Statistical mechanics of multiple scales of neocortical interactions

Lester Ingber

Lester Ingber Research, P.O. Box 857, McLean, VA 22101

ingber@alumni.caltech.edu

14. Statistical mechanics of multiple scales of neocortical interactions

1. INTRODUCTION

1.1. General philosophy

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 (Haken, 1983; Nicolis and Prigogine, 1973). For example, in classical thermodynamics of equilibrium systems, it is possible to leap from microscopic molecular scales to macroscopic scales, to use the macroscopic concept of temperature to describe the average kinetic energy of microscopic molecular activity, or to use the macroscopic concept of pressure to describe the average rate of change of momentum per unit area of microscopic molecules bombarding the wall of a cavity.

However, 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 thermodynamics typically does not apply (Ma, 1985). For example, the description of weather and ocean patterns, which attempt to include important features such as turbulence, rely on semiphenomenological mesoscopic models, those in agreement with molecular theories but not capable of being rigorously derived from them. Phase transitions in magnetic systems, and many systems similarly modeled (Ma, 1976; K.G. Wilson, 1979; K.G. Wilson and Kogurt, 1974), require careful treatment of a continuum of scales near critical points. In general, rather than having a general theory of nonequilibrium nonlinear process, there are several overlapping approaches, typically geared to classes of systems, usually expanding on nonlinear treatments of stochastic systems (Gardiner, 1983; Haken, 1983; Kubo *et al*, 1973; Nicolis and Prigogine, 1973; van Kampen, 1981). Many biological systems give rise to phenomena at overlapping spatial-temporal scales. For example, the coiling of DNA is reasonably approached by blending microscopic molecular-dynamics calculations with mesoscopic diffusion equations to study angular winding (Cartling, 1989). These approaches have been directed to study electroencephalography (EEG) (Basar, 1980), as well as other biological systems (Goel and Richter-Dyn, 1974).

Therefore, it should not be surprising that the complex human brain supports many phenomena arising at different spatial-temporal scales. What is perhaps surprising is that it seems possible to study truly macroscopic neocortical phenomena such as EEG by appealing to a chain of arguments dealing with overlapping microscopic and mesoscopic scales. A series of papers has developed this statistical mechanics of neocortical interactions (SMNI) (Ingber, 1981c; Ingber, 1982; Ingber, 1983b; Ingber, 1984b; Ingber, 1985c; Ingber, 1985d; Ingber, 1985e; Ingber, 1986b; Ingber, 1988a; Ingber, 1988b; Ingber, 1988c; Ingber, 1991b; Ingber, 1992; Ingber, 1996; Ingber and Nunez, 1990). This approach permits us to find models of EEG whose variables and parameters are reasonably identified with ensembles of synaptic and neuronal interactions. This approach has only recently been made possible by developments in mathematical physics since the late 1970s, in the field of nonlinear nonequilibrium statistical mechanics. The origins of this theory are in quantum and gravitational field theory.

As discussed in Chapter 10, many physical systems have varying degrees of theoretical support leading to Fokker-Planck partial differential equation descriptions. Here, new problems arise in nonlinear nonequilibrium systems, often requiring modeling with the introduction of a nonconstant coefficient of the second-derivative “diffusion” term. The spatial second-derivative term usually represents the kinetic energy, when the first derivative represents the momentum in the differential-equation description of a system. It was early noticed that a similar treatment of the gravitational equation (DeWitt, 1957) required a modification of the potential term of the corresponding Lagrangian. We now better understand the mathematical and physical similarities between classical stochastic processes described by Fokker-Planck equations and quantum processes described by the Schrödinger equation (Langouche *et al*, 1982). 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. As will be demonstrated, this is especially useful to obtain information about the system without solving the time-dependent Fokker-Planck equation; however, we also will describe neocortical phenomena requiring the full solution.

In its differential form, the momentum is proportional to the derivative operator. For classical systems, the coefficient of the square of the momentum is twice the diffusion, e.g., the second moment of a probability distribution describing some systems. The introduction of a nonconstant coefficient of even the first-derivative term requires careful treatment. Such a problem arises for a charged particle in an electromagnetic field (Schulman, 1981), which originally was treated by physical arguments to enforce "minimal coupling," whereby the momentum p is simply replaced by $p - Ae/c$, where A is the electromagnetic potential, e is the electric charge, and c is the speed of light. Minimal coupling excludes other A^2 terms from appearing in the equation of evolution, e.g., Schrödinger's equation. Such problems are related to the operator-ordering of the derivative operators with respect to their nonconstant coefficients. For classical systems, the analogous expression to Ae/c is the drift, e.g., the first moment of a probability distribution describing some systems. In the neocortex, we derive nonlinear expressions for both the drift and diffusion. The detailed mathematical relationships to the physical content of these issues was only clarified in the late 1970s and early 1980s, and is relevant to the mathematics of the neocortex. The first real breakthrough was achieved by noting how these issues should be treated in the context of classical nonlinear nonequilibrium statistical mechanics (Graham, 1977a; Graham, 1977b).

While application has been made of these new mathematical physics techniques in quantum and classical statistical mechanics (Ingber, 1983a; Ingber, 1984a; Ingber, 1986a), we are not merely bringing over techniques to neuroscience from other disciplines out of curiousity. Indeed, the contention appears to be well supported that a mathematical investigation of the neocortex reasonably demands these mathematical techniques, to such an extent that it can be argued that, if the neocortex had been studied and sufficient data collected prior to mathematical developments in quantum or gravitational theory, then these mathematical techniques might have been developed in neuroscience first. The brain is sufficiently complex that it requires the same tools used for similar very complex physical systems. In many ways, we may consider the brain as the prototypical information processing system, and the mathematical techniques used here may be rigorously viewed as filters to describe the processing of this information.

