorization
process holds in all sensory modalities (Freeman and van Dijk, 1987; Barrie, Freeman and Lenhart, 1996; Ohl, Scheich and Freeman, 2001; Freeman and Burke, 2003; Freeman and Rogers, 2003; Freeman, 2006a). The convergence to and holding of a cortical state by an attractor provides a frame that typically includes each entire primary sensory cortex and lasts about a tenth of a second. The action-perception cycle includes 3-5 frames plus transition times between frames that repeat at rates in the theta range (3-7 Hz).

A major aim of cognitive neurodynamics is to model the cycle. On the one hand information-based models of the action-perception cycle succeed in describing the sensory input to cortex, the cortical evoked potentials, and the motor output in response to the input. They fail in modeling the mobilization of knowledge, because the portions of the knowledge base that are recalled from the attractor landscapes cannot be measured or expressed in bits of information (Lucky, 1989), yet those portions are the major determinants of cortical responses to stimulus. A further difficulty is that convergence in numerical models is by tree search and gradient descent, which are computationally intensive. von Neumann (1958) noted: “…the mathematical or logical language truly used by the central nervous system is characterized by less logical and arithmetical depth than what we are normally used to. … We require exquisite numerical precision over many logical steps to achieve what brains accomplish in very few short steps (pp. 80-81)”.

On the other hand there is no question that brains are open thermodynamic systems operating far from equilibrium. Brains burn glucose to store energy in glycogen (“animal starch”) and high-energy adenosinetriphosphate (ATP), and in transmembrane ionic gradients; they dissipate free energy in proportion to the square of the dendritic ionic current densities that are manifested in epiphenomenal electric fields, and that mediate the action-perception cycle (Freeman and Vitiello, 2006) by controlling the pulse frequencies of axons. Brain imaging techniques such as fMRI are indirect measures of metabolic dissipation of free energy, relying on secondary increases in blood flow and oxygen depletion. The dendrites dissipate 95% of the metabolic energy in summed excitatory and inhibitory ionic currents; axons dissipate 5% of the energy in action potentials that carry the summed output of dendrites by analog pulse frequency modulation. The aim of this report is to sketch how records of dendritic potentials recorded from the cortical surfaces of human and animal brains (electrocorticogram, ECoG) can be used to describe how brains that produce mind and behavior act as thermodynamic engines held at pseudo-equilibrium in self-organized criticality. Thereby von Neumann’s “very few short steps” that enable action-perception cycles turn out to be sequential phase transitions (Kozma and Freeman, 2002; Kozma et al., 2005).

2. Spatial and spectral properties of dendritic potentials in ECoG

Examples of cognitive brain dynamics are drawn from the olfactory system, which is the simplest system from the standpoint of sensory pre-processing yet prototypical in being phylogenetically the precursor of the other sensory systems (Freeman, 2001). In sleep and at rest the work of dendrites is revealed by robust background activity with Gaussian amplitude distributions and $1/f^2$ ‘brown noise’
power spectra. Amplitude increases with arousal (Fig. 1, A), and spectral peaks emerge on engagement with the environment. Oscillations occur in the gamma range (30-80 Hz) in frames triggered by inhalations at frame rates in the theta range (3-7 Hz). In the presence of input carried by a learned odorant substance, the gamma waves synchronize over the olfactory bulb and mobilize an attractor in a landscape of attractors (Fig. 1, B) governing the expression of spatial patterns of amplitude modulation (AM) and phase modulation (PM) of a gamma carrier wave (Fig. 2, A). Each pattern is recorded with 8x8 electrode arrays and plotted as a point in 64-space. Projection into 2-space of points derived from multiple trials with different substances reveals clusters (Fig. 2, B) that reflect the underlying attractor landscape.

