%A L. Ingber
%T Statistical mechanics of neocortical interactions: Multiple scales of EEG
%J Electroencephal. clin. Neurophysiol.
%N
%V
%D 1994
%P (to be published)
Expanded presentation of invited talk to Frontier Science in EEG Symposium, New Orleans, 9 Oct 1993

Statistical Mechanics of Neocortical Interactions: Multiple Scales of EEG

Lester Ingber

Lester Ingber Research P.O. Box 857 McLean, VA 22101 (U.S.A.)

ingber@alumni.caltech.edu

Summary: The statistical mechanics of neocortical interactions (SMNI) approach derives a theoretical model for aggregated neuronal activity that defines the "dipole" assumed by many EEG researchers. This defines a nonlinear stochastic filter to extract EEG signals.

Key Words: EEG; Statistical Mechanics; Nonlinear

1. Introduction

A plausible model of the statistical mechanics of neocortical interactions (SMNI), spans several important neuroscientific phenomena including phenomena measured by electroencephalography (EEG) (Ingber, 1982; Ingber, 1983; Ingber, 1984; Ingber, 1985a; Ingber, 1985b; Ingber, 1991; Ingber, 1992; Ingber, 1994; Ingber, 1995a; Ingber and Nunez, 1990). Fitted SMNI functional forms to EEG data may help to explicate some underlying biophysical mechanisms responsible for the normal and abnormal behavioral states being investigated.

Section 2 gives a "top down" description of the present paradigm generally accepted by people in the EEG world. This description of EEG phenomena at the scale of centimeters is the focal point of the SMNI development. Section 3 gives an outline of the SMNI "bottom up" development across multiple scales. At each of the scales developed by SMNI, it is reasonable to look for experimental phenomena to at least check the systematics of this development. Section 4 describes the description of short-term memory capacity by SMNI columnar dynamics. Section 5 describes the systematics of EEG by early SMNI theory at larger scales. Section 6 describes some recent work fully developing nonlinear probability distributions to fit EEG data. Section 7 describes some mathematical and numerical aspects of the SMNI development that are not only useful for SMNI but also have quite generic utility in other disciplines. Section 8 is the conclusion, highlighting the utility of SMNI for EEG studies.

2. SMNI Rationale—"Top Down"

In order to detail a model of EEG phenomena, it is useful to seek guidance from "top-down" models; e.g., the nonlinear string model representing nonlinear dipoles of neuronal columnar activity, a paradigm currently accepted by the EEG community (Ingber and Nunez, 1990; Nunez, 1981).

2.1. Noninvasive Recordings of Brain Activity

There are several noninvasive experimental or clinical methods of recording brain activity, e.g., electroencephalography (EEG), magnetoencephalography (MEG), magnetic resonance imaging (MRI), positron-emission tomography (PET), and single-photon-emission-computed tomography (SPECT). While MRI, PET, and SPECT offer better three-dimensional presentations of brain activity, EEG and MEG offer superior temporal resolutions on the order of neuronal relaxation times, i.e., milliseconds.

Recently, it also has been shown under special experimental conditions that EEG and MEG offer comparable spatial resolutions on the order of several millimeters; a square millimeter is the approximate resolution of a macrocolumn representing the activity of approximately 10^5 neurons (Cohen *et al*, 1990). This is not quite the same as possessing the ability to discriminate among alternative choices of sets of dipoles giving rise to similar electric fields.

2.2. EEG Electrodes

A typical map of EEG electrode sites is given in Fig. 1. Many neuroscientists are becoming aware that higher electrode densities are required for many studies. For example, if each site in this figure represented 5 closely spaced electrodes, a numerical Laplacian can be calculated to represent this site. Such Laplacian techniques may offer relatively reference-free recordings and better estimates of localized sources of activity.

Figure 1

2.3. EEG Power Spectra

Limiting cases of linear (macroscopic) theories of intracortical interaction predict local wave phenomena and obtain dispersion relations with typical wave numbers k = 10 to 100 cm^{-1} and dominant frequencies in the general range of human spontaneous EEG (1-20 Hz). However, human scalp potentials are spatially filtered by both distance and tissue between cortical current sources and surface electrodes so that scalp EEG power is attenuated to about 1%.

Fig. 2 gives a high resolution estimate of spectral density function $|\Phi(\mathrm{k}, \omega)|^2$ for EEG recorded from an awake human subject (eyes closed) using 16 scalp recording sites over the right hemisphere: (a) mu rhythm at 8.0 Hz; (b) alpha rhythm at 9.5 Hz from the same 1 minute record. The alpha mode (9.5 Hz) is consistent with standing waves, whereas the mu mode (8.0 Hz) is more consistent with posterior to anterior traveling waves (k_y) across the electrode array.

Figure 2

2.4. Single Electrode Recording of Dipole Activity

As illustrated in Fig. 3, macrocolumns may be considered as “point sources” of dipole-like interactions, mainly due to coherent current flow of top-layer afferent interactions to bottom-layer efferent interactions (Nunez, 1990). However, there is a problem of non-uniqueness of the electric potential that arises from such source activity. Laplacian measurements can help to address this problem.

Figure 3

2.5. EEG of Mechanical String

A model of data processed over many electrodes is the mechanical string. The mechanical string has linear properties and is connected to local nonlinear oscillators. Local cortical dynamics in dipole layers is here considered analogous to the nonlinear mechanical oscillators which influence global modes. Macroscopic scalp potentials are analogous to the lower modes of string displacement.

For purposes of illustration, e.g., as in Fig. 4, a linear string with attached oscillators, e.g., nonlinear springs may be compared to a one-dimensional strip of neocortex.

Figure 4

2.6. String Equation

The equation describing the string displacement Φ is

$$\frac{\partial^2\Phi}{\partial t^2} - c^2 \frac{\partial^2\Phi}{\partial x^2} + [\omega_0^2 + f(\Phi)]\Phi = 0 \ , \tag{1}$$

for a linear array (length L) of sensors (electrodes) of size s. Wave-numbers in the approximate range

$$\frac{\pi}{L} \le k \le \frac{\pi}{s} \tag{2}$$

can be observed. If the center to center spacing of sensors is also s, $L = Ms$, where M = (number of sensors - 1), $k = 2n\pi/R$ for $n = \{1, 2, 3, \cdots\}$ (string forms closed loop), and sensors span half the string (brain), $L = R/2$, then

$$1 \le n \le M \tag{3}$$

for some maximum M, which is on the order of 3 to 7 in EEG studies using 16 to 64 electrodes in two-dimensional arrays on the cortical surface.

For scalp recordings, the wavenumber restriction is more severe. For example, a typical circumference of the neocortex following a coordinate in and out of fissures and sulci is $R = 100$ cm (about 50 cm along the scalp surface). If EEG power is mostly restricted to $k < 0.5\ \text{cm}^{-1}$, only modes $n < 4$ are observed, independent of the number of electrodes.

Theory should be able to be similarly "filtered," e.g., in order to properly fit scalp EEG data.

2.7. String Observables

The string displacement (potential within the cortex) is given by

$$\Phi(x, t) = \sum_{n=1}^{\infty} G_n(t) \sin k_n x \ , \tag{4}$$

but the observed Φ is given by

$$\Phi^\dagger(x, t) = \sum_{n=1}^{M} G_n(t) \sin k_n x \ . \tag{5}$$

As can be seen here, it has been noted that spatial filtering may also induce temporal filtering of EEG (Nunez and Srinivasan, 1993).

In the linear case, where $f(\Phi) = 0$ (equal linear oscillators to simulate local circuit effects in cortical columns), then

$$\frac{\partial^2\Phi}{\partial t^2} - c^2 \frac{\partial^2\Phi}{\partial x^2} + \omega_0^2\Phi = 0 \ ,$$

$$\Phi = \sum_{n=1}^{\infty} A_n \cos \omega_n t \sin k_n x \ ,$$

$$\omega_n^2 = \omega_0^2 + c^2 k_n^2 \ , \tag{6}$$

giving a dispersion relation $\omega_n(k_n)$. For the nonlinear case, $f(\Phi) \ne 0$, the restoring force of each spring is amplitude-dependent. In fact, local oscillators may undergo chaotic motion.

What can be said about

$$\Phi^\dagger(x, t) = \sum_{n=1}^{M} G_n(t) \sin k_n x \ , \tag{7}$$

the macroscopic observable displacement potential on the scalp or cortical surface?

It would seem that $\Phi^\dagger$ should be described as a linear or quasi-linear variable, but influenced by the local nonlinear behavior which crosses the hierarchical level from mesoscopic (columnar dipoles) to macroscopic.

How can this intuition be mathematically articulated, for the purposes of consistent description as well as to lay the foundation for detailed numerical calculations?

3. SMNI Development—"Bottom Up"

In order to construct a more detailed "bottom-up" model that can give reasonable algebraic functions with physical parameters to be fit by data, a wealth of empirical data and modern techniques of mathematical physics across multiple scales of neocortical activity are developed up to the scale described by the top-down model. At each of these scales, reasonable procedures and submodels for climbing from scale to scale are derived. Each of these submodels were tested against some experimental data to see if the theory was on the right track.

For example, at the mesoscopic scale consistency of SMNI was checked with known aspects of visual and auditory short-term memory (STM), e.g., the 4±2 and 7±2 STM capacity rules, respectively, the detailed duration and stability of such states, and the primacy versus recency rule of error rates of learned items in STM (Ingber, 1984; Ingber, 1985b; Ingber, 1994).

At the macroscopic scale SMNI consistency was checked with most stable frequencies being in the high alpha to low beta range, and the velocities of propagation of information across minicolumns being consistent with other experimental data (Ingber, 1983; Ingber, 1985a).

More recently, SMNI has demonstrated that the currently accepted dipole EEG model can be derived as the Euler-Lagrange equations of an electric-potential Lagrangian, describing the trajectories of most likely states, making it possible to return to the "top-down" EEG model, but now with a derivation and detailed structure given to the dipole model (Ingber, 1991; Ingber and Nunez, 1990). The SMNI approach of fitting scaled nonlinear stochastic columnar activity directly to EEG data goes beyond the dipole model, making it possible to extract more signal from noise.

3.1. Scales Illustrated

Illustrated in Fig. 5 are three biophysical scales of neocortical interactions: (a)-(a^*)-(a') microscopic neurons; (b)-(b') mesocolumnar domains; (c)-(c') macroscopic regions. In (a^*) synaptic interneuronal interactions, averaged over by mesocolumns, are phenomenologically described by the mean and variance of a distribution Ψ (Ingber, 1982; Korn *et al*, 1981; Perkel and Feldman, 1979). Similarly, in (a) intraneuronal transmissions are phenomenologically described by the mean and variance of Γ. Mesocolumnar averaged excitatory (E) and inhibitory (I) neuronal firings are represented in (a'). In (b) the vertical organization of minicolumns is sketched together with their horizontal stratification, yielding a physiological entity, the mesocolumn. In (b') the overlap of interacting mesocolumns is sketched. In (c) macroscopic regions at the scale recognized as dipole sources of neocortex are illustrated as arising from many mesocolumnar domains. These are the regions designated for study here. (c') sketches how regions may be coupled by long-ranged interactions.

Figure 5

3.2. SMNI vs Artificial Neural Networks

A quite different approach to neuronal systems is taken by artificial neural networks (ANN) (Hertz *et al*, 1991). Both ANN and SMNI structures are represented in terms of units with algebraic properties greatly simplifying specific realistic neuronal components. Of course, there is a clear logical difference between considering a small ensemble of simple ANN units (each unit representing an "average" neuron) to study the properties of small ensembles of neurons, versus considering distributions of interactions between model neurons to develop a large ensemble of units (each unit representing a column of neurons) developed by SMNI to study properties of large ensembles of columns.

Unlike SMNI, ANN models may yield insights into specific mechanisms of learning, memory, retrieval, information processing among small ensembles of model neurons, etc. However, consider that there are several million neurons located under a cm^2 area of neocortical surface. Current estimates are that 1 to several percent of coherent neuronal firings account for the amplitudes of electric potential measured on the scalp. This translates into measuring firings of hundreds of thousands of neurons as contributing to activity measured under a typical electrode. Even when EEG recordings are made directly on the brain surface, tens of thousands of neurons are contributing to activity measured under electrodes. ANN models cannot approach the order of magnitude of neurons participating in phenomena at the scale of EEG, just as neither ANN nor SMNI can detail relatively smaller scale activity at the membrane or atomic levels. Attempts to do so likely would require statistical interpretations such as are made by SMNI; otherwise the output of the models would just replace the data collected from huge numbers of neuronal firings—a regression from 20th century Science back to Empiricism. Thus, as is the case in many physical sciences, the SMNI approach is to perform prior statistical analyses up to the scale of interest (here at EEG scales). The ANN approach must perform statistical analyses after processing its units.

While ANN models use simplified algebraic structures to represent real neurons, SMNI models develop the statistics of large numbers of realistic neurons representing huge numbers of synaptic interactions—there are 10^4 to 10^5 synapses per neuron. Furthermore, unlike most ANN approaches, SMNI accepts constraints on all its macrocolumnar averaged parameters to be taken from experimentally determined ranges of synaptic and neuronal interactions (Braitenberg, 1978; Shepherd, 1979; Sommerhoff, 1974; Vu and Krasne, 1992); there are no unphysical parameters. The stochastic and nonlinear nature of SMNI development is directly derived from experimentally observed synaptic interactions and from the mathematical development of observed minicolumns and macrocolumns of neurons (Fitzpatrick and Imig, 1980; Goldman and Nauta, 1977; Hubel and Wiesel, 1962; Hubel and Wiesel, 1977; Imig and Reale, 1980; Jones *et al*, 1978; Mountcastle, 1978). SMNI has required the use of mathematical physics techniques first published in the late 1970's in the context of developing an approach to multivariate nonlinear nonequilibrium statistical mechanics (Dekker, 1980; Grabert and Green, 1979; Graham, 1977a; Graham, 1977b; Langouche *et al*, 1982; Schulman, 1981).

Table I summarizes the comparison between the SMNI and ANN approaches.

Table I

3.3. Microscopic Neurons

A derivation has been given of the physics of chemical inter-neuronal and electrical intra-neuronal interactions (Ingber, 1982; Ingber, 1983). This derivation generalized a previous similar derivation (Shaw and Vasudevan, 1974). The derivation yields a short-time probability distribution of a given neuron firing due to its just-previous interactions with other neurons. Within τ_j~5−10 msec, the conditional probability that neuron j fires ($\sigma_j = +1$) or does not fire ($\sigma_j = -1$), given its previous interactions with k neurons, is

$$p_{\sigma_j} \approx \Gamma\,\Psi \approx \frac{\exp(-\sigma_j F_j)}{\exp(F_j) + \exp(-F_j)} ,$$

$$F_j = \frac{V_j - \sum_k a^*_{jk} v_{jk}}{\left(\pi \sum_{k'} a^*_{jk'}(v_{jk'}{}^2 + \phi_{jk'}{}^2)\right)^{1/2}} ,$$

$$a_{jk} = \frac{1}{2} A_{jk}(\sigma_k + 1) + B_{jk} . \qquad (8)$$

This particular derivation assumed that simple algebraic summation of excitatory depolarizations and inhibitory hyperpolarizations at the base of the inner axonal membrane determines the firing depolarization response of a neuron within its absolute and relative refractory periods (Shepherd, 1979). Many other neuroscientists agree that this assumption is reasonable when describing the activity of large ensembles of neocortical neurons, each one typically having many thousands of synaptic interactions. Recently, explicit experimental evidence has been found to support this premise (Jagadeesh *et al*, 1993).

Γ represents the "intra-neuronal" probability distribution, e.g., of a contribution to polarization achieved at an axon given activity at a synapse, taking into account averaging over different neurons, geometries, etc. Ψ represents the "inter-neuronal" probability distribution, e.g., of thousands of quanta of neurotransmitters released at one neuron's postsynaptic site effecting a (hyper-)polarization at another neuron's presynaptic site, taking into account interactions with neuromodulators, etc. This development is true for Γ Poisson, and for Ψ Poisson or Gaussian.

V_j is the depolarization threshold in the somatic-axonal region, v_{jk} is the induced synaptic polarization of E or I type at the axon, and ϕ_{jk} is its variance. The efficacy a_{jk}, related to the inverse conductivity across synaptic gaps, is composed of a contribution A_{jk} from the connectivity between neurons which is activated if the impinging k-neuron fires, and a contribution B_{jk} from spontaneous background noise.

Even at the microscopic scale of an individual neuron, with soma ≈ 10 μm, this conceptual framework assumes a great deal of statistical aggregation of molecular scales of interaction, e.g., of the biophysics of membranes, of thickness $\approx 5 \times 10^{-3}$ μm, composed of biomolecular leaflets of phospholipid molecules (Caillé *et al*, 1980; Scott, 1975; von der Heydt *et al*, 1981).

3.4. Mesoscopic Aggregation

This microscopic scale itself represents a high aggregation of sub-microscopic scales, aggregating effects of tens of thousands of quanta of chemical transmitters as they influence membrane scales at 5×10^{-3} μm. This microscopic scale is aggregated up to the mesoscopic scale (Ingber, 1981; Ingber, 1982; Ingber, 1983), using

$$P_q(q) = \int dq_1 dq_2 P_{q_1 q_2}(q_1, q_2) \delta[q - (q_1 + q_2)] \ .$$

The SMNI approach can be developed without recourse to borrowing paradigms or metaphors from other disciplines. Rather, in the course of a logical, nonlinear, stochastic development of aggregating neuronal and synaptic interactions to larger and larger scales, opportunities are taken to use techniques of mathematical physics to overcome several technical hurdles. After such development, advantage can be taken of associated collateral descriptions and intuitions afforded by such mathematical and physics techniques as they have been used in other disciplines, but paradigms and metaphors do not substitute for logical SMNI development.

3.5. Mesoscopic Interactions

For the purposes of mesoscopic and macroscopic investigation, this biological picture can be cast into an equivalent network. However, some aspects must not be simply cast away. At the microscopic scale, we retain independence of excitatory (E) and inhibitory (I) interactions, and the nonlinear development of probability densities. At the mesoscopic scale, we include the convergence and divergence of minicolumnar and macrocolumnar interactions, and use nearest-neighbor (NN) interactions to summarize interactions of up to 16th order in next-nearest neighbor interactions. At the macroscopic scale, we include long-ranged interactions as constraints on mesocolumns.