This statistical-mechanics approach, at many junctions in its development, may be intuitively compared to the approach used in simple magnetic systems, a comparison made early in neuronal modeling (Cragg and Temperley, 1954). While caution must be exercised to respect the integrity of the neocortical system, such comparisons can be useful pedagogically. The mathematical approach presented here also has been useful to describe phenomena in social systems, ranging from military command, control and communications (Ingber, 1991a; Ingber, 1993c; Ingber, Fujio, and Wehner, 1991; Ingber and Sworder, 1991), to political systems (Weidlich and Haag, 1983), to pricing bonds on financial markets (Ingber, 1984c; Ingber, 1990; Ingber, Wehner *et al*, 1991). In this context, it has been noted that the activity of neurons may resemble the activity of a throng of people, in which interactions take place at multiple hierarchical levels (Bullock, 1980). The numerical algorithms used in SMNI were developed in part in the process of investigating these other systems.

1.2. Top-down versus bottom-up

There are at least two ways to present this admittedly complex technical approach. First, summarizing material presented in Chapters 2 and 10, we will take the top-down approach, essentially examining some macroscopic issues in EEG measurement. This will motivate us to then look to a bottom-up approach, starting with microscopic synaptic activity and neuronal interactions, then scaling up through mesocolumnar activity of columns of neurons, to finally achieve a reasonable macroscopic description of EEG activity. The confluence of these approaches is expected to yield a tractable approach to EEG analyses (Ingber, 1985c; Ingber, 1991b; Ingber and Nunez, 1990; Nunez, 1989).

This chapter's organization closely follows a more technical paper (Ingber, 1991b), where more details and references can be found. Here, we attempt to convey the foundation, general description, and verification of SMNI. The typical method of presentation in each section below is to first give an "executive summary" of the main concepts being developed, then to fortify this with some sampling of more mathematical arguments; for more detail, the reader is invited to explore the references.

2. TOP-DOWN DIPOLE STRING MODEL

2.1. Modeling of observables

As discussed previously in Chapter 2, the string displacement Φ (potential within the cortex) is given by the wave equation

$$\frac{\partial^2 \Phi}{\partial t^2} - c^2 \frac{\partial^2 \Phi}{\partial x^2} + [\omega_0^2 + f(\Phi)]\Phi = 0 \ , \tag{14.1}$$

in terms of the string displacement Φ, represented by its Fourier decomposition,

$$\Phi(x, t) = \sum_{n=1}^{\infty} G_n(t) \sin k_n x \ , \tag{14.2}$$

where we have assumed zero boundary conditions at the ends of the string for simplicity of presentation, but our observed $\hat{\Phi}$ (referenced as $\Phi\dagger$ in the SMNI papers) is given by

$$\hat{\Phi}(x, t) = \sum_{n=1}^{M} G_n(t) \sin k_n x \ . \tag{14.3}$$

With respect to the cortical medium, we address the question: What can we say about $\hat{\Phi}(x, t)$, the macroscopic observable displacement potential on the scalp or cortical surface? On the basis of previous studies of EEG dispersion relations (Ingber, 1985c; Nunez, 1981a), it would seem that we should be able to describe $\hat{\Phi}$ as a linear or quasilinear variable, but influenced by the local nonlinear behavior that crosses the hierarchical level from mesoscopic to macroscopic (columnar dipoles). How do we mathematically articulate this intuition, for the purposes of consistent description as well as lay the foundation for future detailed numerical calculations? We suggest answers to these questions in the "bottom-up" approach part of this paper.

We examine these issues by taking reasonable synaptic parameters, developing the statistical mechanics of neocortical interactions, and then determining whether they are consistent with observed EEG data. In fact, here we report fits of multichannel human scalp EEG data to these algebraic forms. A current project is investigating the response of the cortical system to given initial driving forces that match or mismatch firing patterns of columnar firings possessed by a given set of synaptic parameters, and under conditions of plastically changing synaptic parameters reflecting changes of these patterns. This should help in clinical diagnoses using the EEG tool.

2.2. Outline of derivation of the nonlinear string model

We use a mechanical-analog model, the string model, derived explicitly for neocortical interactions using SMNI (Ingber and Nunez, 1990). This defines a probability distribution of firing activity, which can be used to further investigate the existence of other nonlinear phenomena, e.g., bifurcations or chaotic behavior, in brain states (Rapp *et al*, 1989).

Previous 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 (Ingber, 1984b). Here, G represents E or I, M^E represents contributions to columnar firing from excitatory neurons, and M^I represents contributions to columnar firing from inhibitory neurons. The object of interest within a short refractory time, τ, approximately 5 to 10 msec, is the Lagrangian $\underline{L}$ for a mesocolumn, detailed further below. $\tau \underline{L}$ can vary by as much as a factor of 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. (Here, "potential" refers to the electric potential, not the potential term in the Lagrangian derived below.) In an abbreviated notation subscripting the time-dependence,

$$\Phi_t - \ll \Phi \gg = \Phi(M_t^E, M_t^I) \approx a(M_t^E - \ll M^E \gg) + b(M_t^I - \ll M^I \gg) , \tag{14.4}$$

where a and b are constants, and $\ll \Phi \gg$ and $\ll M^G \gg$ represent typical minima in the trough. In the context of fitting data to the dynamic variables, there are three effective constants, $\{ a, b, \phi \}$,

$$\Phi_t - \phi = a M_t^E + b M_t^I . \tag{14.5}$$

We scale and aggregate the mesoscopic probability distributions, P, over this columnar firing space to obtain the macroscopic conditional probability distribution over the scalp-potential space:

$$P_\Phi[\Phi] = \int dM^E dM^I P[M^E, M^I] \delta[\Phi - \Phi'(M^E, M^I)] , \tag{14.6}$$

in an abbreviated notation expanded on below. The parabolic trough described above justifies a form