Cortex is bistable, having a receiving phase during which the landscape is latent, and a transmitting phase during which the landscape is brought on line during the sensory receptor input barrage brought by inhalation. Selection by sensory input of one of the basins of attraction precipitates spontaneous symmetry breaking (Freeman and Vitiello, 2006) in the form of a phase transition from the receiving phase to the transmitting phase. During exhalation another phase transition returns the bulb to the receiving phase. These properties are schematized by adapting the phase diagram for water (Fig. 8, A), which is the static, time-invariant relation between energy and entropy at equilibrium, to the relation (B) between the rate of increase in order (negentropy) and power (rate of energy dissipation) far from equilibrium. The order parameter is indexed (Fig. 8, B) by the ratio, \( \frac{H_e(t)}{\text{Euclidean distance}} \) of mean analytic power to the Euclidean distance between feature vectors at successive digitizing steps in 64-space (Fig. 2, B); a small step within a cluster indicates a stable pattern and therefore a high degree of order. A large step is required in jumping from one cluster to another cluster, as in moving from the control phase to the stimulated phase. Power is estimated (Fig. 8, B) from mean square analytic amplitude, \( A^2(t) \), derived with the Hilbert transform applied to the ECoG to calculate the analytic signal (Freeman, 2004a,b; 2005; 2006a).

3. **Piece-wise linear analysis of dendritic potentials by root locus**
The dendrites are normally kept within a narrow near-linear range by the time-invariant nonlinearity of the axonal trigger zones, so the dynamics of local populations is approximated with a 2nd or 3rd order ordinary differential equation (Freeman, 1975/2004). The time-invariant sigmoid function (Fig. 3) is evaluated by fitting an equation derived from the Hodgkin-Huxley equation for the neuron to the normalized conditional pulse probability on the ECoG amplitude and frequency (A). Piece-wise linearization is by taking the derivative of the curve to calculate the slope of the tangent at the operating point (B), which at rest is at unity gain, giving the condition required for the steady state.

**Fig. 2.** A. Contour plots show ECoG amplitude from a 8x8 (6x6 mm) surface array. B. Projection of points from 64-space to 2-space is by step-wise discriminant analysis; the origin of the display space is translated and the coordinate axes are rotated to minimize within-group variance and maximize between-group variance. From Freeman and Viana Di Prisco (1986)

Perturbation by impulse input (electric stimulus to an afferent pathway, Fig. 4, A) gives a brief increase in pulse density in proportion to intensity and an exponential decay with rate also proportional to input intensity. Extrapolation to zero amplitude at threshold gives zero decay rate (a step response, Fig. 4, B), which implies that excitatory populations are self-stabilized by homeostatic regulation at unity gain. Stability is expressed by a pole \( \Delta \) at the origin of the complex plane and by a point attractor in the phase space of brain state space.

The impulse response of the interactive excitatory-inhibitory population oscillates about an exponentially decaying baseline shift (Fig. 5, A) that is imposed by periglomerular excitation (Fig. 4, A), and with a fixed frequency and an exponentially decay envelope of which the rate likewise increases with input intensity. Piece-wise linear analysis using the root locus technique (Fig. 5, B) extrapolates to a complex conjugate pole pair on the imaginary axis, which is governed by the pole at the origin within the limits indicated by the vertical and diagonal limits partitioning the complex plane. Excitatory impulse input that bypasses the excitatory bias gives a vertical root locus (Fig. 6, A) indicating the importance of the basal excitatory bias for establishing oscillations that have the same carrier frequency in the distributed population with normal input.
Fig. 3. A. Normalized pulse probability conditional on ECoG amplitude in the gamma range with the numerical derivative. B. Unity gain (steady state) is offset to a non-zero value dependent on arousal (Fig. 1, A) for excitatory populations and at zero wave density for excitatory-inhibitory populations generating oscillations From Freeman (1979).

Fig. 4. A. Impulse responses of the periglomerular mutually excitatory population. B. Root locus plot of the rate constants calculated by piece-wise linearization of the equations for dynamics of mutual excitation at an observed operating point. From Freeman (pp. 289-292, 1975/2004)

The amplitude of the bias is time-varying, as revealed by repeated sampling of the impulse response at fixed near-threshold intensity. The spontaneous variation gives a very different root locus (Mode 2), because it crosses the imaginary axis into the right half of the complex plane with increasing response amplitude (Fig. 6, B) and converges back to the imaginary axis in the gamma range, indicating the existence of a limit cycle attractor in the gamma range and thus the stable transmitting state that must be accessed by phase transition.