3.6. Mathematical Development

A derived mesoscopic Lagrangian $\underline{L}_M$ defines the short-time probability distribution of firings in a minicolumn (Mountcastle, 1978), composed of $\sim 10^2$ neurons, given its just previous interactions with all other neurons in its macrocolumnar surround. G is used to represent excitatory (E) and inhibitory (I) contributions. $\bar{G}$ designates contributions from both E and I.

The Lagrangian, essentially equal to the kinetic energy minus the potential energy, to first order in an expansion about the most likely state of a quantum or stochastic system, gives a global formulation and generalization of the well-known relation, *force equals mass times acceleration* (Feynman *et al*, 1963). In the neocortex, the velocity corresponds to the rate of firing of a column of neurons, and a potential is derived which includes nearest-neighbor interactions between columns. The Lagrangian formulation also accounts for the influence of fluctuations about such most likely paths of the evolution of a system, by use of a variational principle associated with its development. The Lagrangian is therefore often more useful than the Hamiltonian, essentially equal to the kinetic energy plus the potential energy, related to the energy in many systems. This is especially useful to obtain information about the system without solving the time-dependent problem; however, we also will describe neocortical phenomena requiring the full solution.

For the purposes of this paper, the essential technical details to note are that a conditional probability of columnar firing states is exponentially sensitive to a "Lagrangian" function, $\underline{L}$. The Lagrangian is a nonlinear function of a "threshold factor", F^G, itself a nonlinear function of firings $M^{\bar{G}}$. Columnar firings may be simply represented; e.g., a set of 80 excitatory firings may be represented as $M^E = -80$ if all are not firing, $M^E = 0$ if half are firing, or $M^E = 80$ if all are firing.

$$P_M = \prod_G P_M^G[M^G(r; t+\tau)|M^{\bar{G}}(r'; t)]$$

$$= \sum_{\sigma_j} \delta\left(\sum_{jE} \sigma_j - M^E(r; t+\tau) \right) \delta\left(\sum_{jI} \sigma_j - M^I(r; t+\tau) \right) \prod_j^N p_{\sigma_j}$$

$$\approx \prod_G (2\pi\tau g^{GG})^{-1/2} \exp(-N\tau \underline{L}_M^G) ,$$

$$P_M \approx (2\pi\tau)^{-1/2} g^{1/2} \exp(-N\tau \underline{L}_M) ,$$

$$\underline{L}_M = \underline{L}_M^E + \underline{L}_M^I = (2N)^{-1}(\dot{M}^G - g^G)g_{GG'}(\dot{M}^{G'} - g^{G'}) + M^G J_G/(2N\tau) - \underline{V}' ,$$

$$\underline{V}' = \sum_G \underline{V}''^G_{G'}(\rho\nabla M^{G'})^2 ,$$

$$g^G = -\tau^{-1}(M^G + N^G \tanh F^G) ,$$

$$g^{GG'} = (g_{GG'})^{-1} = \delta_G^{G'} \tau^{-1} N^G \mathrm{sech}^2 F^G ,$$

$$g = \det(g_{GG'}) ,$$

$$F^G = \frac{(V^G - a_{G'}^{|G|} v_{G'}^{|G|} N^{G'} - \frac{1}{2} A_{G'}^{|G|} v_{G'}^{|G|} M^{G'})}{(\pi[(v_{G'}^{|G|})^2 + (\phi_{G'}^{|G|})^2](a_{G'}^{|G|} N^{G'} + \frac{1}{2} A_{G'}^{|G|} M^{G'}))^{1/2}} ,$$

$$a_{G'}^G = \frac{1}{2} A_{G'}^G + B_{G'}^G , \tag{9}$$

where $A_{G'}^G$ and $B_{G'}^G$ are minicolumnar-averaged inter-neuronal synaptic efficacies, $v_{G'}^G$ and $\phi_{G'}^G$ are averaged means and variances of contributions to neuronal electric polarizations. $M^{G'}$ and $N^{G'}$ in F^G are afferent macrocolumnar firings, scaled to efferent minicolumnar firings by $N/N^* \sim 10^{-3}$, where N^* is the number of neurons in a macrocolumn, $\sim 10^5$. Similarly, $A_G^{G'}$ and $B_G^{G'}$ have been scaled by $N^*/N \sim 10^3$ to keep F^G invariant. This scaling is for convenience only, and does not affect any numerical details of calculations. J_G are Lagrange multipliers, originally introduced to include constraints imposed by long-ranged interactions, but further SMNI development makes these unnecessary.

3.7. Inclusion of Macroscopic Circuitry

The most important features of this development are described by the Lagrangian $\underline{L}^G$ in the negative of the argument of the exponential describing the probability distribution, and the "threshold factor" F^G in the means and variances, describing an important sensitivity of the distribution to changes in its variables and parameters.

To more properly include long-ranged fibers, when it is possible to numerically include interactions among macrocolumns, the J_G terms can be dropped, and more realistically replaced by a modified threshold factor F^G,

$$F^G = \frac{(V^G - a^{|G|}_{G'} v^{|G|}_{G'} N^{G'} - \frac{1}{2} A^{|G|}_{G'} v^{|G|}_{G'} M^{G'} - a^{\ddagger E}_{E'} v^E_{E'} N^{\ddagger E'} - \frac{1}{2} A^{\ddagger E}_{E'} v^E_{E'} M^{\ddagger E'})}{(\pi[(v^{|G|}_{G'})^2 + (\phi^{|G|}_{G'})^2](a^{|G|}_{G'} N^{G'} + \frac{1}{2} A^{|G|}_{G'} M^{G'} + a^{\ddagger E}_{E'} N^{\ddagger E'} + \frac{1}{2} A^{\ddagger E}_{E'} M^{\ddagger E'}))^{1/2}} ,$$

$$a^{\ddagger E}_{E'} = \frac{1}{2} A^{\ddagger E}_{E'} + B^{\ddagger E}_{E'} . \tag{10}$$

This is further modified for use in mesoscopic neural networks (MNN) discussed below.

Here, afferent contributions from $N^{\ddagger E}$ long-ranged excitatory fibers, e.g., cortico-cortical neurons, have been added, where $N^{\ddagger E}$ might be on the order of 10% of N^*: Of the approximately 10^{10} to 10^{11} neocortical neurons, estimates of the number of pyramidal cells range from 1/10 to 2/3. Nearly every pyramidal cell has an axon branch that makes a cortico-cortical connection; i.e., the number of cortico-cortical fibers is of the order 10^{10}.

3.8. Equivalent Nearest-Neighbor Interactions

Even representing a macrocolumn by 1000 minicolumnar interactions can become a formidable task. Including all such minicolumnar interactions would require 16th-order next-neighbor interactions. (Visualize a minicolumn in the center of a 33×33 grid of other minicolumns.) The numerical details of SMNI and experimentally determined ranges of neuronal parameters justify a nearest-neighbor approximation.

Nearest-neighbor (NN) interactions between mesocolumns are illustrated in Fig. 6. Afferent minicolumns of $\sim 10^2$ neurons are represented by the inner circles, and efferent macrocolumns of $\sim 10^5$ neurons by the outer circles. Illustrated are the NN interactions between a mesocolumn, represented by the thick circles, and its nearest neighbors, represented by thin circles. The area outside the outer thick circle represents the effective number of efferent macrocolumnar nearest-neighbor neurons. This is the number of neurons outside the macrocolumnar area of influence of the central minicolumn. This approximation, albeit successful (Ingber, 1983), can be replaced by the more sophisticated MNN algorithm (Ingber, 1992).

Figure 6

3.9. Minima Structure of Nonlinear Lagrangian

Examination of the minima structure of the spatially-averaged and temporally-averaged Lagrangian provides some quick intuitive details about most likely states of the system. This is supported by further analysis detailing the actual spatial-temporal minima structure. Illustrated in Fig. 7 is the surface of the static (time-independent) mesoscopic neocortical Lagrangian $\underline{\bar{L}}$ over the excitatory-inhibitory firing plane $(\bar{M}^E - \bar{M}^I)$, for a specific set of synaptic parameters. All points on the surface higher than $5 \times 10^{-3}/\tau$ have been deleted to expose this fine structure.

Figure 7

---

4. SMNI Applications—STM

4.1. Derivation of Short-Term Memory (STM)

Within a time scale of several seconds, the human brain can store only about 7±2 auditory *chunks* of information (Ericsson and Chase, 1982; Miller, 1956), modified as 4±2 chunks for visual memory (Zhang and Simon, 1985). There is more awareness in the neuroscience community that phenomena such as STM can have mechanisms at several neocortical scales (Eichenbaum, 1993).

At the mesoscopic scale, properties of STM— its capacity, duration and stability—have been calculated, and found to be consistent with empirical observations (Ingber, 1984).

The maximum STM capacity, consistent with the 7 ± 2 rule, is obtained when a "centering mechanism" is invoked. This occurs when the threshold factor F^G takes minima in the interior of M^G firing-space (i.e., not the corners of this space), as empirically observed. In the SMNI papers, the background noise $B^G_{G'}$ was reasonably adjusted to center F^G, with $J_G = 0$, but similar results could have been obtained by adjusting the influence of the long-ranged fibers $M^{\ddagger G}$.

To derive this capacity rule of STM, choose empirical ranges of synaptic parameters corresponding to a predominately excitatory case (EC), predominately inhibitory case (IC), and a balanced case (BC) in between. For each case, also consider a "centering mechanism" (EC', IC', BC'), whereby some synaptic parameter is internally manipulated, e.g., some chemical neuromodulation or imposition of patterns of firing, such that there is a maximal efficiency of matching of afferent and efferent firings:

$$M^G \approx M^{*G} \approx 0 \ . \tag{11}$$

This sets conditions on other possible minima of the *static* Lagrangian $\underline{\bar{L}}$.

4.2. Centering Mechanism

The centering effect is quite easy for the neocortex to accommodate. For example, this can be accomplished simply by readjusting the synaptic background noise from B^G_E to B'^G_E,

$$B'^G_E = \frac{V^G - (\frac{1}{2} A^G_I + B^G_I) v^G_I N^I - \frac{1}{2} A^G_E v^G_E N^E}{v^G_E N^G} \tag{12}$$

for both G = E and G = I.

This is modified straightforwardly when regional influences from $M^{\ddagger E}$ are included, as used in MNN. In general, B^G_E and B^G_I (and possibly A^G_E and A^G_I due to actions of neuromodulators, and J_G or $M^{\ddagger E}$ constraints from long-ranged fibers) are available to force the constant in the numerator to zero, giving an extra degree(s) of freedom to this mechanism. In this context, it is experimentally observed that the synaptic sensitivity of neurons engaged in selective attention is altered, presumably by the influence of chemical neuromodulators on postsynaptic neurons (Mountcastle *et al*, 1981).

The threshold factors greatly influence when and how smoothly the "step functions" $\tanh F^G_{IC}$ in $g^G(t)$ change $M^G(t)$ to $M^G(t+\theta)$. That is, assuming the drifts are a major driving force,

$$M(t+\Delta t) \approx M(t) - \frac{\Delta t}{\tau} \left(M^G(t) + N^G \tanh F^G(t) \right) \tag{13}$$

together with $\Delta t \le \tau$ can be used to approximately describe the influence on efferent firings from their afferent inputs.

$$M(t+\Delta t) \sim -N^G \tanh F^G(t) \tag{14}$$

can be used as a first approximation.

An important side result is to drive most probable states, i.e., small L which is driven largely by small F^G, to regions where

$$v^G_E A^G_E M^E \approx |v^G_I| A^G_I M^I \ . \tag{15}$$

Since I–I efficacies typically are relatively quite small, the conditional probability density under the centering mechanism is strongly peaked along the line

$$v^E_E A^E_E M^E \approx |v^E_I| A^E_I M^I \ . \tag{16}$$

These numerical details are used to advantage to further develop SMNI to dipole scales.

4.3. Applying the Centering Mechanism—"Inhibitory" State

A model of dominant inhibition describes how minicolumnar firings are suppressed by their neighboring minicolumns. For example, the averaged effect is established by inhibitory mesocolumns (IC) by setting $A^I_E = A^E_I = 2A^E_E = 0.01N^*/N$. Since there appears to be relatively little $I - I$ connectivity, set $A^I_I = 0.0001N^*/N$. The background synaptic noise is taken to be $B^E_I = B^I_E = 2B^E_E = 10B^I_I = 0.002N^*/N$. As nonvisual minicolumns are observed to have ~110 neurons and as there appear to be a predominance of E over I neurons, here take $N^E = 80$ and $N^I = 30$. Use $N^*/N = 10^3$, $J_G = 0$ (absence of long-ranged interactions), and V^G, $v^G_{G'}$, and $\phi^G_{G'}$ as estimated previously, i.e., $V^G = 10$ mV, $|v^G_{G'}| = 0.1$ mV, $\phi^G_{G'} = 0.1$ mV. The "threshold factors" F^G_{IC} for this IC model are then

$$F^E_{IC} = \frac{0.5\bar{M}^I - 0.25\bar{M}^E + 3.0}{\pi^{1/2}(0.1\bar{M}^I + 0.05\bar{M}^E + 9.80)^{1/2}} ,$$

$$F^I_{IC} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E - 45.8}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 11.2)^{1/2}} . \tag{17}$$

F^I_{IC} will cause efferent $\bar{M}^I(t + \Delta t)$ to fire for most afferent input firings, as it will be positive for most values of $\bar{M}^G(t)$ in F^I_{IC}, which is already weighted heavily with a term -45.8. Looking at F^E_{IC}, it is seen that the relatively high positive weights of afferent $\bar{M}^I$ require at least moderate values of positive afferent $\bar{M}^E$ to cause firings of efferent $\bar{M}^E$, diminishing the influence of $\bar{M}^E$.

Using the centering mechanism, $B'^E_E = 1.38$ and $B'^I_I = 15.3$, and F^G_{IC} is transformed to $F^G_{IC'}$,

$$F^E_{IC'} = \frac{0.5\bar{M}^I - 0.25\bar{M}^E}{\pi^{1/2}(0.1\bar{M}^I + 0.05\bar{M}^E + 10.4)^{1/2}} ,$$

$$F^I_{IC'} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 20.4)^{1/2}} . \tag{18}$$

4.4. Contours of "Inhibitory" State

In Fig. 8, contours of the Lagrangian illustrate "valleys" that trap firing-states of mesocolumns. ($\tau\underline{\bar{L}}$ can be as large as 10^3.) I.e., valleys of the Lagrangian represent peaks in the conditional probability density over the two M^G firing variables, giving the most likely states of columnar firing. No interior stable states are observed at scales of $\tau\underline{\bar{L}}$ ranging from 10^3 down to 10^{-2}, until the "centering mechanism" is turned on.

Figure 8

4.5. Applying the Centering Mechanism—"Excitatory" State

The other "extreme" of normal neocortical firings is a model of dominant excitation, effected by establishing excitatory mesocolumns (EC) by using the same parameters $\{ B^G_{G'}, v^G_{G'}, \phi^G_{G'}, A^I_I \}$ as in the IC model, but setting $A^E_E = 2A^I_E = 2A^E_I = 0.01N^*/N$. This yields

$$F^E_{EC} = \frac{0.25\bar{M}^I - 0.5\bar{M}^E - 24.5}{\pi^{1/2}(0.05\bar{M}^I + 0.10\bar{M}^E + 12.3)^{1/2}} ,$$

$$F^I_{EC} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.05\bar{M}^E + 7.24)^{1/2}} . \tag{19}$$

The negative constants in the numerators of F^G_{EC} enhance efferent firings for both E and I afferent inputs. However, the increased coefficient of $\bar{M}^E$ in F^E_{EC} (e.g., relative to its corresponding value in F^E_{IC}), and the fact that $\bar{M}^E$ can range up to $N^E = 80$, readily enhance excitatory relative to inhibitory firings throughout most of the range of $\bar{M}^E$. This is only a first approximation, and the full Lagrangian must be used to determine the actual evolution.

Using the centering mechanism, $B'^E_E = 10.2$ and $B'^I_I = 8.62$, and F^G_{EC} is transformed to $F^G_{EC'}$,

$$F^E_{EC'} = \frac{0.25\bar{M}^I - 0.5\bar{M}^E}{\pi^{1/2}(0.05\bar{M}^I + 0.10\bar{M}^E + 17.2)^{1/2}} ,$$

$$F^I_{EC'} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E}{\pi^{1/2}(0.001\bar{M}^I + 0.05\bar{M}^E + 12.4)^{1/2}} . \tag{20}$$

4.6. Contours of "Excitatory" State

In Fig. 9, contours of the Lagrangian illustrate ''valleys'' that trap firing-states of mesocolumns. ($\tau\underline{\bar{L}}$ can be as large as 10^3.) No interior stable states are observed at scales of $\tau\underline{\bar{L}}$ ranging from 10^3 down to 10^{-2}, until the "centering mechanism" is turned on.

Figure 9

4.7. Applying the Centering Mechanism—"Balanced" State

Now it is natural to examine a balanced case intermediate between IC and EC, labeled BC. This is accomplished by changing $A^E_E = A^I_E = A^E_I = 0.005N^*/N$. This yields

$$F^E_{BC} = \frac{0.25\bar{M}^I - 0.25\bar{M}^E - 4.50}{\pi^{1/2}(0.050\bar{M}^E + 0.050\bar{M}^I + 8.30)^{1/2}} ,$$

$$F^I_{BC} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.050\bar{M}^E + 7.24)^{1/2}} . \tag{21}$$

Here the constant in the numerator of F^E_{BC}, while still negative to promote E efferent firings, is much greater than that in F^E_{EC}, thereby decreasing the net excitatory activity to a more moderate level. A similar argument applies in comparing F^I_{BC} to F^I_{IC}, permitting a moderate level of inhibitory firing.

Applying the centering mechanism to BC, $B'^E_E = 0.438$ and $B'^I_I = 8.62$, and F^G_{BC} is transformed to $F^G_{BC'}$,

$$F^E_{BC'} = \frac{0.25\bar{M}^I - 0.25\bar{M}^E}{\pi^{1/2}(0.050\bar{M}^I + 0.050\bar{M}^E + 7.40)^{1/2}} ,$$

$$F^I_{BC'} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E}{\pi^{1/2}(0.001\bar{M}^I + 0.050\bar{M}^E + 12.4)^{1/2}} . \tag{22}$$

4.8. Contours of "Balanced" State

No interior stable states are observed at scales of $\tau\underline{\bar{L}}$ ranging from 10^3 down to 10^{-2}, until the "centering mechanism" is turned on, as illustrated in Fig. 10. In (a), Contours for values less than 0.04 are drawn for $\tau\underline{\bar{L}}_{BC}$. The $\bar{M}^E$ axis increases to the right, from $-N^E = -80$ to $N^E = 80$. The $\bar{M}^I$ axis increases to the right, from $-N^I = -30$ to $N^I = 30$. In each cluster, the smaller values are closer to the center. In (b), Contours for values less than 0.04 are drawn for $\tau\underline{\bar{L}}_{BC'}$.