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

$$L_\Phi = \frac{\alpha}{2} |\partial\Phi/\partial t|^2 + \frac{\beta}{2} |\partial\Phi/\partial x|^2 + \frac{\gamma}{2} |\Phi|^2 + F(\Phi) ,$$

$$\sigma^2 = 2\Delta t / \alpha \ , \tag{14.7}$$

where $F(\Phi)$ contains nonlinearities away from the trough, σ^2 is on the order of $1/N$ given the derivation of $\underline{L}$ above, and the integral over x is taken over the spatial region of interest. In general, there also will be terms linear in $\partial\Phi/\partial t$ and in $\partial\Phi/\partial x$.

Previous calculations of EEG phenomena, described below (Ingber, 1985c), show 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 of the system, e.g., represented by σ in the above equation. I.e., this is described by the Euler-Lagrange equations derived from the variational principle possessed by L_Φ (essentially the counterpart to force equals mass times acceleration), more properly by the "midpoint-discretized" Feynman L_Φ, with its Riemannian terms, as discussed below (Ingber, 1982; Ingber, 1983b; Ingber, 1988b). Hence, we can use the variational principle,

$$0 = \frac{\partial}{\partial t} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial t)} + \frac{\partial}{\partial x} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial x)} - \frac{\partial L_\Phi}{\partial\Phi} \ . \tag{14.8}$$

The result is

$$\alpha \frac{\partial^2\Phi}{\partial t^2} + \beta \frac{\partial^2\Phi}{\partial x^2} + \gamma\Phi - \frac{\partial F}{\partial\Phi} = 0 \ . \tag{14.9}$$

If there exist regions in neocortical parameter space such that we can identify $\beta/\alpha = -c^2$, $\gamma/\alpha = \omega_0^2$ (e.g., as explicitly calculated below),

$$\frac{1}{\alpha} \frac{\partial F}{\partial\Phi} = -\Phi f(\Phi) \ , \tag{14.10}$$

and we take x to be one-dimensional, then we recover our nonlinear string, Eq. (14.1) above. Terms linear in $\partial\Phi/\partial t$ and in $\partial\Phi/\partial x$ in L_Φ in Eq. (14.7) can make other contributions, e.g., giving rise to damping terms as discussed in Chapter 10.

The path-integral formulation of the long-time evolution of the distribution P_Φ has a utility beyond its deterministic Euler-Lagrange limit. We have utilized this to explicitly examine the long-time evolution of systems, to compare models to long-time correlations in simulation data (Ingber, 1993c; Ingber, Fujio, and Wehner, 1991), and to approach similar problems in finance (Ingber, 1990; Ingber, Wehner *et al*, 1991) and in EEG modeling as described here. Similar advantages of this approach are available for EEG analyses.

2.3. Macroscopic coarse graining and renormalization

We now are in a position to address the issue posed originally of how to mathematically justify the intuitive coarse-graining of Φ to get $\hat{\Phi}$. In L_Φ above, consider terms of the form

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

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

$$= (2\pi / R) \sum_n G_n^2 \ . \tag{14.11}$$

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

$$p_n[G_n(t + \Delta t) | G_m(t)] \ , \tag{14.12}$$

where we note that in general the $F(\Phi)$ term in L_Φ will require coupling between $G_n(t + \Delta t)$ and $G_m(t)$, $n \neq m$, likely including more than one m. Therefore, we can define

$$P_\Phi = p_1 p_2 \cdots p_\infty \ . \tag{14.13}$$

We now physically and mathematically can define a coarse-graining,

$$P_{\hat{\Phi}} = \int dk_{M+1} dk_{M+2} \cdots dk_\infty \, p_1 p_2 \cdots p_M p_{M+1} p_{M+2} \cdots p_\infty \ . \tag{14.14}$$

I.e., since we have a *bona fide* probability distribution P_Φ, we can integrate over those fine-grained variables, which are not observed. This procedure is one contribution to algorithms used in “renormalization-group” theory (Ma, 1976), to account for multiple intermediate scales of interactions. While other criteria for use of that specific theory certainly are not present here, it is useful to recognize that this is a reasonable phenomenological approach to integrating over many scales of neocortical interactions.

The integration over the fine-grained wave numbers tends to smooth out the influence of the k_n's for $n > M$, effectively "renormalizing"

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

$$\Phi \rightarrow \hat{\Phi} \,,$$

$$L_{\Phi} \rightarrow L^{\dagger}{}_{\hat{\Phi}} \,. \tag{14.15}$$

Eventually, laminar circuitry should be included in both the local and global models. Previous papers have detailed how this can be realized, but more numerical study is needed to determine the degree to which this can be accomplished. As reported here, the solutions are being tested by their goodness of fit to existing EEG data using methods of very fast simulated reannealing (VFSR) (Ingber, 1989; Ingber, 1993b; Ingber and Rosen, 1992), the precursor to a more powerful code, adaptive simulated annealing (ASA) (Ingber, 1993a).