Interactions of excitatory and inhibitory neurons in cortex give oscillations in the beta (12-30 Hz) and gamma (30-80 Hz) ranges through a wide range of feedback delays, giving broad spectral bands that can be modeled with multiple negative feedback loops (Freeman, 1975/2004). Superposition holds in small-signal ranges, so that the pass bands can be simulated with transfer functions for linear filters. When brown noise is filtered in the beta or gamma range, the analytic signals from the multichannel ECoG both at rest and at work give fluctuating peaks of mean amplitude, $\Delta(t)$, separated by sharp
null spikes at which $\Delta(t)$ approaches zero (Figure 7, A). At null spikes the phase, $\phi(t)$, is undefined, so the spatial standard deviation as a time function, $SD_X(t)$, across channels increases in a sharp spike, contrasting with fixed $SD_X(t)$ and mean $\phi(t)$ in peaks of mean $\Delta(t)$.

Fig. 5, A. Impulse responses of excitatory-inhibitory population with excitatory bias. B. Root locus plot showing the dependence of decay rate on stimulus intensity at a constant frequency of oscillation. From Freeman (pp. 361-364, 1975/2004).

This operation reproduces the inverse relation of $\Delta(t)$ and $SD_X(t)$ (Figure 7, B). The null spikes are beats in a type of Rayleigh noise (Freeman, 1975/2004). They are defects that resemble vortices and show that the lower is the coherent mesoscopic background analytic amplitude, the higher is the amplification of the response to impulse input. The root loci suggest that the null spike exposes a singularity in cortical dynamics, at which the cortical mechanism approaches criticality. At criticality all wavelengths coexist, and the mesoscopic system can be captured by microscopic input that is amplified by a Hebbian assembly. Being dependent on the background activity that is self-stabilized, the critical state is self-organized at each level of arousal (Fig. 8, B).
Fig. 7. A. The mean analytic amplitude, $A(t)$, and the spatial SD $X(t)$ are inversely correlated. The inhibitory feedback acts as a band pass filter on the ECoG output, which results in episodic decreases in coherent power by interference without reduction in total power, resembling beats in Rayleigh noise. The wide range of feedback frequencies is supported by the power-law distribution of feedback distances and delays. B. Simulation is by band pass filtered $1/f^2$ brown noise that can be generated by summing multiple time series of random numbers (Schroeder, 1991). From Freeman (2006a).

4. Thermodynamic phase diagram from ECoG analysis

The first step is to display the static, time-invariant phase diagram for water at equilibrium in the coordinates of entropy (pressure) vs. energy (temperature) to identify the two phase boundaries, the triple point, and the critical point (Fig. 8, A). The diagram is adapted to the open, dissipative state by adopting two new coordinates: the rate of increase in order measuring negentropy and the rate of energy dissipation (power) (B). Order is indexed by the pragmatic information, $H(t)$ (Freeman, 2006a), and power by mean square analytic amplitude, $A^2(t)$. At equilibrium the critical point is at the origin (0,0), which holds under deep anesthesia with total suppression of ECoG and action potentials (Fig. 1, B).

Fig. 8 A. Time-invariant thermodynamic phase diagram. From Blauch (2007). B. Adaptation to pseudo-equilibrium by introducing two dynamic state variables and fixing the critical point at the origin. From Freeman (2007b).

With increasing arousal, the order, the power, and the critical point increase together. The critical point that governs cortical dynamics is collocated with the non-zero point attractor at the origin of the complex plane (Fig. 3, B), and the phase boundary between receiving and transmitting phases is collocated with the imaginary axis. Convergent linear dynamics is described with root locus techniques to the left; divergent dynamics is predicted by root loci extending into the right half of the complex plane.

The dynamic operating point of the receiving state ($\Delta$, ‘rest’ in Fig. 9, B) is maintained by extracortical input at a steady state with greater dissipation and lower order than the critical point ($\Delta$, SOC, Fig. 9, B). This operating point (fig. 9, A) shows the mean value
Thermodynamic model of brain dynamics

about which the fluctuations occur in amplitude of the impulse responses. In the complex plane its location corresponds to a pair of complex conjugate poles showing the frequency and decay rate of the cortical impulse response in the presence of inhibitory feedback (Fig. 5, A). During the null spikes of the Rayleigh noise the sensitivity of the cortex increases as the background noise goes to zero, as shown by the increase in amplitude of the impulse response. The decay rate of the impulse response decreases toward zero (Freeman, 1975/2004) as the cortex approaches the critical point. At that point of criticality, the linear analysis predicts that the signal-to-noise ratio in the cortex approaches infinity. Any input containing action potentials that target a nerve cell assembly can activate it and access the attractor that implements recognition and recall. When the Rayleigh noise returns, it bears the AM pattern of that attractor and maintains it for 3 to 5 cycles of the carrier frequency.