Figure 10

Thus, the BC′ model exhibits several stable states, which are calculated as minima of the Lagrangian, or equivalently as peaks in the conditional probability distribution. These stable states are in reasonable ranges of normal neocortex. This is not the case for the EC′ or IC′ models.

4.9. Modeling Visual Cortex STM

As illustrated in Fig. 11, when N = 220, modeling the number of neurons per minicolumn in visual neocortex, then a maximum of only 5-6 minima are found within a given cluster, consistent with visual STM. These minima are narrower, consistent with the sharpness required to store visual patterns.

Figure 11

4.10. Primacy Versus Recency Rule

SMNI also presents an explanation, or at least an upper statistical constraint, on the primacy versus recency rule observed in serial processing (Murdock, 1983). First-learned items are recalled most error-free, and last-learned items are still more error-free than those in the middle. That is, the deepest minima are more likely first accessed, while the more recent memories or newer patterns have synaptic parameters most recently tuned or are more actively rehearsed.

Note that for visual cortex, presentation of 7±2 items would have memories distributed among different clusters, and therefore the recency effect should not be observed. Instead the 4±2 rule should dictate the number of presented items.

4.11. STM Stability and Duration

To be more rigorous, stationary states are located, and possible hysteresis and/or jumps between local minima are explicitly calculated. Detailed calculations identify the inner valleys of the parabolic trough with stable short-term-memory states having durations on the order of tenths of a second (Ingber, 1984; Ingber, 1985a).

Stability is investigated by

$$\delta \dot{M}^G \approx -N^2 \underline{\bar{L}}_{,GG'} \delta M^{G'} \ . \tag{23}$$

Therefore, minima of the static Lagrangian $\underline{\bar{L}}$ are minima of the dynamic transient system defined by L. The time of first passage is calculated as (Agarwal and Shenoy, 1981)

$$t_{vp} \approx \pi N^{-2} \left(|\underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_p) | \underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_v) \right)^{-1/2}$$
$$\times \exp\{CN\tau[\underline{\bar{L}}(\ll \bar{M} \gg_p) - \underline{\bar{L}}(\ll \bar{M} \gg_v)]\} \ . \tag{24}$$

For $\tau \underline{\bar{L}} \sim 10^{-2}$, the *only* values found for all three cases of firing, the time of first passage t_{vp} is found to be several tenths of second for jumps among most minima, up to 9. There is hysteresis for deeper valleys at 10th-11th minima of $\underline{\bar{L}}_{FBC'}$ at the corners of the $\bar{M}^G$ plane. The hysteresis occurs in about a few minutes, which is too long to affect the 7 ± 2 rule. This result is *exponentially* sensitive to N in Φ/D, and *exponentially* sensitive to $(N^* N)^{1/2}$ in F^G, the "threshold factor."

Use is made in later development of EEG analyses of the discovered nature of the line of stable minima lying in a deep parabolic trough, across a wide range of cases of extreme types of firings.

5. SMNI Applications—EEG

5.1. EEG Phenomena—Euler-Lagrange Approximation

When dealing with tens of thousands to millions of neurons, it seems reasonable, as a first approximation, to consider only most likely spatial-temporal states of neocortex, without including the complications of statistical processes. Since a Lagrangian is developed by SMNI, this can be quantitatively articulated.

The variational principle permits derivation of the Euler-Lagrange equations. These equations are then linearized about a given local minima to investigate oscillatory behavior (Ingber, 1983; Ingber, 1985a). Here, long ranged constraints in the form of Lagrange multipliers J_G were used to efficiently search for minima, corresponding to roots of the Euler-Lagrange equations. This illustrates how macroscopic constraints can be imposed on the mesoscopic and microscopic systems. About a given minima, a Taylor expansion is made, and spatial-temporal Fourier transformations are made to derive a "dispersion relation."

$$0 = \hat{\delta} \underline{L}_F = \underline{L}_{F,\dot{G}:t} - \hat{\delta}_G \underline{L}_F$$

$$\approx - \underline{f}_{|G|} \ddot{\underline{M}}^{|G|} + \underline{f}^1_G \dot{\underline{M}}^{G^\neg} - \underline{g}_{|G|} \nabla^2 \underline{M}^{|G|} + \underline{b}_{|G|} \underline{M}^{|G|} + \underline{b}\, \underline{M}^{G^\neg} \ ,$$

$$G^\neg \neq G \ ,$$

$$\underline{M}^G = M^G - \ll \bar{M}^G \gg \ ,$$

$$\underline{M}^G = \mathrm{Re}\, \underline{M}^G_{osc} \exp[-i(\underline{\xi} \cdot r - \omega t)] \ ,$$

$$\underline{M}^G_{osc}(r, t) = \int d^2\underline{\xi} d\omega\, \hat{\underline{M}}^G_{osc}(\underline{\xi}, \omega) \exp[i(\underline{\xi} \cdot r - \omega t)] \ ,$$

$$\omega\tau = \pm\{ -1.86 + 2.38(\xi\rho)^2 ; -1.25i + 1.51i(\xi\rho)^2 \} \ , \ \xi = |\underline{\xi}| \ . \tag{25}$$

It is calculated that

$$\omega \sim 10^2 \ \mathrm{sec}^{-1} \ , \tag{26}$$

which is equivalent to

$$\nu = \omega/(2\pi) \ 16 \ \mathrm{cps\ (Hz)} \ . \tag{27}$$

This is approximately within the experimentally observed ranges of the *alpha* and *beta* frequencies.

5.2. Nearest-Neighbor Contours

As illustrated in Fig. 12, a numerical calculation of the coefficients, $\underline{g}_G$, of nearest-neighbor interactions, $(\nabla M^G)^2$, shows that SMNI can support/describe both spreading activation of firings as well as local containment of firings. The $\underline{g}_G$ terms are responsible for the spatial dependence of the EEG dispersion relations.

Figure 12

5.3. Propagation of Information

The propagation velocity v is calculated from

$$v = d\omega/d\xi \approx 1 \ \mathrm{cm/sec} \ , \ \xi \sim 30\rho \ , \tag{28}$$

which tests the NN interactions, including the spatial content of SMNI at this stage of development. Thus, within 10^{-1} sec, short-ranged interactions over several minicolumns of 10^{-1} cm may simultaneously interact with long-ranged interactions over tens of cm, since the long-ranged interactions are speeded by myelinated fibers and have velocities of 600–900 cm/sec. In other words, interaction among different neocortical modalities, e.g., visual, auditory, etc., may simultaneously interact within the same time scales, as observed.

This propagation velocity is consistent with the observed movement of attention (Tsal, 1983) and with the observed movement of hallucinations across the visual field (Cowan, 1982), which moves at ~1/2 mm/sec, about 5 times as slow as v. (The observed movement is about 8 msec/°, and a macrocolumn of extent a mm processes 180° of visual field.)

5.4. Local and Global EEG

The derived mesoscopic dispersion relations also are consistent with other global macroscopic dispersion relations, described by long-range fibers interacting across regions (Ingber, 1985a; Nunez, 1981). Other investigators have detailed specific neuronal mechanisms for local generators of EEG (Lopes da Silva, 1991).

This SMNI model yields oscillatory solutions consistent with the alpha rhythm, i.e., $\omega \approx 10^2$ $\sec^{-1}$, equivalent to $\nu = \omega/(2\pi) \approx 16$ Hz. This suggests that these complementary local and global theories may be confluent, considered as a joint set of dispersion relations evolving from the most likely trajectories of a joint Lagrangian, referred to as the "equations of motion," but linearly simplified in neighborhoods of minima of the stationary Lagrangian.

These two approaches, i.e., local mesocolumnar versus global macrocolumnar, give rise to important alternative conjectures:

(1) Is the EEG global resonance of primarily long-ranged cortical interactions? If so, can relatively short-ranged local firing patterns effectively modulate this frequency and its harmonics, to enhance their information processing across macroscopic regions?

(2) Or, does global circuitry imply boundary conditions on collective mesoscopic states of local firing patterns, and is the EEG a manifestation of these collective local firings?

(3) Or, is the truth some combination of (1) and (2) above? For example, the possibility of generating EEG rhythms from multiple mechanisms at multiple scales of interactions, e.g., as discussed above, may account for weakly damped oscillatory behavior in a variety of physiological conditions.

This theory has allowed the local and global approaches to complement each other at a common level of formal analysis, i.e., yielding the same dispersion relations derived from the "equations of motion," analogous to $\sum$(forces) = d(momentum)/dt describing mechanical systems.

6. Nonlinear Stochastic Fit of SMNI to EEG

6.1. Linearization Aids Probability Development

More recently, the nonlinear statistical description of SMNI was further developed up to the regional scales described as "dipoles." This permitted fitting of the full SMNI probability distribution to EEG data (Ingber, 1991).

As illustrated in Figs. 8-10, previous STM studies have detailed that the predominant physics of short-term memory and of (short-fiber contribution) to EEG phenomena takes place in a narrow "parabolic trough" in M^G-space, roughly along a diagonal line. That is, $\tau \underline{L}_M$ can vary by as much as 10^5 from the highest peak to the lowest valley in M^G-space. Therefore, it is reasonable to assume that a single independent firing variable might offer a crude description of this physics. Furthermore, the scalp potential Φ can be considered to be a function of this firing variable. In an abbreviated notation subscripting the time-dependence,

$$\Phi_t - \ll \Phi \gg = \Phi(M_t^E, M_t^I) \approx a(M_t^E - \ll M^E \gg) + b(M_t^I - \ll M^I \gg) \,, \tag{29}$$

where a and b are constants of the same sign, and $\ll \Phi \gg$ and $\ll M^G \gg$ represent a minima in the trough. This determines an SMNI approach to study EEG under conditions of selective attention.

Laplacian techniques help to better localize sources of activity, and thereby present data more suitable for modeling. Then Φ is more directly related to columnar firings, instead of representing the electric potential produced by such activity.

6.2. EEG Macrocolumnar Lagrangian

Again, aggregation is performed,

$$P_\Phi[\Phi_{t+\Delta t}|\Phi_t] = \int dM^E_{t+\Delta t} dM^I_{t+\Delta t} dM^E_t dM^I_t P_M[M^E_{t+\Delta t}, M^I_{t+\Delta t}|M^E_t, M^I_t]$$

$$\delta[\Phi_{t+\Delta t} - \Phi(M^E_{t+\Delta t}, M^I_{t+\Delta t})]\delta[\Phi_t - \Phi(M^E_t, M^I_t)] \ . \qquad (30)$$

Under conditions of selective attention, within the parabolic trough along a line in $\bar{M}^G$ space, the parabolic shapes of the multiple minima, ascertained by the stability analysis, justifies a form

$$P_\Phi = (2\pi\sigma^2 dt)^{-1/2} \exp[-(dt/2\sigma^2) \int dx L_\Phi] \ ,$$

$$L_\Phi = \frac{1}{2} |\partial\Phi/\partial t|^2 - \frac{1}{2} c^2 |\partial\Phi/\partial x|^2 - \frac{1}{2} \omega_0^2 |\Phi|^2 - F(\Phi) \ , \qquad (31)$$

where $F(\Phi)$ contains nonlinearities away from the trough, and where σ^2 is on the order of $1/N$, given the derivation of $\underline{L}_M$ above.

6.3. EEG Variational Equation

Previous calculations of EEG phenomena showed that the short-fiber contribution to the alpha frequency and the movement of attention across the visual field are consistent with the assumption that the EEG physics is derived from an average over the fluctuations σ of the system. This is described by the Euler-Lagrange equations derived from the variational principle possessed by L_Φ, more properly by the "midpoint-discretized" L_Φ, with its Riemannian terms, discussed below. Hence,

$$0 = \frac{\partial}{\partial t} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial t)} + \frac{\partial}{\partial x} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial x)} - \frac{\partial L_\Phi}{\partial\Phi} \ . \qquad (32)$$

When expressed in the firing state variables, this leads to the same results published earlier (Ingber, 1983).

The result for the Φ equation is:

$$\frac{\partial^2\Phi}{\partial t^2} - c^2 \frac{\partial^2\Phi}{\partial x^2} + \omega_0^2\Phi + \frac{\partial F}{\partial\Phi} = 0 \ . \qquad (33)$$

If the identification

$$\frac{\partial F}{\partial\Phi} = \Phi f(\Phi) \qquad (34)$$

is made, then

$$\frac{\partial^2\Phi}{\partial t^2} - c^2 \frac{\partial^2\Phi}{\partial x^2} + [\omega_0^2 + f(\Phi)]\Phi = 0 \qquad (35)$$

is recovered, i.e., the dipole-like string equation, eqn. (1).

The previous application of the variational principle was at the scale of minicolumns and, with the aid of nearest-neighbor interactions, the spatial-temporal Euler-Lagrange equation gave rise to dispersion relations consistent with STM experimental observations. Here, the scale of interactions is at the macrocolumnar level, and spatial interactions must be developed taking into account specific regional circuitries.

6.4. Macroscopic Coarse-Graining

Now the issue posed previously, how to mathematically justify the intuitive coarse-graining of Φ to get $\Phi^\dagger$, can be approached.

In L_Φ above, consider terms of the form

$$\int \Phi^2 dx = \int dx \sum_n^\infty \sum_m^\infty G_n G_m \sin k_n x \sin k_m x$$

$$= \sum_n \sum_m G_n G_m \int dx \sin k_n x \sin k_m x$$

$$= (2\pi/R) \sum_n G_n^2 \ . \tag{36}$$

By similarly considering all terms in L_Φ, a short-time probability distribution for the change in node n is defined,

$$p_n[G_n(t + \Delta t)|G_n(t)] \ . \tag{37}$$

Note that in general the $F(\Phi)$ term in L_Φ will require coupling between G_n and G_m, $n \neq m$. This defines

$$P_\Phi = p_1 p_2 \cdots p_\infty \ . \tag{38}$$

Now a coarse-graining can be defined that satisfies physical and mathematical rigor:

$$P_{\Phi^\dagger} = \int dk_{M+1} dk_{M+2} \cdots dk_\infty \, p_1 p_2 \cdots p_M p_{M+1} p_{M+2} \cdots p_\infty \ . \tag{39}$$

Since SMNI is developed in terms of *bona fide* probability distributions, variables which are not observed can be integrated out. The integration over the fine-grained wave-numbers tends to smooth out the influence of the k_n's for $n > M$, effectively ‘‘renormalizing’’

$$G_n \rightarrow G^\dagger{}_n \ ,$$

$$\Phi \rightarrow \Phi^\dagger \ ,$$

$$L_\Phi \rightarrow L^\dagger{}_\Phi \ . \tag{40}$$

This development shows how this probability approach to EEG specifically addresses experimental issues at the scale of the more phenomenological dipole model.

6.5. Development of Macrocolumnar EEG Distribution

Advantage can be taken of the prepoint discretization, where the postpoint $M^G(t + \Delta t)$ moments are given by

$$m \equiv < \Phi_\nu - \phi > = a < M^E > + b < M^I > = a g^E + b g^I \ ,$$

$$\sigma^2 \equiv < (\Phi_\nu - \phi)^2 > - < \Phi_\nu - \phi >^2 = a^2 g^{EE} + b^2 g^{II} \ . \tag{41}$$

Note that the macroscopic drifts and diffusions of the Φ's are simply linearly related to the mesoscopic drifts and diffusions of the M^G's. For the prepoint $M^G(t)$ firings, the same linear relationship in terms of $\{ \phi, a, b \}$ is assumed.

The data being fit are consistent with invoking the “centering” mechanism. Therefore, for the prepoint $M^E(t)$ firings, the nature of the parabolic trough derived for the STM Lagrangian is taken advantage of, and

$$M^I(t) = c M^E(t) \ , \tag{42}$$

where the slope c is determined for each electrode site. This permits a complete transformation from M^G variables to Φ variables.

Similarly, as appearing in the modified threshold factor F^G, each regional influence from electrode site μ acting at electrode site ν, given by afferent firings $M^{\ddagger E}$, is taken as

$$M^{\ddagger E}_{\mu \rightarrow \nu} = d_\nu M^E_\mu (t - T_{\mu \rightarrow \nu}) \ , \tag{43}$$

where d_ν are constants to be fitted at each electrode site, and $T_{\mu \rightarrow \nu}$ is the delay time estimated for inter-electrode signal propagation, typically on the order of one to several multiples of $\tau = 5$ msec. In future fits, some experimentation will be performed, taking the T's as parameters.

This defines the conditional probability distribution for the measured scalp potential Φ_ν,

$$P_\nu[\Phi_\nu(t+\Delta t)|\Phi_\nu(t)] = \frac{1}{(2\pi\sigma^2\Delta t)^{1/2}} \exp(-L_\nu \Delta t) ,$$

$$L_\nu = \frac{1}{2\sigma^2} (\dot{\Phi}_\nu - m)^2 . \tag{44}$$

The probability distribution for all electrodes is taken to be the product of all these distributions:

$$P = \prod_\nu P_\nu ,$$

$$L = \sum_\nu L_\nu . \tag{45}$$

6.6. Development of EEG Dipole Distribution

The model derived for $P[M^G(t+\Delta t)|M^{\bar{G}}(t)]$ is for a macrocolumnar-averaged minicolumn. Hence, it is expected to be a reasonable approximation to represent a macrocolumn, scaled to its contribution to Φ_ν. Hence L is used to represent this macroscopic regional Lagrangian, scaled from its mesoscopic mesocolumnar counterpart $\underline{L}$.

However, the expression for P_ν is extended to include the dipole assumption, to also use this expression to represent several to many macrocolumns present in a region under an electrode: A macrocolumn has a spatial extent of about a millimeter. Often most data represents a resolution more on the order of up to several centimeters, many macrocolumns.

A scaling is tentatively assumed, to use the expression for the macrocolumnar distribution for the electrode distribution, then checked to see if the fits are consistent with this scaling. One argument in favor of this procedure is that it is generally acknowledged that only a small fraction of firings, those that fire coherently, are responsible for the observed activity being recorded.