3. BOTTOM-UP SMNI MODEL

3.1. Rationale

We begin our "bottom-up" approach by taking the viewpoint that, since there has been much progress made in mathematically describing the physics at finer spatial-temporal scales, we should use these descriptions to derive a development of the coarser EEG macroscopic scale described above. SMNI has reasonably detailed a consistent physics which at least parallels, with striking numerical specificity, short-term memory (STM) and EEG phenomena at the mesocolumnar scale of hundreds of neurons to the macrocolumnar scale of hundreds of thousands of neurons, in terms of aggregated physics transpiring at the single-neuronal level. The details of this SMNI development of STM and EEG will be used to support the "top-down" development described above.

A major contribution of SMNI is its analytic treatment of minicolumns (Mountcastle, 1978). Minicolumns are observed to span approximately 7×10^2 μm^2, as discussed in Chapter 2. Mesocolumnar domains are defined here as physiological (functional) units, with input converging from an anatomical (structural) macrocolumn extent of approximately 1000 minicolumns, and output diverging from an anatomical minicolumn out to a macrocolumnar extent. Calculations support observations of periodically alternating firings of columnar structures (Fitzpatrick and Imig, 1980; Goldman and Nauta, 1977; Hubel and Wiesel, 1962; Hubel and Wiesel, 1977; Jones *et al*, 1978). As pictured in Fig. 14-1, this microscopic scale is orders of magnitude larger than the molecular scale of membrane biophysics. Also note that "macrocolumns" spanning roughly 7×10^5 μm^2 have been defined as another physiological entity observed in the neocortex (Mountcastle, 1978), but the macroscopic regions considered here are orders of magnitude larger than these. Mesocolumnar domains are sufficiently close to the scale of microscopic neurons to allow direct dependence of this theory on neuronal chemical and electrical properties. The proper stochastic treatment of their interaction permits their development into macroscopic regions responsible for global neocortical information processing. "Thermodynamic" entities corresponding to the "free-energy" potential, "temperature," and order parameters of these macroscopic regions are derived by a statistical-mechanics paradigm, but without recourse to any quasi-equilibrium limits (Landau *et al*, 1980).

Figure 14-1.

Relative to other biological entities, the intrinsic synaptic activity of the most highly evolved mammalian human neocortex functions via the most degenerate and the shortest-ranged neuronal interactions (on the order of micrometers). Here, "degenerate" reflects the mesoscopic state of approximate redundancy of connectivity among microscopic neurons, for purposes of describing this coarser scale. This suggests that many collective aspects of this system may be fruitfully studied similarly to other collective systems, e.g., including magnetic systems, lasers, and more general information-theoretic systems (Haken, 1988; van Kampen, 1981). Collective effects, from clustering (Szentágothai, 1975; Szentágothai and Arbib, 1974) or from statistical interactions (Katchalsky *et al*, 1974), are proposed to be mechanisms of information processing, in addition to the "hard-wiring" mechanisms also possessed by other more ordered cortical entities (Afifi and Bergman, 1978; Sommerhoff, 1974).

Reasonable criteria for any physical approach to the neocortex should include the following three basic features. These also serve to illustrate the appropriate analogies between the neocortex and other collective physical systems.

i. Interactions. Short-ranged neuronal interactions over time periods of several milliseconds should be derived from even more microscopic synaptic activities (Shepherd, 1979). [See Fig. 14-1(a).] Long-ranged spatial interactions from specific neuronal pathways, primarily composed of the relatively low population of long excitatory fibers from ipsilateral association, contralateral commissural, and thalamo-cortical processes must be consistently treated. These long-ranged interactions are also important for collective activity in the mammalian cortex (Braitenberg, 1978), and they are included in this study. Longer-time, weaker and modulatory nonsynaptic influences arising from humoral and electrotonic interactions (Dismukes, 1979; McGeer *et al*, 1978; Schmitt *et al*, 1976) are included, only as their averaged properties affect synaptic parameters.

ii. Statistics. Neurons separated by large distances, across 10^3 to 10^8 neurons, can be statistically coupled via their short-ranged interactions. [See Fig. 14-1(c).] Order parameters, the underlying independent variables at the appropriate scale of the theory, must be identified, and intrinsic fluctuations from the microscopic synaptic and neuronal systems, diffusion effects, must be included. There also are fluctuations of the mesoscopic system due to their aggregated neuronal interactions, derived here as gradient couplings between neighboring mesoscopic cells. These spatially ordered mesoscopic domains respect the observed anatomy and physiology of the neocortex (Szentágothai, 1975; Szentágothai and Arbib, 1974), complementing earlier theories hypothesizing random neural networks (Griffith, 1971; McCulloch and Pitts, 1943).

iii. Dynamics. A viable formalism must be adopted to describe the statistical time evolution of the macroscopic nonequilibrium system over scales of 10^2 to 10^4 msec.

Although cooperativity between distant neurons is typically quite low (Adey, 1978), except perhaps in homologous regions of separate hemispheres, macroscopic regions reflect cooperative behavior, proposed here to best be understood as initiated at the mesoscopic level of interaction. The existence of collective spatial-temporal activity, embedded in a spontaneous noisy background, is supported by statistical analyses of electroencephalographic and magnetoencephalographic recordings (John, 1972; Williamson *et al*, 1979). As long as collective mechanisms arising in a physical system characterized by the above three features are considered to be viable sources of collective neocortical phenomena, then these features must be correctly formulated.