The key piece of evidence is that the operating point is shifted in the direction of decreased dissipation (compare Fig. 6, B and Fig. 9, B). The phase transition to from reception to transmission with a discontinuity in analytic phase is followed by an early increase in order, then by a massive increase in both order and dissipation (Freeman, 2005, 2006a). The second phase transition without a discontinuity in analytic phase returns the cortex to its receiving ‘rest’ state. The loop repeats 3 to 5 times in an action-perception cycle, between a conditioned stimulus and the conditioned response in a helix in time that projects into the ellipse seen in Fig. 9.

Fig. 9 A. The critical point is translated by arousal toward increased dissipation and order (right upward). The operating point is displaced toward increased dissipation and decreased order by centrifugal input from other parts of the brain (Mode 1i). The conjoint increase in order and dissipation is replicated by orthodromic afferent electrical stimulation (Mode 1e). Spontaneous variation leading to symmetry breaking (Mode 2) is attributed to the amplitude fluctuations shown in Fig. 7. The inset square shows the part of the complex plane that is used for linear analysis (Freeman, 1975/2004). B. Early in a phase transition the power and order both decrease (Freeman, 2005). Next order increases; thereafter power increases; after peaking both decrease into the receiving phase, so that the cycle rotates clockwise. The ellipse is a projection into the plane of a helix in time, with 3 to 5 repetitions in an action-perception cycle. From Freeman (2007b).

5. Discussion

This neural mechanism depends on robust maintenance by cortex approximating scale-free dynamics at criticality (Linkenkaer-Hansen et al, 2001; Freeman, 2007a ), where all frequencies and wavelengths of activity may simultaneously be expressed, as revealed in the power-law distributions of spectra and other parameters and variables. Its operation can explain the fact that the first step in a phase transition is reduction in amplitude, not in the
surge of dissipation expected following sensory impact. It explains how a faint whiff, whisper, frisson or glimpse can capture an entire prepared sensory cortex in a literal eye blink without necessity for logical tree search or gradient descent. It accounts for the fact that the AM pattern thus retrieved is not a representation of a stimulus but the mobilization of the knowledge that a subject has accumulated about the stimulus through reinforcement and Hebbian learning. It accounts for the attentional blink (Fragopanagos, Kockelkoren and Taylor, 2005). It explains the cross-spectral correlation of beta-gamma with the peaks of activity in the theta range, and for the fact that psychometric studies of frame repetitions do not closely correlate with theta or alpha, because percept formation requires coincidence of null spikes and limbic sampling through its control of snacks, whisk, and saccades.

The key finding on which this study is based is the demonstration that the neural correlates of conditioned stimuli (CS) in all sensory cortices, not just olfaction, have the form of spatial patterns of amplitude modulation (AM) of chaotic carrier waves with frequencies in the beta and gamma ranges. The 64 electrodes with which the AM patterns are measured are close enough together to be well within the correlation diameters of the patterns, yet not so close as to significantly over-sample the textures. The optimal form for classification of AM patterns with respect to CS is the analytic amplitude from the Hilbert transform (Freeman, 2005). The times of onset and offset of the individual frames are best demarcated by the analytic phase differences divided by the duration of the digitizing step, giving the analytic frequency. Owing to the shared wave form that includes the conic spatial pattern of phase modulation (Freeman, 2004a), each frame can be described by a 64x1 column feature vector that specifies a point in 64-space for classification, and that serves also as the vectorial order parameter. The rate of change of AM pattern in normalized 64-space gives a scalar index of the rate of change in the order parameter, corresponding to the rate of decrease in entropy in a system at equilibrium. The square of the analytic amplitude corresponds to an index of the rate of energy dissipation by the dendrites driving ionic currents across the dendritic and axonal membranes, which is echoed in the minute passive potential differences as the current flows across the extracellular resistance, giving the ECoG. The electric field is a mean-field quantity, owing to sharing of the extracellular current pathway by all neurons in the neighborhood of each electrode.