The results obtained here seem to confirm that this approximation is in fact quite reasonable. For example, for the nonvisual neocortex, taking the extreme of permitting only unit changes in M^G firings, it seems reasonable to always be able to map the observed electric potential values Φ from a given electrode onto a mesh a fraction of $4N^E N^I \approx 10^4$.

It is expected that the use of MNN will make this scaling approximation unnecessary, because these multiple scales can now be explicitly calculated and their interactions explicitly included.

6.7. Key Indicators of EEG Correlates to Brain States

The SMNI probability distribution can be used directly to model EEG data, instead of using just the variational equations. Some important features not previously considered in this field were (Ingber, 1991):

- *Intra-Electrode Coherency* is determined by the standard deviations of excitatory and inhibitory firings under a given electrode as calculated using SMNI. Once the SMNI parameters are fit, then these firings are calculated as transformations on the EEG data, as described in terms of the SMNI derived probability distributions. This is primarily a measure of coherent columnar activity.
- *Inter-Electrode Circuitry* is determined by the fraction of available long-ranged fibers under one electrode which actively contribute to activity under another electrode, within the resolution of time given in the data (which is typically greater than or equal to the relative refractory time of most neurons, about 5−10 msec). This is primarily a measure of inter-regional activity/circuitry. Realistic delays can be modeled and fit to data.

The electrical potential of each electrode, labeled by G, is represented by its dipole-like nature, $M^G(t)$, which is influenced by its underlying columnar activity as well as its interactions with other electrodes, $M^{G'}$, $G \neq G'$. This can be expressed as:

$$\dot{M}^G = g^G + \hat{g}^G_i \eta^i ,$$

$$g^G = -\tau^{-1}(M^G + N^G \tanh F^G) \ ,$$

$$\hat{g}^G_i = (N^G/\tau)^{1/2} \mathrm{sech} F^G \ ,$$

$$F^G = \frac{(V^G - a^{|G|}_{G'} v^{|G|}_{G'} N^{G'} - \frac{1}{2} A^{|G|}_{G'} v^{|G|}_{G'} M^{G'} - a^{\ddagger E}_{E'} v^E_{E'} N^{\ddagger E'} - \frac{1}{2} A^{\ddagger E}_{E'} v^E_{E'} M^{\ddagger E'})}{(\pi[(v^{|G|}_{G'})^2 + (\phi^{|G|}_{G'})^2](a^{|G|}_{G'} N^{G'} + \frac{1}{2} A^{|G|}_{G'} M^{G'} + a^{\ddagger E}_{E'} N^{\ddagger E'} + \frac{1}{2} A^{\ddagger E}_{E'} M^{\ddagger E'}))^{1/2}} \ . \tag{46}$$

The equivalent Lagrangian is used for the actual fits.

6.8. Pilot Study—EEG Correlates to Behavioral States

A pilot study (Ingber, 1991) used sets of EEG data, given to the author by Henri Begleiter, Neurodynamics Laboratory at the State University of New York Health Center at Brooklyn, which were obtained from subjects while they were reacting to pattern-matching “odd-ball”-type tasks requiring varying states of selective attention taxing their short-term memory (Porjesz and Begleiter, 1990). Based on psychiatric and family-history evaluations, 49 subjects were classified into two groups, 25 possibly having high-risk and 24 possibly having low-risk genetic propensities to alcoholism.

After each subject’s data set, representing averages over 190 points of EEG data, was fitted to its probability distribution, the data were again filtered through the fitted Lagrangian, and the mean and mean-square values of M^G were recorded as they were calculated from Φ. Then, the group’s averages and standard deviations were calculated, the latter simply from $\{[< (M^G)^2 > - < M^G >^2]n/(n-1)\}^{1/2}$, where $n = 49$. This procedure gives the means and standard deviations of the effective firings, M^G, aggregated from all subjects under each electrode, as well as the weights d of the time-delayed inter-electrode inputs M^{*E}.

Although M^G were permitted to roam throughout their physical ranges of $\pm N^E = \pm 80$ and $\pm N^I = \pm 30$ (in the nonvisual neocortex, true for all these regions), their observed effective regional-averaged firing states were observed to obey the centering mechanism. That is, this numerical result is consistent with the assumption that the most likely firing states are centered about the region $M^G \approx 0 \approx M^{*E}$ in F^G.

Fitted parameters were used to calculate equivalent columnar firing states and time delays between regions. No statistical differences were observed between the total group, the high-risk group, and the low-risk group. It now is generally understood that a genetic predisposition for alcoholism, if it exists, is a multi-factor phenomena requiring a very large study group.

6.9. Sample Results — Total Group

As detailed in Table II, means and standard deviations of averages over EEG recordings from 49 subjects, representing 190 points of data per subject, are consistent with the centering mechanism during selective attention tasks. Under each electrode the means and standard deviations of M^G are given. Also given for each electrode are the means and standard deviations of the individual-subject standard deviations, here labeled as σ, aggregated from each subject. The physical bounds for all M^E under these nonvisual regions are $\pm N^E = \pm 80$. Also given are the weights d of the regional time-delayed contributions dM^{*E}. The physical bounds for all M^E and M^{*E} under these nonvisual regions are $\pm N^E = \pm N^{*E} = \pm 80$; the physical bounds for all M^I are $\pm N^I = \pm 30$.

Table II

6.10. Future Sources of Data

Laplacian techniques help to better localize sources of activity, and thereby become more suitable for the MNN modeling. By virtue of Poisson’s equation,

$$\nabla \cdot \sigma(\tilde{r}) \nabla \phi(\tilde{r}, t) = s(\tilde{r}, t) \ , \tag{47}$$

where $\sigma(\tilde{r})$ is the tissue conductivity, $s(\tilde{r}, t)$ is the macrocolumnar-averaged microsource typically $\approx 0.1-1$ μA/mm^3, and $\phi(\tilde{r}, t)$ is the micropotential. Thus, the Laplacian of the EEG potential, directly related to ϕ, presents an EEG variable directly related to columnar firings. Φ, instead of representing the EEG electric potential, then is a direct measure of current flows described by temporal changes in firings M^G.

There is at least one problem associated with the use of Laplacian filtering, a second-order differentiation process. Differentiation generally tends to emphasize the noise contributions owing to their typically sharper curvatures. As in similar modeling projects in combat analyses (Ingber, 1993a; Ingber, Fujio, and Wehner, 1991; Ingber and Sworder, 1991) and finance (Ingber, 1990; Ingber, Wehner *et al*, 1991), this can be alleviated by using the path-integral Lagrangian to determine the proper meshes. Then, this resolution of data must be available for further modeling.

Future work will fit single-sweep data to the SMNI model, not averaged evoked potential (AEP) data over 100's of trials as in the previous study. As the SMNI model suggests, EEG "noise" likely possesses non-constant structure developed from the statistical mechanics of neocortical interactions, and the model should be fit directly to the single-sweep data to be able to extract the maximum signal.

6.11. Precursor Information of Epileptic Seizures

As an example of how SMNI might be applied in a clinical setting, its spatial and temporal aspects can be become decision aids for treatment of epilepsy.

Improve Temporal Prediction of Seizures

If these SMNI techniques can find patterns of such upcoming activity some time before the trained eye of the clinician, then the costs of time and pain in preparation for surgery can be reduced. This project will determine inter-electrode and intra-electrode activities prior to spike activity to determine likely electrode circuitries highly correlated to the onset of seizures. This can only do better than simple averaging or filtering of such activity, as typically used as input to determine dipole locations of activity prior to the onset of seizures.

Improve Spatial Resolution

If a subset of electrode circuitries are determined to be highly correlated to the onset of seizures, then their associated regions of activity can be used as a first approximate of underlying dipole sources of brain activity affecting seizures. This first approximate may be better than using a spherical head model to deduce such a first guess. Such first approximates can then be used for more realistic dipole source modeling, including the actual shape of the brain surface to determine likely localized areas of diseased tissue.

7. Relevant Mathematical/Numerical Aspects

7.1. Adaptive Simulated Annealing (ASA)

Optimal fitting of nonlinear/stochastic models to EEG data (Ingber, 1991) requires a very powerful optimization algorithm. Adaptive simulated annealing (ASA) (Ingber, 1989; Ingber, 1993b; Ingber, 1993c), an enhanced algorithm derived from very fast simulated annealing (VFSR) (Ingber, 1989), was developed for such applications.

This algorithm fits empirical data to a theoretical cost function over a D-dimensional parameter space, adapting for varying sensitivities of parameters during the fit. Heuristic arguments have been developed to demonstrate that this algorithm is faster than the fast Cauchy annealing, $T_i = T_0/k$, and much faster than Boltzmann annealing, $T_i = T_0/\ln k$.

For parameters

$$\alpha^i_k \in [A_i, B_i] , \tag{48}$$

sampling with the random variable x^i,

$$x^i \in [-1, 1] ,$$

$$\alpha^i_{k+1} = \alpha^i_k + x^i(B_i - A_i) , \tag{49}$$

define the generating function

$$g_T(x) = \prod_{i=1}^{D} \frac{1}{2\ln(1 + 1/T_i)(|x^i| + T_i)} \equiv \prod_{i=1}^{D} g_T^i(x^i) ,$$

in terms of parameter "temperatures"

$$T_i = T_{i0} \exp(-c_i k^{1/D}) . \tag{50}$$

The cost-functions $\underset{\sim}{L}$ under consideration are of the form

$$h(M; \alpha) = \exp(-\underset{\sim}{L}/T) ,$$

$$\underset{\sim}{L} = L\Delta t + \frac{1}{2} \ln(2\pi\Delta t g_t^2) , \tag{51}$$

where L is a Lagrangian with dynamic variables M(t), and parameter-coefficients α to be fit to data. g_t is the determinant of the metric, and T is the cost "temperature."

For several test problems, ASA has been shown to be orders of magnitude more efficient than other similar techniques, e.g., genetic algorithms. ASA has been applied to many complex systems, including specific problems in neuroscience, finance (Ingber, 1990; Ingber, Wehner *et al*, 1991) and combat systems (Ingber, 1993a; Ingber, Fujio, and Wehner, 1991; Ingber and Sworder, 1991).

7.2. Path Integral of Nonlinear Nonequilibrium Processes

7.2.1. Representations of Path Integral

There are three mathematically equivalent representations of multivariate multiplicative Gaussian-Markovian systems. Their equivalence for quite arbitrarily nonlinear systems was established in the late 1970's by mathematical physicists. SMNI was the first physical system to take advantage of the path-integral representation, deriving it from neuronal interactions. All three representations have been useful in the SMNI development.

The Langevin Rate-Equation exhibits a stochastic equation, wherein drifts can be arbitrarily nonlinear functions, and multiplicative noise is added.

$$M(t + \Delta t) - M(t) \sim \Delta t\, f[M(t)] ,$$

$$\dot{M} = \frac{dM}{dt} \sim f ,$$

$$\dot{M} = f + \hat{g}\eta ,$$

$$< \eta(t) >_\eta = 0 , < \eta(t)\eta(t') >_\eta = \delta(t - t') . \tag{52}$$

The Diffusion Equation is another equivalent representation of Langevin equations. The first moment "drift" is identified as f, and the second moment "diffusion," the variance, is identified as $\hat{g}^2$.

$$\frac{\partial P}{\partial t} = \frac{\partial(-fP)}{\partial M} + \frac{1}{2} \frac{\partial^2(\hat{g}^2 P)}{\partial M^2} . \tag{53}$$

The Path-Integral Lagrangian represents yet another equivalent representation of Langevin equations. It has been demonstrated that the drift and diffusion, in addition to possibly being quite general nonlinear functions of the independent variables and of time explicitly, may also be explicit functions of the distribution P itself (Wehner and Wolfer, 1987).

$$P[M_{t+\Delta t}|M_t] = (2\pi\hat{g}^2\Delta t)^{-1/2} \exp(-\Delta t L) ,$$

$$L = (\dot{M} - f)^2/(2\hat{g}^2) ,$$

$$P[M_t|M_{t_0}] = \int \cdots \int dM_{t-\Delta t} dM_{t-2\Delta t} \cdots dM_{t_0+\Delta t}$$

$$\times P[M_t|M_{t-\Delta t}] P[M_{t-\Delta t}|M_{t-2\Delta t}] \cdots P[M_{t_0+\Delta t}|M_{t_0}] \ ,$$

$$P[M_t|M_{t_0}] = \int \cdots \int \underline{D}M \exp(-\sum_{s=0}^{u} \Delta t L_s) \ ,$$

$$\underline{D}M = (2\pi \hat{g}_0^2 \Delta t)^{-1/2} \prod_{s=1}^{u} (2\pi \hat{g}_s^2 \Delta t)^{-1/2} dM_s \ ,$$

$$\int dM_s \to \sum_{\alpha=1}^{N} \Delta M_{\alpha s} \ , M_0 = M_{t_0} \ , M_{u+1} = M_t \ . \tag{54}$$

This representation is useful for fitting stochastic data to parameters in L.

7.2.2. Calculation of Path Integral

Given a form for L, we use the path-integral to calculate the long-time distribution of variables. This is impossible in general to calculate in closed form, and we therefore must use numerical methods. Techniques and codes have been developed for calculating highly nonlinear multivariate Lagrangians. A code has been developed to perform a detailed non-Monte Carlo path integral (Ingber, 1994), to calculate the long-time probability distribution of rather arbitrarily nonlinear Lagrangians, based on a histogram approach (Ingber, Fujio, and Wehner, 1991; Wehner and Wolfer, 1983a; Wehner and Wolfer, 1983b; Wehner and Wolfer, 1987).

The path-integral calculation of the long-time distribution, in addition to being a predictor of upcoming information, provides an internal check that the system can be well represented as a nonlinear Gaussian-Markovian system. This calculation also serves to more sensitively distinguish among alternative Lagrangians which may approximately equally fit the sparse data. The use of the path integral to compare different models is akin to comparing long-time correlations. Complex boundary conditions can be cleanly incorporated into this representation, using a variant of "boundary element" techniques.

The histogram procedure recognizes that the distribution can be numerically approximated to a high degree of accuracy as sum of rectangles at points M_i of height P_i and width ΔM_i. For convenience, just consider a one-dimensional system. The above path-integral representation can be rewritten, for each of its intermediate integrals, as

$$P(M; t + \Delta t) = \int dM' [g_s^{1/2} (2\pi \Delta t)^{-1/2} \exp(-\underline{L}_s \Delta t)] P(M'; t)$$

$$= \int dM' G(M, M'; \Delta t) P(M'; t) \ ,$$

$$P(M; t) = \sum_{i=1}^{N} \pi(M - M_i) P_i(t) \ ,$$

$$\pi(M - M_i) = \begin{cases} 1 \ , (M_i - \frac{1}{2} \Delta M_{i-1}) \le M \le (M_i + \frac{1}{2} \Delta M_i) \ , \\ 0 \ , \text{otherwise} \ , \end{cases} \tag{55}$$

which yields

$$P_i(t + \Delta t) = T_{ij}(\Delta t) P_j(t) \ ,$$

$$T_{ij}(\Delta t) = \frac{2}{\Delta M_{i-1} + \Delta M_i} \int_{M_i - \Delta M_{i-1}/2}^{M_i + \Delta M_i/2} dM \int_{M_j - \Delta M_{j-1}/2}^{M_j + \Delta M_j/2} dM' G(M, M'; \Delta t) \ . \tag{56}$$

7.2.3. PATHINT Path-Integral Evolution of Fitted EEG

The author has generalized the Wehner-Wolfer algorithm to arbitrary dimensions. The utility of this code extends to many sciences in addition to physics and chemistry, e.g., to simulations analyses (Ingber, Fujio, and Wehner, 1991) and finance (Ingber, 1990; Ingber, 1995b; Ingber, Wehner *et al*, 1991). PATHINT has been used to rigorously evolve STM constraints (Ingber, 1994), previously approximately analysed using asymptotic limits of the SMNI mesocolumnar probability distribution (Ingber, 1984; Ingber, 1985b).

This PATHINT code is an important partner to the ASA code previously developed. ASA has made it possible to perform fits of complex nonlinear probability distributions to EEG data. Now, using ASA, the parameters of the fitted SMNI distribution can be used to determine a distribution of firings in a short initial time epoch of EEG. Then, PATHINT can be used to predict the evolution of the system, possibly to predict oncoming states, e.g., epileptic seizures of patients baselined to a fitted distribution.

As an example of a recent application of PATHINT to STM, Fig. 13 is the evolution of model BC′ at τ, after 100 foldings of the path integral (Ingber and Nunez, 1995).

Figure 13

In agreement with previous studies (Ingber, 1984; Ingber, 1985b), models BC′ and BC′_VIS support multiple stable states in the interior physical firing M^G-space for time scales of a few tenths of a second. Models EC′ and IC′ do not possess these attributes.

Fig. 14 examines the interior of M^G-space of model BC′_VIS at time τ (Ingber and Nunez, 1995).

Figure 14

7.3. Invariants

7.3.1. Induced Riemannian Geometry

A Riemannian geometry is derived as a consequence of nonlinear noise, reflecting that the probability distribution is invariant under general nonlinear transformations of these variables (Langouche *et al*, 1982).