There is a large body of literature dealing with neuronal mechanisms that intuits phenomenological differential equations from rates of change of average excitatory and inhibitory neuronal firings, and then proceeds to search for collective behavior, limit cycles, and oscillatory behavior (Babloyantz and Kaczmarek, 1979; Freeman, 1975; Kaczmarek and Babloyantz, 1977; H.R. Wilson, 1977; H.R. Wilson and Cowan, 1972; H.R. Wilson and Cowan, 1973). Mechanisms are sought to explain varied phenomena such as hysteresis in perception (H.R. Wilson, 1977), perception and learning (Hebb, 1949; Takeuchi and Amari, 1979), and ontogenesis of columnar formation (Takeuchi and Amari, 1979; von der Marlsburg, 1979). Comparisons with applications of these techniques to those used in other physical systems (Haken, 1983), illustrates that the pioneering application of these appropriate formalisms to the neocortical system still has much to offer. Much inspiration for these applications has come from work in nonequilibrium thermodynamics, which has been applied to specific systems, e.g., chemical reactions, lasers, magnetic systems, fluids, spin glasses, etc., as well as to the general formulation of complex nonlinear systems (Haken, 1983; Katchalsky *et al*, 1974; Kubo *et al*, 1973; Nicolis and Prigogine, 1973; van Kampen, 1976).

This study also distinguishes between neuronal mechanisms the neocortex uses to process information and the structures of information the neocortex processes. A Lagrangian is derived that operates on firings of the system. When integrated over a time period, this yields the nonequilibrium equivalent of a "thermodynamic potential." This Lagrangian is derived, not conveniently defined or hypothesized, from the short-time evolution of the probability distribution of columnar firing states. The exponential of minus the Lagrangian, essentially this short-time distribution up to an important normalization factor, operates as a weighting factor on all possible states, filtering or transforming (patterns of) input firings into output firings. "Information" is a concept well defined in terms of the probability eigenfunctions of electrical-chemical activity of this Lagrangian. The path-integral formulation presents an accurate intuitive picture of an initial probability distribution of patterns of firings being filtered by the (exponential of the) Lagrangian, resulting in a final probability distribution of patterns of firing.

In the conclusion below is a brief summary of the advantages of the SMNI model in the context of its practical application to enhancing the extraction of signal out of noise from EEG data.

3.2. Microscopic neurons

3.2.1. *General description*

Figure 14-1(a) illustrates the microscopic neuronal interaction scale, on the order of several micrometers. Neocortical neurons typically have many dendrites that receive quanta of chemical postsynaptic stimulation from many other neurons. The distribution of quanta transmitted across synapses takes place on the scale of 10^{-2} μm, as illustrated in the inset of Fig. 14-1(a*). Each quantum has thousands of molecules of chemical neurotransmitters that affect the chemically gated postsynaptic membrane. Chemical transmissions in the neocortex are believed to be either excitatory (E), such as glutamic acid, or inhibitory (I), such as γ aminobutyric acid. There exist many transmitters as well as other chemicals that modulate their effects, but it is assumed that after millions of synapses between hundreds of neurons are averaged over, then it is reasonable to ascribe a distribution function Ψ with a mean and variance for E and I interneuronal interactions.

Some neuroscientists do not accept the assumption 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), i.e., including the absolute refractory time after a firing during which no new spikes can be generated, and the relative refractory period during which spikes can be produced only at a decreased sensitivity (Sommerhoff, 1974). However, 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.

This same averaging procedure makes it reasonable to ascribe a distribution function Γ with a mean and variance for E and I intraneuronal interactions. A Gaussian Γ is taken to describe the distribution of electrical polarizations caused by chemical quanta impinging on the postsynaptic membrane. These polarizations give a resultant polarization at the base of the neuron, the axon [extension in Fig. 14-1(a) cut by the double broken line]. The base of the axon of a large fiber may be myelinated. However, smaller neurons typically lack these distinguishing features. Experimental techniques are not yet sufficiently advanced to attempt the explicit averaging procedure necessary to establish the means and variances of Ψ and Γ, and their parameters, *in vivo* (Vu and Krasne, 1992). Differential attenuations of polarizations from synapses to the base of an axon are here only phenomenologically accounted for by including these geometric and physiological effects into Γ.

With a sufficient depolarization of approximately 10 to 20 mV at the soma, within an absolute and relative refractory period of approximately 5 msec, an action potential is pulsed down the axon and its many collaterals, affecting voltage-gated presynaptic membranes to release quanta of neurotransmitters. Not detailed here is 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). At present, Ψ and Γ are taken to approximate this biophysics for use in macroscopic studies. The formalism adopted in this study is capable of using new microscopic functional dependences, gleaned from other experimental or theoretical investigations, and cranking them through to obtain similar macroscopic descriptions. Chemical independence of excitatory depolarizations and inhibitory hyperpolarizations are well established in the neocortex, and this independence is retained throughout this study.

It should be noted that experimental studies initially used to infer Ψ and Γ (e.g., at neuromuscular junctions) were made possible by deliberately reducing the number of quanta by lowering external calcium concentrations (Boyd and Martin, 1956; Katz, 1966). Ψ was found to be Poissonian, but in that system, where hundreds of quanta are transmitted *in vivo*, Ψ may well be otherwise; for example, Gaussian with independent mean and variance. Current research suggests a binomial distribution, having a Poisson limit (Ingber, 1982; Korn, Mallet, and Faber, 1981; Perkel and Feldman, 1979). Note that some investigators have shown a Bernoulli distribution to be more accurate in some cases (Ingber, 1982; Korn and Mallet, 1984; Perkel and Feldman, 1979), and that the very concept of quantal transmission, albeit that good fits to experimental data are achieved with this concept, is under review. In the neocortex, probably small numbers of quanta are transmitted at synapses, but other effects, such as nonuniformity and nonstationarity of presynaptic release sites, and nonlinear summation of postsynaptic potentials, may detract from a simple phenomenological Poisson description (Shepherd, 1979). This short description serves to point out possible differences in Ψ resulting from many sources. However, the derivation of synaptic interactions given here makes it plausible that for reasonable neuronal parameters, the statistical folding of Ψ and Γ is essentially independent of the functional form assumed for Ψ, just requiring specification of its numerical mean and variance.