The simulation of the AM patterns and their classification (Freeman, 2006a) has shown that each AM pattern emerges by synchronization of the pre-existing background activity. The main support for this conclusion comes from recognition that the AM patterns are not representations of CS, because they change when the significance and context of fixed CS are modified. They are expressions of knowledge by the subject about the class to which the CS belongs. That knowledge is embedded in the synaptic matrix of the sensory and limbic cortices. Each class is based in a microscopic Hebbian nerve cell assembly that has been formed by incremental learning during the training process. The role of the Hebbian assembly is to support induction, abstraction and generalization, which it does by responding as a whole to excitation of any small number of inputs in any combination. Then the assembly selects the AM pattern through the pre-existing constraints embodied in cortical synapses.
Therefore, as widely believed, the volley of sensory input does retrieve a memory embedded at least in part in sensory cortex. What is new in the thermodynamic model is the explanation of how such a Hebbian assembly can bring an entire sensory area into the stable, ordered fluctuation of an AM pattern. The proposed process is that each phase transition taking cortex from a receiving state to transmitting state begins with the onset of a null spike, in which the mesoscopic analytic power approaches zero. The analytic phase and frequency are then undefined, but the microscopic activity continues unabated, as reflected in the high spatial standard deviation, \( SD_X(t) \). Well before the amplitude returns to a high value, oscillations re-synchronize to a new value of carrier frequency across the null spike, as shown by the increase in the index of the order parameter indicating stabilization of the new AM pattern.

The origin of the pre-existing background activity has been traced to the interaction among excitatory neurons, which is stabilized by refractory periods and not by thresholds or inhibition. The role of inhibitory neurons is to impose negative feedback, facilitated by their interactions through gap junctions (Steyn-Ross et al., 2007) and by the tuning of membrane permeabilities (Traub et al., 1996; Whittington et al., 2000), but depending on negative feedback as shown by the near quarter-cycle lag of the oscillations of the inhibitory neurons behind those of the excitatory neurons at the same frequency (Freeman, 1975/2004; Ahrens and Freeman, 2001). The gamma frequencies arise by short, local delays; the beta frequencies emerge by longer, more distant delays. Negative feedback acts as a band enhancement filter and gives rise to the Rayleigh noise with null spikes at rates proportional to the center and width of the pass band (Freeman, 2006a). The value of piece-wise linear analysis using Hamiltonians becomes apparent in demonstrating the stability boundary of linear approximations along the imaginary axis of the complex plane, which translates into the phase boundary across which the phase transition occurs with a discontinuity in the analytic phase.

The linear analysis yields two critical exponents: the zero eigenvalue representing the point attractor that governs the background activity, and the complex eigenvalue near 40 Hz that governs the gamma fluctuations. The existence of both the implied point attractor and limit cycle attractor were predicted by extrapolation to zero input and have been confirmed by surgical isolation of parts of the olfactory system (Freeman, 1975/2004; Gray and Skinner, 1988). Both attractors are engulfed by chaotic attractors in the intact system (Freeman, 1987). The key event in the phase transition is the abatement in the null spike of mesoscopic (classic) power that reveals the microscopic (quantal) power. The cortex approaches criticality at which microscopic input that is provided by the Hebbian cell assembly that captures the system and re-expresses its microscopic pattern in a mesoscopic pattern over the whole of the receiving cortex (Freeman and Vitiello, 2007).

6. Conclusion and Summary

Perhaps the most significant departure here from conventional views of brain function is the shift from information-based dynamics to knowledge-based dynamics, in which microscopic information processing extends from sensory receptors to the selection of a Hebbian cell assembly, after which mesoscopic knowledge
processing takes over. The key role is played by the AM pattern, which is the behavioral connection with the CS, the feature vector of the knowledge fragment, and the order parameter by which the phase transition is verified. The analytic phase shows that minor state transitions occur as frequently as bubbles in boiling water, but a phase transition can only be assured when it enables the construction of a behaviorally related AM pattern. The nature of the phase boundary is revealed by piece-wise linear exploration of the dynamics using the impulse responses and root locus techniques, with describing functions to express the relations of the parameters of the dynamics to the physiological and pharmacological properties of the neurons, their synaptic connections, and the subjects’ behaviors. Further interpretation of the root loci in the receiving and transmitting states and the critical exponents of the transition through the singularity between them will require modeling with with equations relying on the tools of classical physics, neuropercolation (Kozma et al., 2005), dissipative quantum field theory (Freeman and Vitiello, 2007), and renormalization group theory (Freeman and Cao, 2007).

References