This becomes explicit under a transformation to the midpoint discretization, in which the standard rules of differential calculus hold for the *same* distribution:

$$M^G(\bar{t}_s) = \frac{1}{2}(M^G_{s+1} + M^G_s) \,, \; \dot{M}^G(\bar{t}_s) = (M^G_{s+1} - M^G_s)/\theta \,,$$

$$\tilde{P} = \prod_{\nu} P \,, \; P = \int \cdots \int \underline{D}M \exp(-\sum_{s=0}^{u} \Delta t L_{Fs}) \,,$$

$$\underline{D}M = g_{0_+}^{1/2}(2\pi\Delta t)^{-1/2} \prod_{s=1}^{u} g_{s_+}^{1/2} \prod_{G=1}^{\Theta} (2\pi\Delta t)^{-1/2} dM^G_s \,,$$

$$\int dM^G_s \to \sum_{\alpha=1}^{N^G} \Delta M^G_{\alpha s} \,, \; M^G_0 = M^G_{t_0} \,, \; M^G_{u+1} = M^G_t \,,$$

$$L_F = \frac{1}{2}(\dot{M}^G - h^G) g_{GG'} (\dot{M}^{G'} - h^{G'}) + \frac{1}{2} h^G{}_{;G} + R/6 - V \,,$$

$$[\cdots]_{,G} = \frac{\partial[\cdots]}{\partial M^G} \,,$$

$$h^G = g^G - \frac{1}{2} g^{-1/2}(g^{1/2}g^{GG'})_{,G'} ,$$

$$g_{GG'} = (g^{GG'})^{-1} ,$$

$$g_s[M^G(\bar{t}_s), \bar{t}_s] = \det(g_{GG'})_s , g_{s_+} = g_s[M^G_{s+1}, \bar{t}_s] ,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF} h^G = g^{-1/2}(g^{1/2}h^G)_{,G} ,$$

$$\Gamma^F_{JK} \equiv g^{LF}[JK, L] = g^{LF}(g_{JL,K} + g_{KL,J} - g_{JK,L}) ,$$

$$R = g^{JL}R_{JL} = g^{JL}g^{JK}R_{FJKL} ,$$

$$R_{FJKL} = \frac{1}{2}(g_{FK,JL} - g_{JK,FL} - g_{FL,JK} + g_{JL,FK}) + g_{MN}(\Gamma^M_{FK}\Gamma^N_{JL} - \Gamma^M_{FL}\Gamma^N_{JK}) . \tag{57}$$

This geometry presents us with possible invariants of the system. This can function like a “fuzzy” pattern matcher, to extract full invariant patterns when only a smaller signal can be extracted from noise.

7.3.2. Calculation of Information

Information is well defined in terms of a path integral:

$$I = \int \underline{D}\tilde{M}\, \tilde{P} \ln(\tilde{P}/\bar{P}) \tag{58}$$

with respect to a reference distribution $\bar{P}$ (Graham, 1978; Haken, 1983). Like the probability distribution $\tilde{P}$ which defines it, information also is an invariant under general nonlinear transformations.

Some investigators have tried to directly fit statistical white-noise models of mutual information under different electrodes (Gersch, 1987; Mars and Lopes da Silva, 1987). SMNI first fits observables (electric potentials or current flows) to probability distributions, describing inter-electrode interactions via long-ranged fibers and intra-electrode interactions via short-ranged fibers, before calculating the information.

For example, sensory cortex may transmit information to motor cortex, although they have somewhat different neuronal structures or neuronal languages. This information flow can be relatively independent of information flows that take place at finer resolutions, e.g., across a subset of synaptic gaps or individual neurons.

7.4. Mesoscopic Neural Networks (MNN)

7.4.1. Generic Mesoscopic Neural Networks

The SMNI methodology also defines an algorithm to construct a mesoscopic neural network (MNN), based on realistic neocortical processes and parameters, to record patterns of brain activity and to compute the evolution of this system (Ingber, 1992). MNN makes it possible to add a finer minicolumnar scale to the explicit SMNI development at the mesoscopic and regional scales. MNN permits an overlap in scales being investigated by SMNI and ANN.

Furthermore, this new algorithm is quite generic, and can be used to similarly process information in other systems, especially, but not limited to, those amenable to modeling by mathematical physics techniques alternatively described by path-integral Lagrangians, Fokker-Planck equations, or Langevin rate equations. This methodology is made possible and practical by a confluence of techniques drawn from SMNI itself, modern methods of functional stochastic calculus defining nonlinear Lagrangians (Langouche *et al*, 1982), ASA, and parallel-processing computation.

- SMNI describes reasonable mechanism for information processing in neocortex at columnar scales.
- Modern stochastic calculus permits development of alternative descriptions of path-integral Lagrangians, Fokker-Planck equations, and Langevin rate equations. The induced Riemannian geometry affords invariance of probability distribution under general nonlinear transformations.

- ASA presents a powerful global optimization that has been tested in a variety of problems defined by nonlinear Lagrangians.
- Parallel-processing computations can be applied to ASA as well as to a neural-network architecture, as illustrated by the MNN algorithm (Ingber, 1992).

7.4.2. Further Development of SMNI for MNN

While the development of nearest-neighbor interactions into a potential term V' was useful to explore local EEG dispersion relations, for present purposes this is not necessary. As permitted in the development of SMNI, minicolumnar interactions with firings $M^{\dagger G}$ are simply incorporated into F^G:

$$F^G = \frac{V^G - v^{|G|}_{G'} T^{|G|}_{G'}}{(\pi[(v^{|G|}_{G'})^2 + (\phi^{|G|}_{G'})^2] T^{|G|}_{G'})^{1/2}} \ ,$$

$$T^{|G|}_{G'} = a^{|G|}_{G'} N^{G'} + \frac{1}{2} A^{|G|}_{G'} M^{G'}$$

$$+ a^{\dagger |G|}_{G'} N^{\dagger G'} + \frac{1}{2} A^{\dagger |G|}_{G'} M^{\dagger G'}$$

$$+ a^{\ddagger |G|}_{G'} N^{\ddagger G'} + \frac{1}{2} A^{\ddagger |G|}_{G'} M^{\ddagger G'} \ ,$$

$$a^{\dagger G}_{G'} = \frac{1}{2} A^{\dagger G}_{G'} + B^{\dagger G}_{G'} \ ,$$

$$A^{\ddagger I}_E = A^{\ddagger E}_I = A^{\ddagger I}_I = B^{\ddagger I}_E = B^{\ddagger E}_I = B^{\ddagger I}_I = 0 \ ,$$

$$a^{\ddagger E}_E = \frac{1}{2} A^{\ddagger E}_E + B^{\ddagger E}_E \ . \tag{59}$$

M^G represent firings internal to a given minicolumn. $M^{\dagger G}$ represent firings among minicolumns within a macrocolumn. $M^{\ddagger E}$ represent firings among minicolumns in macrocolumns under different electrodes (only $G = E$ firings exist).

7.4.3. SMNI MNN Representation

SMNI can be further developed using MNN as a set of nodes, each described by a short-time probability distribution interacting with the other nodes, on at least two scales.

If each node represents a macrocolumn, or a "macro-node," a set of 100 such macro-nodes could represent a dipole source at the scale recorded by EEG. Even a 10 macro-node representation would tone down the order of magnitude scaling of macrocolumnar probability distributions to represent dipole sources. This should give a more accurate representation of inter-electrode circuitry, possibly even of intra-electrode coherencies.

If a "meso-node" represents a minicolumn, a set of 1000 such meso-nodes represents a macrocolumn. If at least two foldings of temporal interactions are permitted to develop interactions during time epochs typically measured in EEG recordings, then at least one "hidden" layer of meso-nodes should be included. This would make unnecessary the use of nearest-neighbor interactions to represent local spatial interactions across minicolumns. While the inclusion of this meso-node structure might not seem essential to model EEG, this enhancement of SMNI is at least important to clarify other neocortical mechanisms involved in STM, and this could give a more accurate representation of intra-electrode coherencies.

A circuitry among patches of macrocolumns, "regional-nodes," represents a typical circuit of activity correlated to specific behavioral states as recorded by EEG under specific experimental conditions.

It is clear that the addition of "micro-nodes,", with 100 neuronal micro-nodes representing a minicolumn, would quickly overwhelm computer resources if applications to model macroscopic phenomena such as EEG are required. Statistical mechanics plays an important role in making neocortical modeling not only accessible to numerical computation, but also to human conceptual understanding of processes that operate at different scales, a worthy goal of any modeling effort that elevates science from empiricism (Jammer, 1974). The human brain, as well as other complex systems, develops mechanisms at mesoscopic scales to more efficiently process information.

7.4.4. Minicolumnar Interactions

As illustrated in Fig. 15, minicolumnar interactions are represented across three scales: intra-macrocolumnar within a given macrocolumn, intra-regional and inter-macrocolumnar within a given region and between macrocolumns, and inter-regional between regions. The large solid circles represent regions, the intermediate long-dashed circles represent macrocolumns, and the small short-dashed circles represent minicolumns.

Figure 15

7.4.5. Generic MNN

This SMNI MNN can be generalized to model other large-scale nonlinear stochastic multivariate systems, by considering general drifts and diffusions to model such systems, now letting G represent an arbitrary number of variables. Ideally, these systems inherently will be of the Fokker-Planck type,

$$\frac{\partial P}{\partial t} = \frac{\partial(-g^G P)}{\partial M^G} + \frac{1}{2}\frac{\partial^2(g^{GG'}P)}{\partial M^G \partial M^{G'}} . \tag{60}$$

The topology, geometry, and connectivity of the MNN can of course be generalized. There need not be any restriction to nearest-neighbor interactions, although this is simpler to implement especially on parallel processors. Also, "hidden layers" can be included to increase the complexity of the MNN, although the inclusion of nonlinear structure in the drifts g^G and diffusions $g^{GG'}$ may make this unnecessary for many systems.

This addresses some concerns in the neural network community relating to the ability of neural networks to be trusted to generalize to new contexts: If the nodes can be described by mechanisms inherently consistent with system, then more confidence can be justified for generalization. This is more desirable and likely more robust, than using additional "hidden layers" to model such nonlinear structures.

7.4.6. MNN Learning

"Learning" takes place by presenting the MNN with data, and parametrizing the data in terms of the "firings," or multivariate M^G "spins." The "weights," or coefficients of functions of M^G appearing in the drifts and diffusions, are fit to incoming data, considering the joint "effective" Lagrangian (including the logarithm of the prefactor in the probability distribution) as a dynamic cost function.

The cost function is a sum of effective Lagrangians from each node and over each time epoch of data. This program of fitting coefficients in Lagrangian uses methods of adaptive simulated annealing (ASA). This maximum likelihood procedure (statistically) avoids problems of trapping in local minima, as experienced by other types of gradient and regression techniques.

7.4.7. MNN Prediction

"Prediction" takes advantage of a mathematically equivalent representation of the Lagrangian path-integral algorithm, i.e., a set of coupled Langevin rate-equations. The Itô (prepoint-discretized) Langevin equation is analyzed in terms of the Wiener process dW^i, which is rewritten in terms of Gaussian noise $\eta^i = dW^i/dt$ in the limit:

$$M^G(t + \Delta t) - M^G(t) = dM^G = g^G dt + \hat{g}^G_i dW^i ,$$

$$\frac{dM^G}{dt} = \dot{M}^G = g^G + \hat{g}_i^G \eta^i \ ,$$

$$M = \{ M^G ; G = 1, \cdots, \Lambda \} \ \ , \ \ \eta = \{ \eta^i ; i = 1, \cdots, N \} \ \ ,$$

$$< \eta^j(t) >_\eta = 0 \ \ , \ \ < \eta^j(t), \eta^{j'}(t') >_\eta = \delta^{jj'} \delta(t - t') \ . \tag{61}$$

Moments of an arbitrary function $F(\eta)$ over this stochastic space are defined by a path integral over η^i. The Lagrangian diffusions are calculated as

$$g^{GG'} = \sum_{i=1}^{N} \hat{g}_i^G \hat{g}_i^{G'} \ . \tag{62}$$

The calculation of the evolution of Langevin systems has been implemented in the above-mentioned systems using ASA. It has been used as an aid to debug the ASA fitting codes, by first generating data from coupled Langevin equations, relaxing the coefficients, and then fitting this data with the effective Lagrangian cost-function algorithm to recapture the original coefficients within the diffusions defined by $g^{GG'}$.

7.4.8. MNN Parallel Processing

The use of parallel processors can make this algorithm even more efficient, as ASA lends itself well to parallelization.

- During "learning," blocks of random numbers are generated in parallel, and then sequentially checked to find a generating point satisfying all boundary conditions.
- Advantage is taken of the low ratio of acceptance to generated points typical in ASA, to generate blocks of cost functions, and then sequentially checked to find the next best current minimum.
- Additionally, when fitting dynamic systems, e.g., the three physical systems examined to date, parallelization is attained by independently calculating each time epoch's contribution to the cost function.
- Similarly, during "prediction," blocks of random numbers are generated to support the Langevin-equation calculations, and each node is processed in parallel.

7.5. Nonlinear Dynamics — Chaos?

What if EEG has chaotic mechanisms that overshadow the above stochastic considerations? The real issue is whether the scatter in data can be distinguished between being due to noise or chaos (Pool, 1989). In this regard, we note that several studies have been proposed with regard to comparing chaos to simple filtered (colored) noise (Grassberger, 1986; Pool, 1989). Since we have previously derived the existence of multiplicative noise in neocortical interactions, then the previous references must be generalized, such that we must investigate whether EEG scatter can be distinguished from multiplicative noise.

Similar to serious work undertaken in several fields (Brock, 1986; Grassberger, 1986), the impulse to identify ''chaos'' in a complex system often has been premature. It is not supported by the facts, tentative as they are because of sparse data. Similar caution should be exercised regarding neocortical interactions.

A more purposeful project is to compare stochastic with deterministic models of data. Today much attention is turning to the use of deterministic chaotic models for short-time predictions of systems. For example, if only short-time predictions are required, and if a deterministic chaotic model could well describe stochastic data within these epochs, then this model might be more computationally efficient instead of a more ''correct'' stochastic model, which would be necessary for long-time predictions. The scales of time involved are of course system dependent, and the deterministic chaotic modeling of data is still in its infancy (Abarbanel *et al*, 1990).

Similarly, the above SMNI-derived distributions can be used to help determine if chaos is a viable mechanism in EEG. The probability distribution itself is a mathematical measure to which tests can be applied to determine the existence of other nonlinear mechanisms. The path integral has been used to compare long-time correlations in data to predictions of models, while calculating their sensitivity, e.g., of second moments, to initial conditions (Ingber, Fujio, and Wehner, 1991). This also helps to compare alternative models, previously having their short-time probability distributions fit to data, with respect to their predictive power over long time scales.

While these studies are concerned with neocortical interactions, it is of interest to note a series of experimental and theoretical studies of nonlinear dynamics of the olfactory bulb in small mammals, in which distinctive EEG patterns on the bulb surface are shown to be associated with specific odors (Baird, 1986; Freeman, 1987; Freeman and Skarda, 1985). These studies demonstrating chaos are very much model dependent, and as such it is only fair to present the models as possessing chaos, not necessarily the actual physical system.

For example, it has been widely noted that the correlation dimension of data is difficult to calculate; perhaps it is often not even a well-founded concept, e.g., since the EEG of event-related potentials is likely nonstationary and very much context and subject dependent (Rapp *et al*, 1989). Its calculation, e.g., using the popular Grassberger-Procaccia algorithm (Grassberger and Procaccia, 1983), even when supplemented with finer statistical tests (Brock *et al*, 1988) and noise reduction techniques (Broomhead and King, 1986), may prove fruitful, but likely only as a sensitivity index relative to shifting contexts and complementary to other models of EEG data.

8. Conclusion

8.1. Scaling Paradigm

These results, in addition to their importance in reasonably modeling EEG with SMNI, also have a deeper theoretical importance with respect to the scaling of neocortical mechanisms of interaction across disparate spatial scales and behavioral phenomena.

Summarizing some of the above equations, the derivation of the mesoscopic probability distribution, yields similar algebraic structures of the threshold factors at the neuronal scale,

$$F_j = \frac{V_j - \sum_k a_{jk} v_{jk}}{(\pi \sum_{k'} a_{jk'} (v_{jk'}^2 + \phi_{jk'}^2))^{\frac{1}{2}}} , \tag{63}$$

to be compared with the mesoscopic threshold factor,

$$F^G = \frac{(V^G - a_{G'}^{|G|} v_{G'}^{|G|} N^{G'} - \frac{1}{2} A_{G'}^{|G|} v_{G'}^{|G|} M^{G'})}{(\pi [(v_{G'}^{|G|})^2 + (\phi_{G'}^{|G|})^2](a_{G'}^{|G|} N^{G'} + \frac{1}{2} A_{G'}^{|G|} M^{G'}))^{1/2}} , \tag{64}$$

which contribute to the drifts g^G and diffusions g^{GG},

$$g^G = -\tau^{-1}(M^G + N^G \tanh F^G) ,$$

$$g^{GG'} = (g_{GG'})^{-1} = \delta_G^{G'} \tau^{-1} N^G \mathrm{sech}^2 F^G . \tag{65}$$

The common algebraic factors in the F_j and F^G factors illustrate the similar mechanisms of threshold activity at their appropriate scales of neurons and columns of neurons, respectively.

The macroscopic regional probability distribution has the same functional form as the mesoscopic distribution, where the macroscopic drifts and diffusions of the potentials described by the Φ's are simply linearly related to the (nonlinear) mesoscopic drifts and diffusions of the columnar firing states given by the M^G's.

$$m \equiv < \Phi_\nu - \phi > = a < M^E > + b < M^I > = a g^E + b g^I ,$$

$$\sigma^2 \equiv <(\Phi_\nu - \phi)^2> - <\Phi_\nu - \phi>^2 = \mathrm{a}^2 \mathrm{g}^{\mathrm{EE}} + \mathrm{b}^2 \mathrm{g}^{\mathrm{II}} \ . \tag{66}$$

It was possible to scale the macrocolumnar distribution to describe electrode recording areas of several cm^2, and have the fitted parameters lie within their experimentally observed ranges.

8.2. Addressing Phenomena at Multiple Scales

Analyzing, understanding, and attempting to predict neocortical phenomena at the spatial scales encountered in EEG scalp recordings involves measuring firings from millions of neurons; in intracortical measurements at least tens of thousands of neurons are involved. SMNI approaches such scales as new systems with emergent phenomena. Just as the physics and chemistry of many large-scale systems, such as encountered in fluid and plasma states, could not have progressed this century if all theoretical and experimental research were inappropriately constrained to be understood at the level of quantum mechanics (or quarks or strings), so neuroscience must soon accept that all brain phenomena is not best understood or perhaps understood at all at the level of simple neuron-neuron interactions (or membrane dynamics or macromolecular interactions). Different scales present new complex systems that must be approached as such and often with different approaches.

8.3. SMNI Features for Further Application and Testing

The following features of SMNI are also a set of guidelines to implement and test in future studies of EEG data.

Logical and Testable Development Across Multiple Scales

In the course of a logical, nonlinear, stochastic development of aggregating neuronal and synaptic interactions to larger and larger scales, opportunities are taken to use techniques of mathematical physics to overcome several technical hurdles. Paradigms and metaphors from other disciplines do not substitute for logical SMNI development.

Validity Across Multiple Scales

The SMNI theoretical model has independent validity in describing EEG dispersion relations, systematics of short-term memory, velocities of propagation of information across neocortical fields, recency versus primacy effects, etc. Fits of such models to data should do better in extracting signal from noise than *ad hoc* phenomenological models.