The result of this analysis is to calculate the transition probability of the firing of neuron j, p_{σ_j}, given its interaction with its neighbors that also may fire or not fire. The result is given as the tabulated error function. Within the range where the total influences of excitatory and inhibitory firings match and exceed the average threshold potential of a given neuron, the probability of that neuron firing receives its major contribution to increase from 0 towards 1. A step function derived as $\tanh F^G$ is defined by the "threshold factor" F_j. That is, for $F^G \gg 1$, $\tanh F^G \to 1$, while for $F^G \ll 1$, $\tanh F^G \to -1$, and so the threshold region of $-1 < F^G < 1$ sensitively controls this important $\tanh F^G$ contribution to the drifts, the driving terms, in the Lagrangian. The mesoscopic development discussed below retains this sensitivity.

This is similar to the mathematical result obtained by others (Little, 1974; Little and Shaw, 1978; Shaw and Vasudevan, 1974) who have modeled the neocortex after magnetic systems (Cragg and Temperley, 1954). However, the following is derived more generally, and has the neural parameters more specifically denoted with different statistical significances given to Ψ and Γ, as described above.

3.2.2. *Conditional probability*

Consider $10^2 < N < 10^3$ neurons, labeled by k, interacting with a given neuron j. Each neuron may contribute many synaptic interactions to many other neurons. A neuron may have as many as $10^4 - 10^5$ synaptic interactions. Within time $\tau_n \approx 5$ msec, Ψ is the distribution of q quanta of chemical transmitter released from neuron k to neuron j ($k \neq j$) with mean a_{jk}, where

$$a_{jk} = A_{jk}(\sigma_k + 1)/2 + B_{jk} \ . \tag{14.16}$$

A_{jk} is the conductivity weighting transmission of polarization, dependent on k firing,

$$\sigma_k = \begin{cases} 1, & k \text{ fires,} \\ -1, & k \text{ does not fire} \end{cases} \tag{14.17}$$

and B_{jk} is a background including some nonsynaptic and long-range activity. Of course, A and B are highly complicated functions of kj. This definition of σ_k permits a decomposition of a_{jk} into two different physical contributions. At this point there is a reasonable analogy to make with magnetic systems, where σ_k might represent a unit spin. However, the details of the interactions between neurons differ from those between magnetic spins, and this greatly affects such comparisons.

Further SMNI development yields the conditional probability, p_{σ_j}, of neuron j firing given previous firings within τ of other neurons k:

$$p_{\sigma_j} = \pi^{-\frac{1}{2}} \int_{(\sigma_j F_j \sqrt{\pi}/2)}^{\infty} dz \, \exp(-z^2)$$

$$= \tfrac{1}{2} \, [1 - \mathrm{erf}(\sigma_j F_j \sqrt{\pi}/2)] ,$$

$$F_j = \frac{V_j - \sum_k a_{jk} v_{jk}}{(\pi \sum_{k'} a_{jk'} (v_{jk'}^2 + \phi_{jk'}^2))^{\frac{1}{2}}} \; . \tag{14.18}$$

"erf" is the tabulated error function, simply related to the normal probability function (Mathews and Walker, 1970). F_j^G is a "threshold factor," as p_{σ_j} increases from 0 to 1 between $\infty > \sigma_j F_j > -\infty$ sharply within the range of $F_j \approx 0$.

If

$$|\sigma_j F_j| < 1 , \tag{14.19}$$

then an asymptotic expression for p_{σ_j} is

$$p_{\sigma_j} \approx \frac{\exp(-\sigma_j F_j)}{\exp\,(F_j) \; + \; \exp(-F_j)} \; . \tag{14.20}$$

3.3. Mesoscopic domains

3.3.1. *General description*

As is found for most nonequilibrium systems, e.g., for lasers, chemical systems, fluids, and ecological systems (Haken, 1983; van Kampen, 1981), a mesoscopic scale is required to formulate the statistical mechanics of the microscopic system, from which the macroscopic scale can be developed (Haken, 1983). The neocortex is particularly interesting in this context in that a clear scale for the mesoscopic system exists, both anatomically (structurally) and physiologically (functionally). "Minicolumns" of about

$N \approx 110$ neurons (about 220 in the visual cortex) comprise modular units vertically oriented relative to the warped and convoluted neocortical surface throughout most, if not all, regions of the neocortex (Gilbert and Wiesel, 1983; Goldman and Nauta, 1977; Hubel and Wiesel, 1962; Imig and Reale, 1980; Jones *et al*, 1978; Mountcastle, 1978). Clusters of about 100 neurons have been deduced to be reasonable from other considerations as well (Bullock, 1980). Since the short-ranged interactions between neurons take place within ~1 mm, which is the extent of a "macrocolumn" comprising ~ 10^3 minicolumns of $N^* \approx 10^5$ neurons, and since macrocolumns also exhibit rather specific information-processing features, this theory has retained the divergence-convergence of macrocolumn-minicolumn, efferent-afferent interactions by considering domains of minicolumns as having similar synaptic interactions within the extent of a macrocolumn. This macrocolumnar-averaged minicolumn is designated in this theory as a "mesocolumn."