Use of ASA and PATHINT on Nonlinear Stochastic Systems

ASA enables the fitting of quite arbitrary nonlinear stochastic models to such data as presented by EEG systems. This means that functional dependences in the noise itself (the diffusion matrix) as well as the functional dependences in the driving terms (the drift vector) can be fit directly. Once fitted, PATHINT can evolve the system, testing long-time correlations between the model(s) and the data, as well as serving to predict events.

Inclusion of Short-Range and Long-Range Interactions

SMNI proposes that models to be fitted to data include models of activity under each electrode, e.g., due to short-ranged neuronal fibers, as well as models of activity across electrodes, e.g., due to long-ranged fibers. These influences can be disentangled by SMNI fits.

Riemannian Invariants

Yet to explore are the ramifications of using the derived (not hypothesized) Riemannian metric induced by multivariate Fokker-Plank-type systems. This seems to form a natural invariant measure of information, that could/should be used to explore flows of information between neocortical regions.

Renormalization of Attenuated Frequencies

The SMNI approach shows how to "renormalize" the spatial activity to get a model that more closely matches the experimental situation, wherein there is attenuation of ranges of wave numbers. This program must be carried out.

MNN Real-Time Processing and Audit Trail to Finer Scales

The MNN parallel algorithm may offer real-time processing of nonlinear modeling and fitting of EEG data for clinical use. Regional EEG data can be interpreted as mechanisms occurring at the minicolumnar scales. This algorithm will be implemented in the future.

8.4. Utility, Parsimony, & Beauty

SMNI presents a computational algorithm faithful to a model of neocortical interactions that has been baselined to experimental observations across multiple scales. This project is constructing a tool to be used for real-time study and diagnoses of brain function.

Similar to the neocortex, in many complex systems, as spatial-temporal scales of observation are increased, new phenomena arise by virtue of synergistic interactions among smaller-scale entities—perhaps more properly labeled ‘‘quasientities’’—which serve to explain much observed data in a parsimonious, usually mathematically aesthetic, fashion (Gardiner, 1983; Haken, 1983; Kubo *et al*, 1973; Nicolis and Prigogine, 1973; van Kampen, 1981). Many complex systems are in nonequilibrium, being driven by nonlinear and stochastic interactions of many external and internal degrees of freedom. For these systems, classical thermodynamical approaches typically do not apply (Ma, 1985). Such systems are best treated by respecting some intermediate mesoscale as “fundamental” to drive larger macroscopic processes.

The use of ASA and PATHINT permit the processing of quite general nonlinear, stochastic, and multivariate descriptions of systems, without being limited to equilibrium energy-type functions. Phenomena specific to given scales of interaction must be modeled and studied, as well as the flow of information across these scales. The stochastic nonlinear SMNI paradigm permits this development.

References

Abarbanel, H.D.I., Brown, R., and Kadtke, J.B. Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra. Phys. Rev. A. 1990, 41: 1782-1807.

Agarwal, G.S. and Shenoy, S.R. Observability of hysteresis in first-order equilibrium and nonequilibrium phase transitions. Phys. Rev. A. 1981, 23: 2719-2723.

Baird, B. Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb. Physica D. 1986, 22: 150-175.

Braitenberg, V. Cortical Architectonics: General and Areal. Raven. New York. 1978.

Brock, W.A. Distinguishing random and deterministic systems: Abridged version. J. Econ. Theory. 1986, 40: 168-195.

Brock, W.A., Dechert, W.D., Scheinkman, J.A., and LeBaron, B. A test for the independence based on the correlation dimension. SSRI Working Paper No. 8702, Revised. U. Wisconsin. Madison, WI. 1988.

Broomhead, D.S. and King, G.P. Extracting qualitative dynamics from experimental data. Physica D. 1986, 20: 217-236.

Caillé, A., Pink, D., de Verteuil, F., and Zuckermann, M.J. Theoretical models for quasi-two-dimensional mesomorphic monolayers and membrane bilayers. Can. J. Phys.. 1980, 58: 1723-1726.

Cohen, D., Cuffin, B.N., Yunokuchi, K., Maniewski, R., Purcell, C., Cosgrove, G.R., Ives, J., Kennedy, J., and Schomer, D. MEG versus EEG localization test using implanted sources in the human brain. Ann. Neurol.. 1990, 28: 811-817.

Cowan, J.D. Spontaneous symmetry breaking in large scale nervous activity. Int. J. Quant. Chem.. 1982, 22: 1059-1082.

Dekker, H. Quantization in curved spaces. In: J.P. Antoine and E. Tirapegui (Ed.). Functional Integration: Theory and Applications. Plenum. New York. 1980, 207-224.

Eichenbaum, H. Thinking about brain cell assemblies. Science. 1993, 261: 993-994.

Ericsson, K.A. and Chase, W.G. Exceptional memory. Am. Sci.. 1982, 70: 607-615.

Feynman, R.P., Leighton, R.B., and Sands, M. The Feynman Lectures on Physics. Addison Wesley. Reading, MA. 1963.

Fitzpatrick, K.A. and Imig, T.J. Auditory cortico-cortical connections in the owl monkey. J. Comp. Neurol.. 1980, 192: 589-601.

Freeman, W.J. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. Biol. Cybern.. 1987, 55: 139-150.

Freeman, W.J. and Skarda, C.A. Spatial EEG patterns, nonlinear dynamics and perception: The neo-Sherringtonian view. Brain Res. Rev.. 1985, 10: 147-175.

Gardiner, C.W. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Springer-Verlag. Berlin, Germany. 1983.

Gersch, W. Non-stationary multichannel time series analysis. In: A.S. Gevins and A. Remond (Ed.). Methods of Brain Electrical and Magnetic Signals. EEG Handbook. Elsevier. New York, NY. 1987, 261-296.

Goldman, P.S. and Nauta, W.J.H. Columnar distribution of cortico-cortical fibers in the frontal association, limbic, and motor cortex of the developing rhesus monkey. Brain Res.. 1977, 122: 393-413.

Grabert, H. and Green, M.S. Fluctuations and nonlinear irreversible processes. Phys. Rev. A. 1979, 19: 1747-1756.

Graham, R. Covariant formulation of non-equilibrium statistical thermodynamics. Z. Physik. 1977a, B26: 397-405.

Graham, R. Lagrangian for diffusion in curved phase space. Phys. Rev. Lett.. 1977b, 38: 51-53.

Graham, R. Path-integral methods on nonequilibrium thermodynamics and statistics. In: L. Garrido, P. Seglar and P.J. Shepherd (Ed.). Stochastic Processes in Nonequilibrium Systems. Springer. New York, NY. 1978, 82-138.

Grassberger, P. Do climatic attractors exist?. Nature. 1986, 323: 609-612.

Grassberger, P. and Procaccia, I. Measuring the strangeness of strange attractors. Physica D. 1983, 9: 189-208.

Haken, H. Synergetics. Springer. New York. 1983.

Hertz, J., Krogh, A., and Palmer, R.G. Introduction to the Theory of Neural Computation. Addison-Wesley. Redwood City, CA. 1991.

Hubel, D.H. and Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol.. 1962, 160: 106-154.

Hubel, D.H. and Wiesel, T.N. Functional architecture of Macaque monkey visual cortex. Proc. R. Soc. London Ser. B. 1977, 198: 1-59.

Imig, T.J. and Reale, R.A. Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex. J. Comp. Neurol.. 1980, 192: 293-332.

Ingber, L. Towards a unified brain theory. J. Social Biol. Struct.. 1981, 4: 211-224.

Ingber, L. Statistical mechanics of neocortical interactions. I. Basic formulation. Physica D. 1982, 5: 83-107.

Ingber, L. Statistical mechanics of neocortical interactions. Dynamics of synaptic modification. Phys. Rev. A. 1983, 28: 395-416.

Ingber, L. Statistical mechanics of neocortical interactions. Derivation of short-term-memory capacity. Phys. Rev. A. 1984, 29: 3346-3358.

Ingber, L. Statistical mechanics of neocortical interactions. EEG dispersion relations. IEEE Trans. Biomed. Eng.. 1985a, 32: 91-94.

Ingber, L. Statistical mechanics of neocortical interactions: Stability and duration of the 7±2 rule of short-term-memory capacity. Phys. Rev. A. 1985b, 31: 1183-1186.

Ingber, L. Very fast simulated re-annealing. Mathl. Comput. Modelling. 1989, 12: 967-973.

Ingber, L. Statistical mechanical aids to calculating term structure models. Phys. Rev. A. 1990, 42: 7057-7064.

Ingber, L. Statistical mechanics of neocortical interactions: A scaling paradigm applied to electroencephalography. Phys. Rev. A. 1991, 44: 4017-4060.

Ingber, L. Generic mesoscopic neural networks based on statistical mechanics of neocortical interactions. Phys. Rev. A. 1992, 45: R2183-R2186.

Ingber, L. Statistical mechanics of combat and extensions. In: C. Jones (Ed.). Toward a Science of Command, Control, and Communications. American Institute of Aeronautics and Astronautics. Washington, D.C.. 1993a, 117-149.

Ingber, L. Simulated annealing: Practice versus theory. Mathl. Comput. Modelling. 1993b, 18: 29-57.

Ingber, L. Adaptive Simulated Annealing (ASA). [ftp.alumni.caltech.edu: /pub/ingber/ASA-shar, ASA-shar.Z, ASA.tar.Z, ASA.tar.gz, ASA.zip]. Lester Ingber Research. McLean, VA. 1993c.

Ingber, L. Statistical mechanics of neocortical interactions: Path-integral evolution of short-term memory. Phys. Rev. E. 1994, 49: 4652-4664.

Ingber, L. Statistical mechanics of multiple scales of neocortical interactions. In: P.L. Nunez (Ed.). Neocortical Dynamics and Human EEG Rhythms. Oxford University Press. New York, NY. 1995a, 628-681.

Ingber, L. Path-integral evolution of multivariate systems with moderate noise. Phys. Rev. E. 1995b, 51: 1616-1619.

Ingber, L., Fujio, H., and Wehner, M.F. Mathematical comparison of combat computer models to exercise data. Mathl. Comput. Modelling. 1991, 15: 65-90.

Ingber, L. and Nunez, P.L. Multiple scales of statistical physics of neocortex: Application to electroencephalography. Mathl. Comput. Modelling. 1990, 13: 83-95.

Ingber, L. and Nunez, P.L. Statistical mechanics of neocortical interactions: High resolution path-integral calculation of short-term memory. Phys. Rev. E. 1995, 51: 5074-5083.

Ingber, L. and Sworder, D.D. Statistical mechanics of combat with human factors. Mathl. Comput. Modelling. 1991, 15: 99-127.

Ingber, L., Wehner, M.F., Jabbour, G.M., and Barnhill, T.M. Application of statistical mechanics methodology to term-structure bond-pricing models. Mathl. Comput. Modelling. 1991, 15: 77-98.

Jagadeesh, B., Wheat, H.S., and Ferster, D. Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. Science. 1993, 262: 1901-1904.

Jammer, M. The Philosophy of Quantum Mechanics. Wiley & Sons. New York, NY. 1974.

Jones, E.G., Coulter, J.D., and Hendry, S.H.C. Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys. J. Comp. Neurol.. 1978, 181: 291-348.

Korn, H., Mallet, A., and Faber, D.S. Fluctuating responses at a central synapse: n of binomial fit predicts number of stained presynaptic boutons. Science. 1981, 213: 898-900.

Kubo, R., Matsuo, K., and Kitahara, K. Fluctuation and relaxation of macrovariables. J. Stat. Phys.. 1973, 9: 51-96.

Langouche, F., Roekaerts, D., and Tirapegui, E. Functional Integration and Semiclassical Expansions. Reidel. Dordrecht, The Netherlands. 1982.

Lopes da Silva, F.H. Neural mechanisms underlying brain waves: From neural membranes to networks. Electroencephal. clin. Neurophysiol.. 1991, 79: 81-93.

Ma, S.-K. Statistical Mechanics. World Scientific. Philadelphia. 1985.

Mars, N.J.I. and Lopes da Silva, F.H. EEG analysis methods based on information theory. In: A.S. Gevins and A. Remond (Ed.). Methods of Brain Electrical and Magnetic Signals. EEG Handbook. Elsevier. New York, NY. 1987, 297-307.

Miller, G.A. The magical number seven, plus or minus two. Psychol. Rev.. 1956, 63: 81-97.

Mountcastle, V.B. An organizing principle for cerebral function: The unit module and the distributed system. In: G.M. Edelman and V.B. Mountcastle (Ed.). The Mindful Brain. Massachusetts Institute of Technology. Cambridge. 1978, 7-50.

Mountcastle, V.B., Andersen, R.A., and Motter, B.C. The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex. J. Neurosci.. 1981, 1: 1218-1235.

Murdock, B.B., Jr. A distributed memory model for serial-order information. Psychol. Rev.. 1983, 90: 316-338.

Nicolis, G. and Prigogine, I. Self-Organization in Nonequilibrium Systems. Wiley. New York, NY. 1973.

Nunez, P.L. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press. London. 1981.

Nunez, P.L. Localization of brain activity with Electroencephalography. In: S. Sato (Ed.). Advances in Neurology, Vol. 54: Magnetoencephalography. Raven Press. New York, NY. 1990, 39-65.

Nunez, P.L. and Srinivasan, R. Implications of recording strategy for estimates of neocortical dynamics with electroencephalography. Chaos. 1993, 3: 257-266.

Perkel, D.H. and Feldman, M.W. Neurotransmitter release statistics: Moment estimates for inhomogeneous Bernoulli trials. J. Math. Biol.. 1979, 7: 31-40.

Pool, R. Is it chaos, or is it just noise?. Science. 1989, 243: 25-28.

Porjesz, B. and Begleiter, H. Event-related potentials individuals at risk for alcoholism. Alcohol. 1990, 7: 465-469.

Rapp, P.E., Bashore, T.R., Marinerie, J.M., Albano, A.M., Zimmerman, I.D., and Mees, A.I. Dynamics of brain electrical activity. Brain Topography. 1989, 2: 99-118.

Schulman, L.S. Techniques and Applications of Path Integration. J. Wiley & Sons. New York. 1981.

Scott, A.C. The electrophysics of a nerve fiber. Rev. Mod. Phys.. 1975, 47: 487-533.

Shaw, G.L. and Vasudevan, R. Persistent states of neural networks and the random nature of synaptic transmission. Math. Biosci.. 1974, 21: 207-218.

Shepherd, G.M. The Synaptic Organization of the Brain. Oxford University Press. New York, NY. 1979.

Sommerhoff, G. Logic of the Living Brain. Wiley. New York, NY. 1974.

Tsal, Y. Movements of attention across the visual field. J. Exp. Psychol.. 1983, 9: 523-530.

van Kampen, N.G. Stochastic Processes in Physics and Chemistry. North-Holland. Amsterdam. 1981.

von der Heydt, I., von der Heydt, N., and Obermair, G.M. Statistical model of current-coupled ion channels in nerve membranes. Z. Physik. 1981, 41: 153-164.

Vu, E.T. and Krasne, F.B. Evidence for a computational distinction between proximal and distal neuronal inhibition. Science. 1992, 255: 1710-1712.

Wehner, M.F. and Wolfer, W.G. Numerical evaluation of path-integral solutions to Fokker-Planck equations. I.. Phys. Rev. A. 1983a, 27: 2663-2670.

Wehner, M.F. and Wolfer, W.G. Numerical evaluation of path-integral solutions to Fokker-Planck equations. II. Restricted stochastic processes. Phys. Rev. A. 1983b, 28: 3003-3011.

Wehner, M.F. and Wolfer, W.G. Numerical evaluation of path integral solutions to Fokker-Planck equations. III. Time and functionally dependent coefficients. Phys. Rev. A. 1987, 35: 1795-1801.

Zhang, G. and Simon, H.A. STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. Memory & Cognition. 1985, 13: 193-201.

Table Captions

TABLE I.
Comparison of ANN and SMNI approaches to EEG scales.

TABLE II.
SMNI fits to EEG data of all 49 subjects in the experiment described in the text, expressed in terms of intra-electrode firing states and inter-electrode connectivity. EEG data courtesy of Henri Begleiter.

Figure Captions

Fig. 1. Standard EEG electrode map.

Fig. 2. Sample EEG power spectra. The upper left figure is mu rhythm at 8.0 Hz. The lower right figure is alpha rhythm at 9.5 Hz. Figure courtesy of Paul Nunez.

Fig. 3. Illustration of scales underlying dipole approximation. Figure courtesy of Paul Nunez.

Fig. 4. Mechanical string analogy to spatial-temporal EEG.

Fig. 5. Multiple scales of neocortical interactions.

Fig. 6. Representation of nearest-neighbor mesocolumnar interactions.

Fig. 7. Example of minima structure of Lagrangian.

Fig. 8. Contours of inhibitory-state Lagrangian.

Fig. 9. Contours of excitatory-state Lagrangian.

Fig. 10. Contours of balanced-state Lagrangian.

Fig. 11. Contours of balanced-state Lagrangian for visual neocortex.

Fig. 12. Contours of nearest-neighbor contributions to Lagrangian.

Fig. 13. Path-integral evolution of minicolumnar Lagrangian at 50τ, approximately 0.5 sec for model BC′.

Fig. 14. Path-integral evolution of minicolumnar Lagrangian at 50τ, approximately 0.5 sec, for model BC′_VIS, using denser minicolumns appropriate to visual neocortex. Probability densities are cut off for values higher than 0.0001 to expose the clustering effect.