This being the observed situation, it is interesting that $N \approx 10^2$ is just the right order of magnitude to permit a formal analysis using methods of mathematical physics just developed for statistical systems in the late 1970s (Graham, 1977a; Langouche *et al*, 1982). N is small enough to permit nearest-neighbor interactions to be formulated, such that interactions between mesocolumns are small enough to be considered gradient perturbations on otherwise independent mesocolumnar firing states. This is consistent with rather continuous spatial gradient interactions observed among columns (Dykes, 1983), and with the basic hypothesis that nonrandom differentiation of properties among broadly tuned individual neurons coexists with functional columnar averages representing superpositions of patterned information (Erickson, 1982). This is a definite mathematical convenience; otherwise, a macrocolumn of ~ 10^3 minicolumns would have to be described by a system of minicolumns with up to 16th-order next-nearest neighbors. (Consider 1000 minicolumns spread out in a two-dimensional grid about 33 by 33 minicolumns, and focus attention on the center minicolumn.) Also, N is large enough to permit the derived binomial distribution of afferent minicolumnar firing states to be well approximated by a Gaussian distribution, a luxury not afforded an "average" neuron, even in this otherwise similar physical context. Finally, mesocolumnar interactions are observed to take place via one to several relays of neuronal interactions, so that their time scales are similarly $\tau \approx 5 - 10$ msec.

Even after statistically shaping the microscopic system, the parameters of the mesoscopic system are still macrocolumnar-averaged synaptic parameters, i.e., reflecting the statistics of millions of synapses with regard to their chemical and electrical properties. Explicit laminar circuitry, and more complicated synaptic interactions, e.g., dependent on all combinations of presynaptic and postsynaptic firings, can be included without loss of detailed analysis (Ingber, 1983b).

The mathematical development of mesocolumns establishes a mesoscopic Lagrangian $\underline{L}$, which may be considered as a "cost function" with variables M^G, $\dot{M}^G$, and ∇M^G, and with parameters defined by the macrocolumnar-averaged chemical-electrical entities developed below. (See Fig. 14-2.)

The Einstein summation convention is used for compactness, whereby any index appearing more than once among factors in any term is assumed to be summed over, unless otherwise indicated by vertical bars, e.g., $|G|$. The mesoscopic probability distribution P is given by the product of microscopic probability distributions p_{σ_i}, constrained such that the aggregate mesoscopic excitatory firings $M^E = \sum_{j \in E} \sigma_j$, and the aggregate mesoscopic inhibitory firings $M^I = \sum_{j \in I} \sigma_j$.

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

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

$$\approx \prod_G (2\pi \tau g^{GG})^{-1/2} \exp(-N \tau \underline{L}^G) , \tag{14.21}$$

where the final form is derived using the fact that $N > 100$. $\bar{G}$ represents contributions from both E and I sources. This defines the Lagrangian, in terms of its first-moment drifts g^G, its second-moment diffusion matrix $g^{GG'}$, and its potential $\underline{V}'$, all of which depend sensitively on threshold factors F^G,

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

$$\underline{L} = \underline{T} - \underline{V} ,$$

$$\underline{T} = (2N)^{-1} (\dot{M}^G - g^G) g_{GG'} (\dot{M}^{G'} - g^{G'}) ,$$

$$\underline{V} = \underline{V}' - M^G J_G / (2N \tau) ,$$

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

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

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

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

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

$$a^G_{G'} = \frac{1}{2} A^G_{G'} + B^G_{G'} \ , \tag{14.22}$$

where $A^G_{G'}$ and $B^G_{G'}$ are macrocolumnar-averaged interneuronal synaptic efficacies, $v^G_{G'}$ and $\phi^G_{G'}$ are averaged means and variances of contributions to neuronal electric polarizations, and nearest-neighbor interactions $\underline{V}'$ are detailed in other SMNI papers (Ingber, 1982; Ingber, 1984b). $M^{G'}$ 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. 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. This mathematical description defines the mesocolumn concept introduced previously.

Figure 14-2.

At this stage, severe approximation in modeling is typically required in order to proceed towards solutions. However, advantage can be taken of experimentally observed columnar structure to first attempt to analytically scale the neuronal system into mesoscopic domains that are still relatively microscopic compared to the macroscopic regions to be described (Fitzpatrick and Imig, 1980; Goldman and Nauta, 1977; Hubel and Wiesel, 1977; Jones *et al*, 1978; Mountcastle, 1978; Takeuchi and Amari, 1979). For purposes of macroscopic description, the minicolumnar structure effectively spatially averages the neuronal interactions within one to several firing periods.

The following development is proposed, which: (1) reasonably includes and averages over millions of synaptic interactions that exist between groups of hundreds of neurons, (2) analytically establishes the integrity of columnar domains and specifies their interactions, and (3) prepares the formulation of (1) and (2) to foresee their analytic inclusion into studies of macroscopic regions.

For purposes of detailing anatomical or physiological properties of neurons, it is simply incorrect to postulate an "average" neuron. However, for the purpose of macroscopic brain function, when considering millions of neurons, we repeat that it is reasonable to at least respect the incredibly similar modular structure present in all regions of the neocortex (Gilbert and Wiesel, 1983; Goldman and Nauta, 1977; Hubel and Wiesel, 1962; Imig and Reale, 1980; Jones *et al*, 1978; Mountcastle, 1978), still allowing for the differentiation among the laminar structure of individual modules and among neurons active at different temporal and spatial scales.