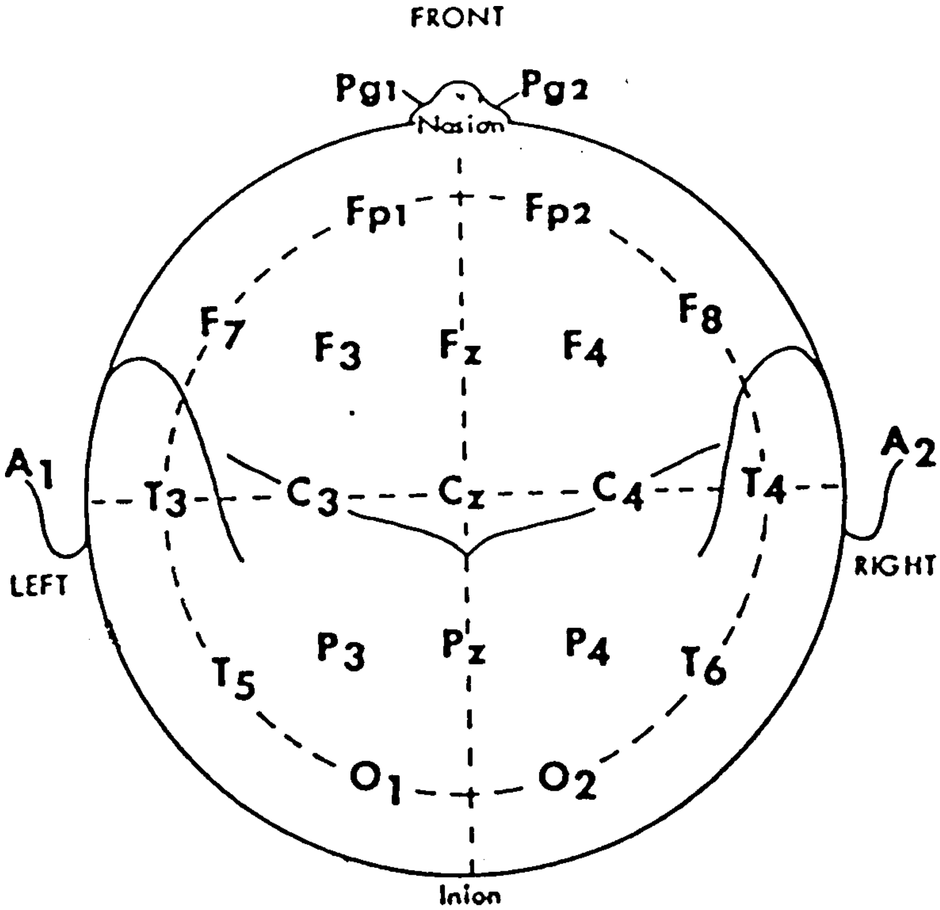
Fig. 15. Illustration of interactions between regions, macrocolumns and minicolumns.

Feature	SMNI	ANN
spatial unit	minicolumn	neuron
spatial scale	$10^5 - 10^6$ neurons	$10^2 - 10^3$ neurons
large sets of neurons	*pre*-analysis	*post*-analysis
neuronal features	statistical constraints	neuronal mechanisms
columnar features	columnar interactions	columnar formation
scalp EEG applications	yes	probably not
intra-cortical EEG	yes	yes
spike EEG	no	yes

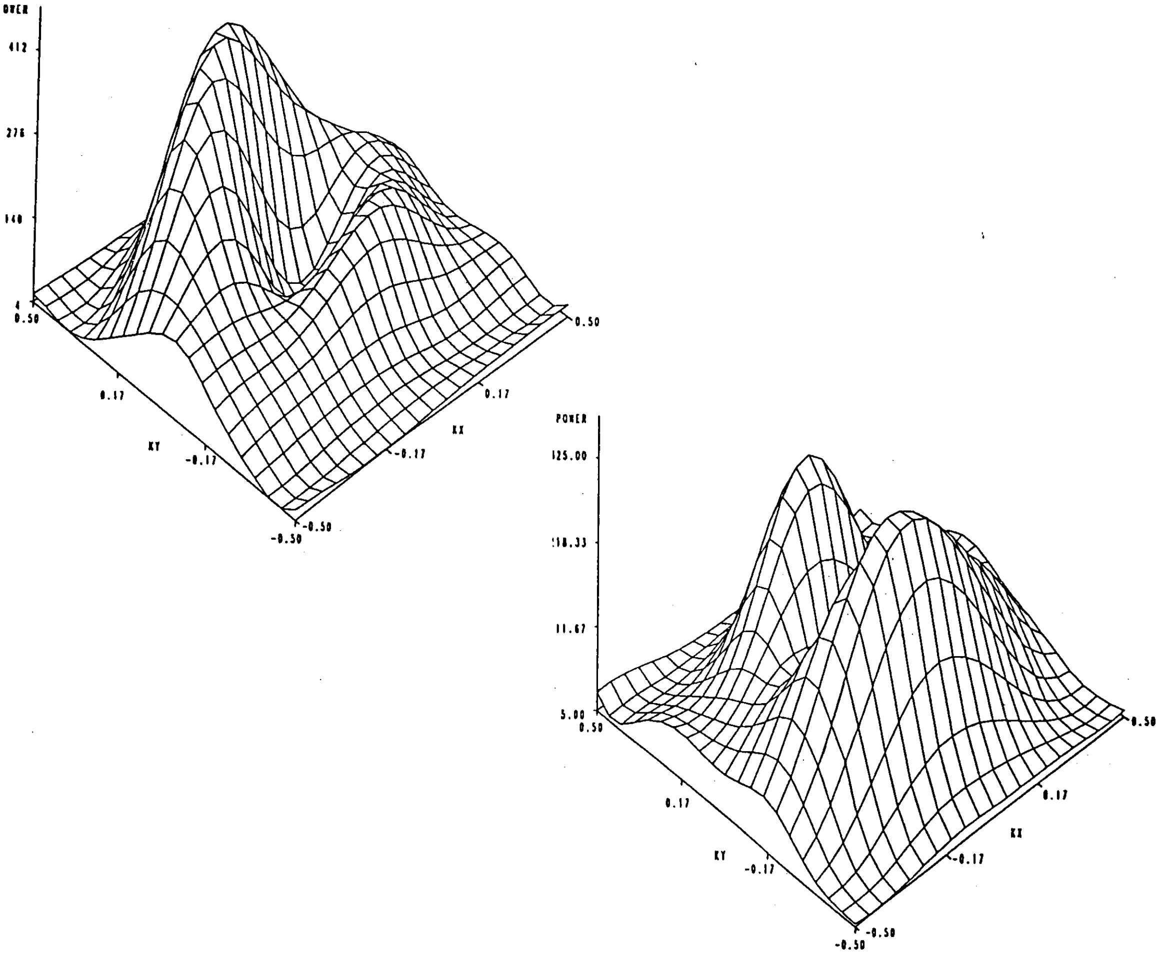
Electrode	Mean M^E	S.D. M^E	Mean $\sigma(M^{E)}$	S.D. $\sigma(M^E)$
Fz	1.13661	3.8754	6.09339	5.30891
Cz	-0.533493	4.83208	6.31146	5.59003
Pz	-0.3158	3.87471	5.44242	5.50453
P3	-0.121703	10.1069	8.152	7.08701
P4	-0.0208276	7.47837	11.0526	7.04522

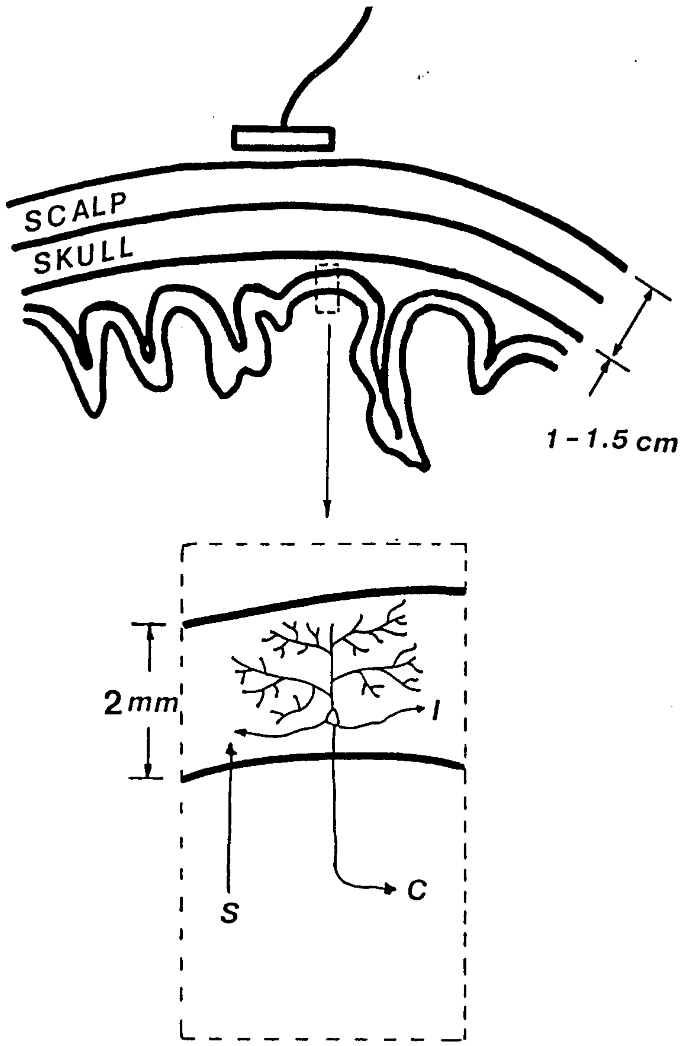
Electrode	Mean M^I	S.D. M^I	Mean $\sigma(M^{I)}$	S.D. $\sigma(M^I)$
Fz	1.83249	7.11368	11.4131	2.71679
Cz	0.229446	5.89307	11.5578	2.68969
Pz	-0.255393	6.37452	12.4699	2.86198
P3	-0.0234756	7.39736	10.5579	3.2693
P4	-0.0271411	6.25121	12.0525	2.52846

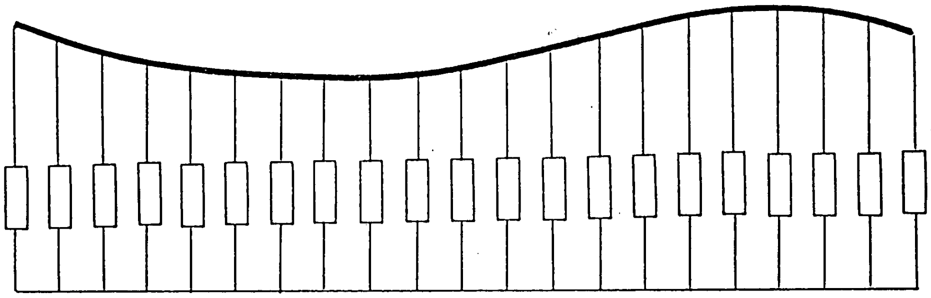
Electrode	Mean d	S.D. d
Fz → Cz	0.389722	0.291677
Cz → Pz	0.377305	0.266958
Pz → P3	0.536313	0.288519
Pz → P4	0.485525	0.294742

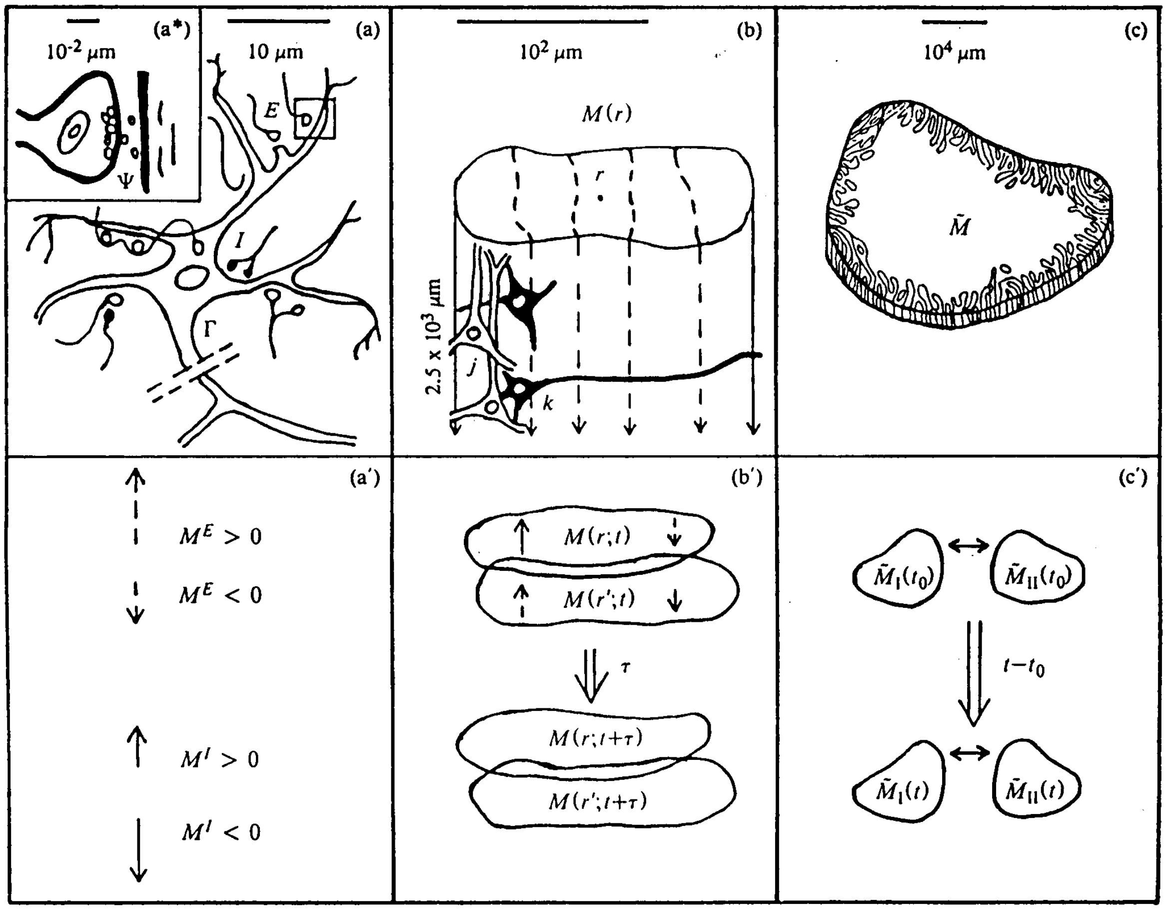

FRONT
Pg1
Pg2
Nasion
Fp1
Fp2
F7
F3
Fz
F4
F8
A1
T3
C3
Cz
C4
T4
A2
LEFT
RIGHT
P3
Pz
P4
T5
T6
O1
O2
Inion

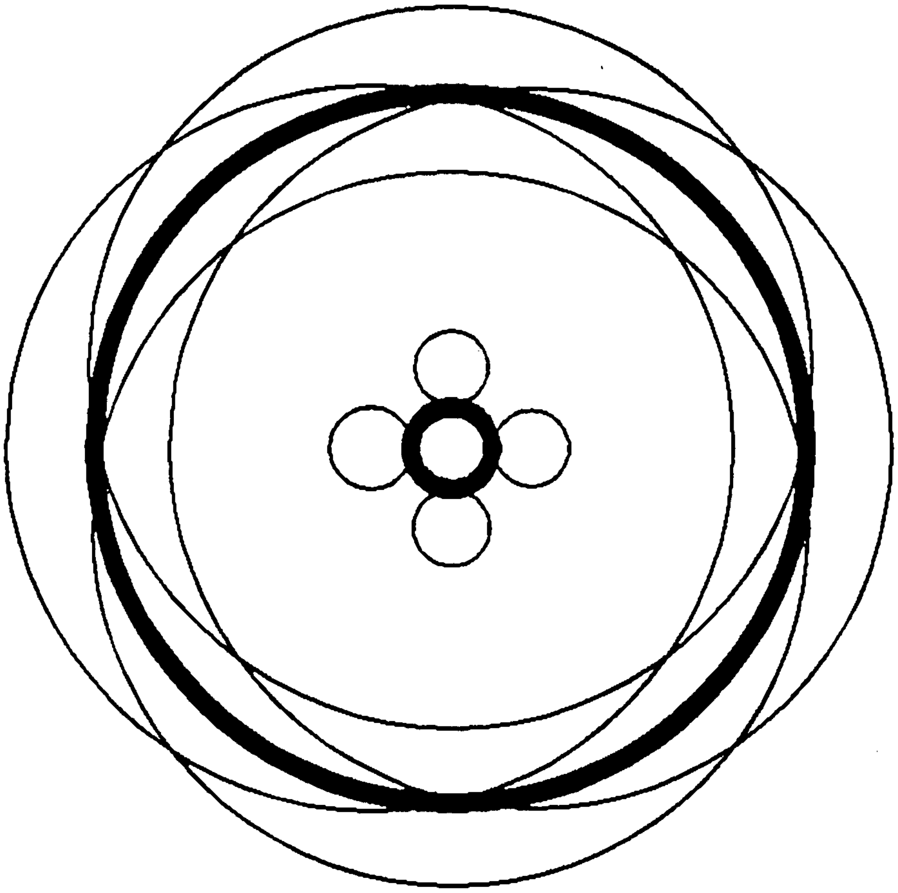
TOP OF HEAD

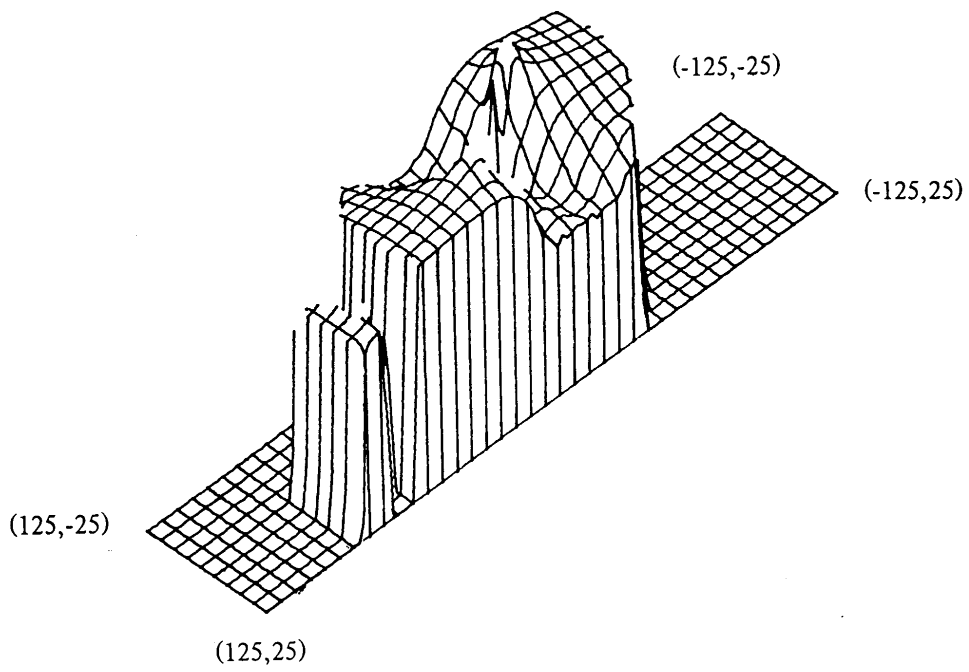

412
276
140
4
0.50
0.17
XY
-0.17
-0.50
-0.50
XX
0.50
0.17
-0.17
POWER
125.00
11.67
5.00
0.50
0.17
XY
-0.17
-0.50
-0.50
XX
0.50
0.17
-0.17

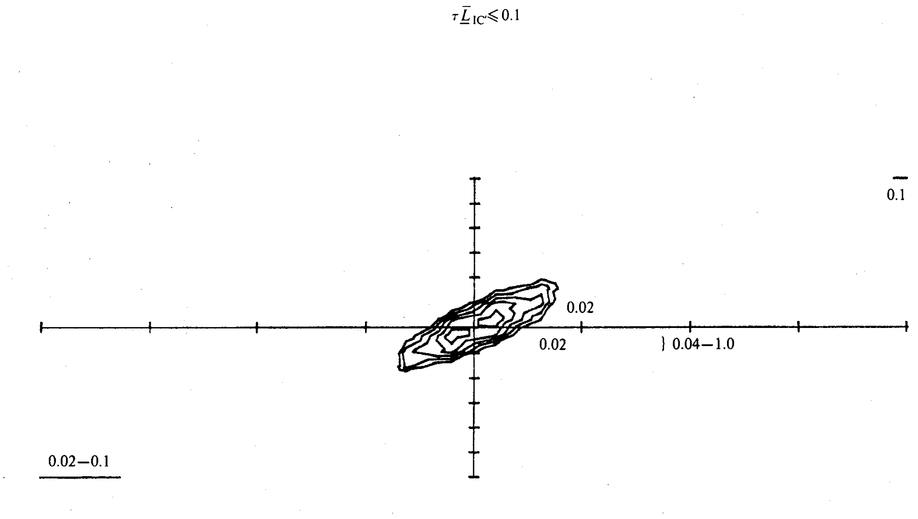

SCALP
SKULL
1 - 1.5 cm
2 mm
I
S
C



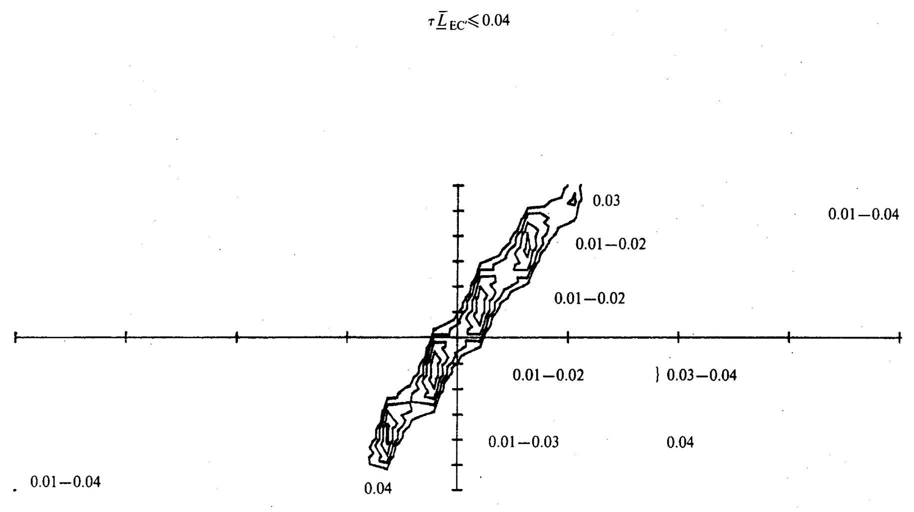

(a*)
10^{-2} μm
Ψ
(a)
10 μm
E
I
Γ
(b)
10^2 μm
M(r)
r
2.5 x 10^3 μm
j
k
(c)
10^4 μm
$\bar{M}$
(a′)
$M^E > 0$
$M^E < 0$
$M^I > 0$
$M^I < 0$
(b′)
$M(r;t)$
$M(r';t)$
τ
$M(r;t+\tau)$
$M(r';t+\tau)$
(c′)
$\bar{M}_{\rm I}(t_0)$
$\bar{M}_{\rm II}(t_0)$
$t-t_0$
$\bar{M}_{\rm I}(t)$
$\bar{M}_{\rm II}(t)$



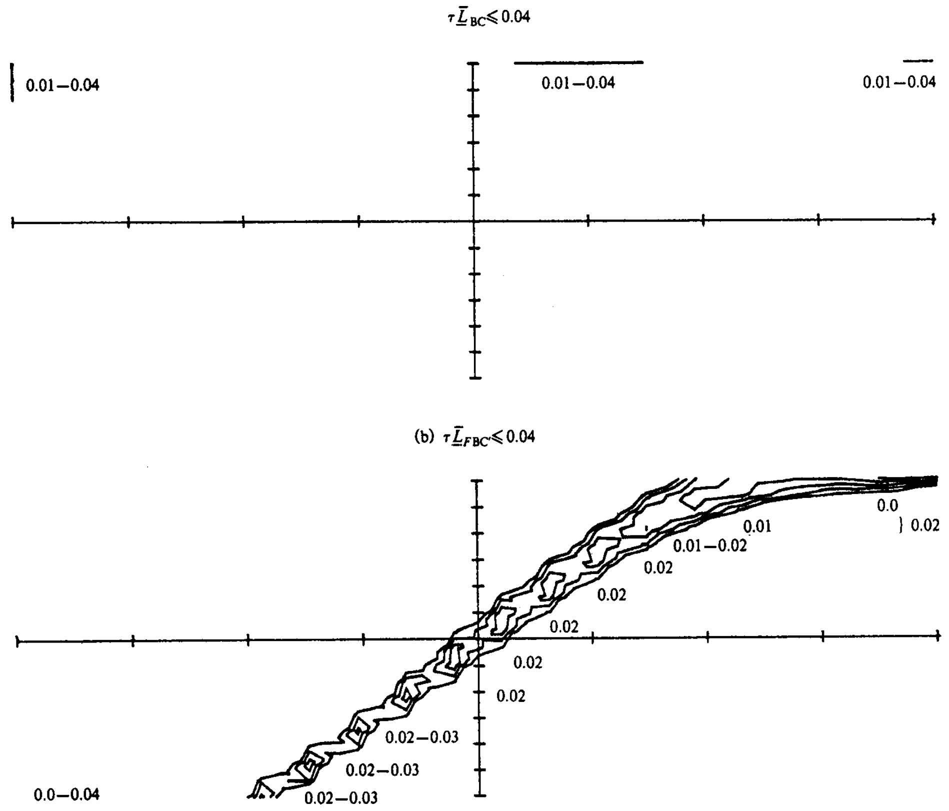

(-125,-25)
(-125,25)
(125,-25)
(125,25)

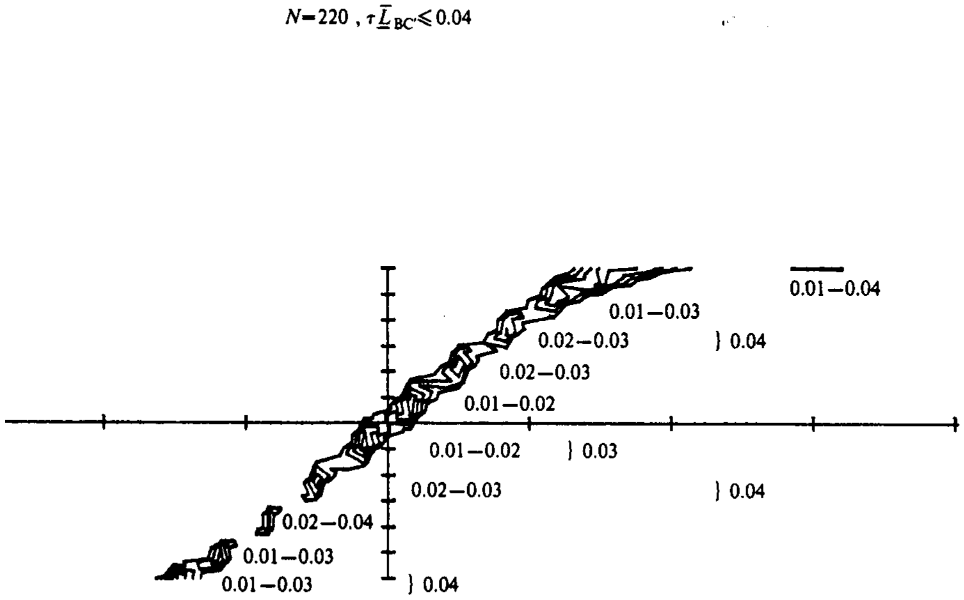


$\tau \bar{L}_{IC} \leqslant 0.1$
0.1
0.02
0.02
} 0.04−1.0
0.02−0.1

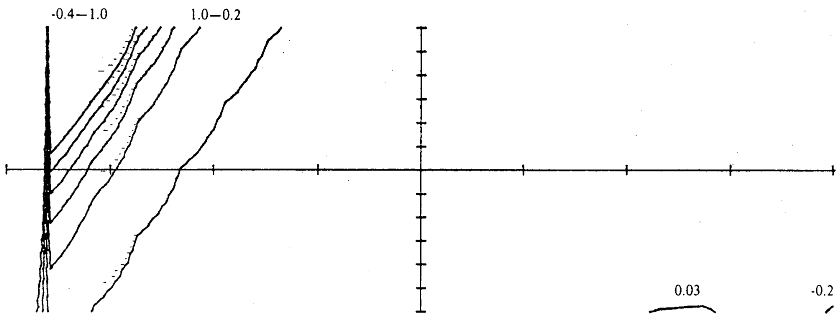


$\tau \bar{L}_{EC'} \leqslant 0.04$
0.03
0.01 − 0.04
0.01 − 0.02
0.01 − 0.02
0.01 − 0.02
} 0.03 − 0.04
0.01 − 0.03
0.04
0.01 − 0.04
0.04

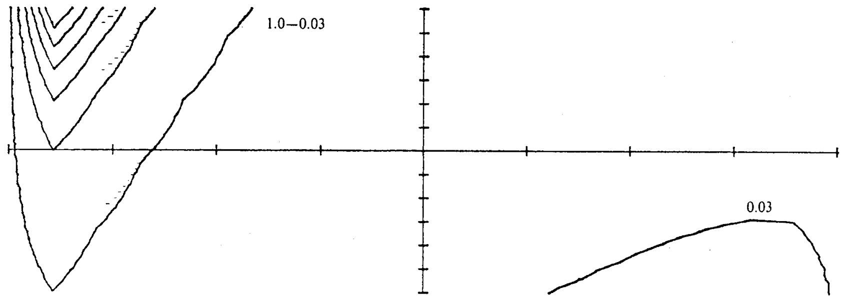

$\tau \underline{\bar{L}}_{BC} \leqslant 0.04$
0.01−0.04
0.01−0.04
0.01−0.04
(b) $\tau \underline{\bar{L}}_{FBC} \leqslant 0.04$
0.0
0.01
} 0.02
0.01−0.02
0.02
0.02
0.02
0.02
0.02
0.02−0.03
0.02−0.03
0.02−0.03
0.0−0.04

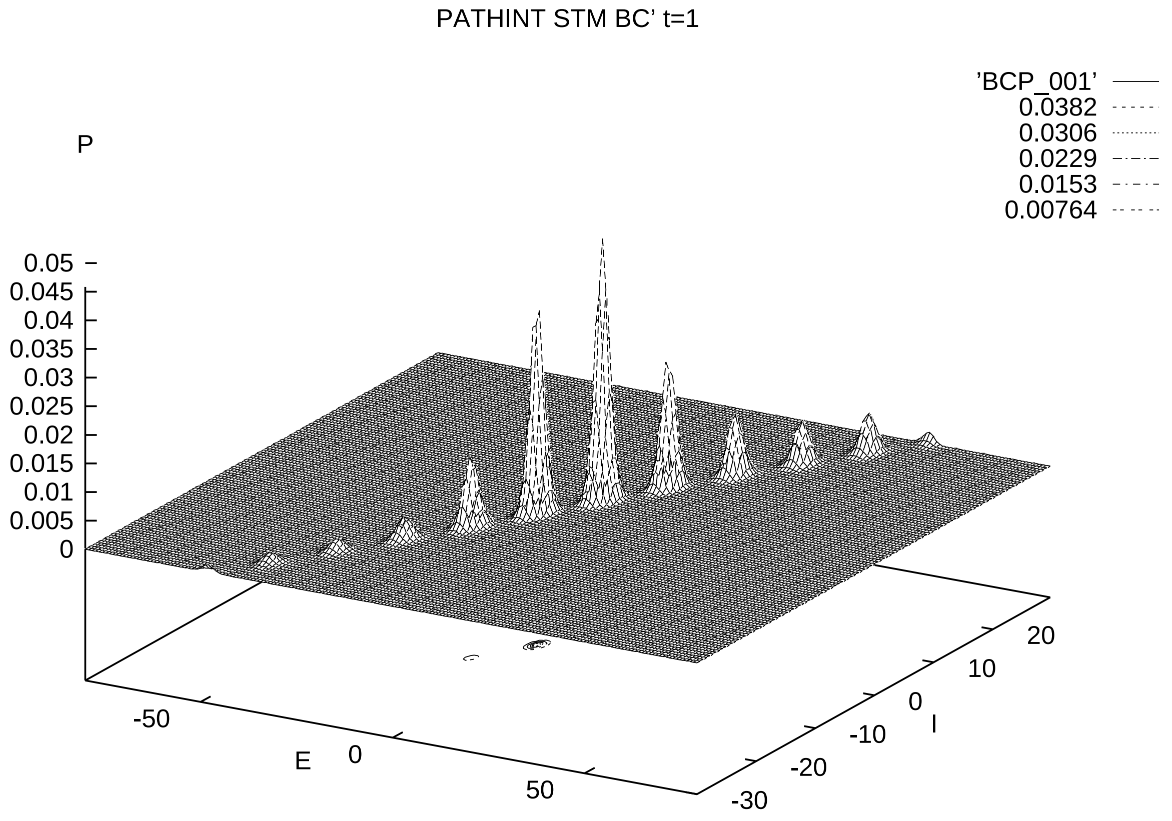

$N=220$, $\tau \underline{\bar{L}}_{BC'} \leqslant 0.04$
0.01−0.04
0.01−0.03
0.02−0.03
} 0.04
0.02−0.03
0.01−0.02
0.01−0.02
} 0.03
0.02−0.03
} 0.04
0.02−0.04
0.01−0.03
0.01−0.03
} 0.04

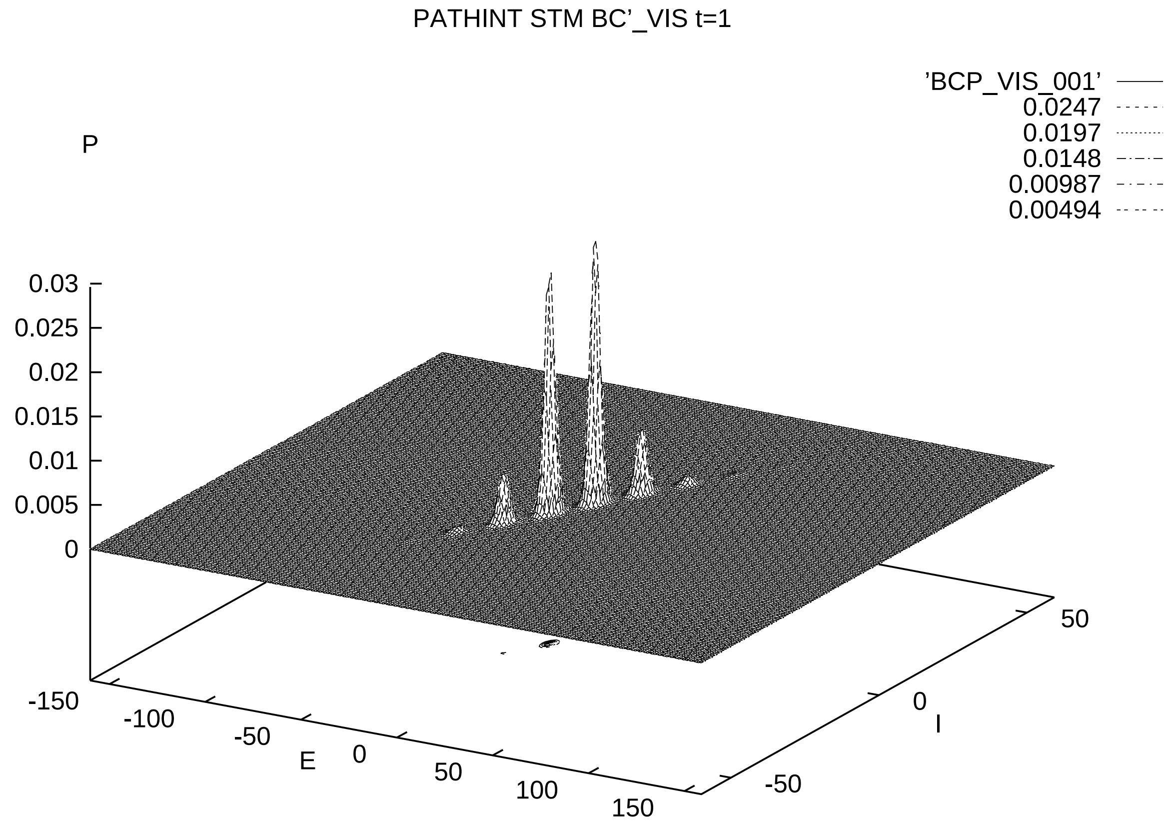
(a) $(\cdots)_{BC}\tau(\rho\nabla M^E)^2 \leqslant 1.0$



(b) $(\cdots)_{BC}\tau(\rho\nabla M^I)^2 \leqslant 1.0$




PATHINT STM BC' t=1
'BCP_001'
0.0382
0.0306
0.0229
0.0153
0.00764
P
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
-50
E
0
50
20
10
0
-10
-20
-30
I


PATHINT STM BC'_VIS t=1
'BCP_VIS_001'
0.0247
0.0197
0.0148
0.00987
0.00494
P
0.03
0.025
0.02
0.015
0.01
0.005
0
-150
-100
-50
E
0
50
100
150
50
0
I
-50