In this context, the neocortex has about 5×10^{10} neurons distributed rather uniformly over approximately 5×10^8 minicolumns. (The visual cortex has double this density.) Within these minicolumns, a "vertical" structure is defined perpendicular to six highly convoluted laminae of total thickness $\approx 2.5 \times 10^3$ μm, principally by the efferent pyramidal cells. They exhibit vertical apical bundling of their dendrites in the upper laminae, and some of their recurrent axonal collaterals also ascend to upper laminae. A number of other fusiform, Martinotti, and stellate cells (granule cells in the sensory cortex and basket cells in the motor cortex) also contribute to this vertical organization. In general, laminae I to IV are afferent and laminae V and VI are efferent (Afifi and Bergman, 1978).

However, "horizontal" dendritic basal arborizations (treelike structures) of the pyramidal cells, tangential to the laminae, horizontal axonal collaterals of the pyramidal cells, and horizontal processes of stellate, Martinotti, and neonatal horizontal cells, all impart horizontal stratification to columnar interactions. Therefore, although the columnar concept has anatomical and physiological support, the minicolumnar boundaries are not so clearly defined (Dykes, 1978). If this stratification and other long-ranged afferent input to groups of minicolumns are incorporated, then it is possible that future work may have to define a physiological unit that encompasses a mesocolumn consisting of one to perhaps several minicolumns outputting to a macrocolumar extent. This study formalizes these circumstances by defining a mesocolumn with extent greater than 10^2 μm, as an intermediate integral physiological unit encompassing one minicolumn. [See Fig. 14-1(b).] Dynamic nearest-neighbor interactions between mesocolumns are analytically defined by their overlapping neuronal interactions, in accordance with the observed horizontal columnar stratifications outlined above. [See Fig. 14-1(b').] This approach permits future analytic modifications, as differences between inter- and intra-minicolumnar interactions and circuitries become experimentally clarified.

The resulting picture of columnar interactions is relatively simpler than a mass of interacting neurons, but not so simple to the point of uselessness. A collection of average excitatory and inhibitory neuronal firings, as depicted in Fig. 14-1(a'), now define a continuum of mesocolumnar firings. A zero-order binomial distribution is easily intuited: Let G denote E or I firings. Using the magnetic analogy, consider E as a spin-up and I as a spin-down magnet. A column of N^G neurons can have a total firing of Nn^G, within time τ, where n^G is the fraction firing, ranging by 2's between $-N^G \le Nn^G \le N^G$. (Count firing as +1, nonfiring as −1.) For convenience, assume $Nn^G > 0$, which arises from Nn^G firings plus $\frac{1}{2}(N^G - Nn^G)$ cancelling pairs of firings and nonfirings. This gives a total of $\frac{1}{2}(N^G - Nn^G) + Nn^G = \frac{1}{2}(N^G + Nn^G)$ firings and $N^G - \frac{1}{2}(N^G + Nn^G)$ nonfirings. The degeneracy factor, as a function of the firing rate Nn^G, is the number of ways N^G neurons can produce a given firing pattern, i.e., the binomial distribution. Note that the binomial coefficient is unity for states of all firing or all nonfiring, and peaks as $N^G!/[(N^G/2)!]^2 \approx 2^{N^G+\frac{1}{2}}(\pi N^G)^{-\frac{1}{2}}$ for $Nn^G = 0$. In the range $Nn^G \approx 0$, there is maximal degeneracy of information encoded by mesocolumnar firings. This argument analytically articulates the meaning of "neuronal degeneracy" and also of the ubiquitous, often ambiguous "average neuron." However, reasonable properties of mesocolumns, not of average neurons, are developed here for macroscopic study.

The properly calculated distribution contains nearest-neighbor mesocolumnar interactions expressed as spatial-derivative correction terms. This verifies that in macroscopic activity, where patterns of mesocolumnar firing vary smoothly over neighboring mesocolumns, it is consistent to approximate mesocolumnar interactions by including only second-order gradient correction terms. We calculate macroscopic states of mesocolumnar firings, which are subject to these constraints. Excitatory and inhibitory sensitivity to the neuronal parameters survives, similar to the sensitivity encountered by single neurons.

Nearest-neighbor interactions are "induced" between minicolumns in the following way. The bulk of short-ranged interactions engaging the neurons in a minicolumn do not take place within the minicolumn, but rather within a spatial extent the size of a macrocolumn, comprising roughly 1000 minicolumns. If we consider the area of influence of a minicolumn as extending out to a macrocolumn's reach, then the area of interactions engaged by nearest-neighbor minicolumns has an offset circle of influence (Ingber, 1984b). (See Fig. 14-3.) Therefore, within one or two epochs spanning the refractory periods of a typical neuron, interactions engaged by a given minicolumn can be extended out to areas of influence engaged by their nearest neighbors. This is what physically is being calculated by a careful mathematical treatment of overlapping interactions. In this manner, microscopic degrees of freedom of many types of neurons (many of which are only crudely classified by the above definitions), synapses, neurotransmitters, cellular architecture, and circuitries, may be practically weighted and averaged for macroscopic considerations.

Figure 14-3.

In the steps outlined above, the mesocolumnar conditional probability that a given mesocolumn will fire is calculated, given its direct interactions with other mesocolumns just previously fired. Thus a transition rate from one state of mesocolumnar firing to another state closely following the first state is obtained. A string, or path of these conditional probabilities connects the mesocolumnar firings at one time to the firing at any time afterwards. Many paths may link the same initial and final state. In this way the long-time conditional probability of all possible mesocolumnar firings at any given time is obtained. A Lagrangian is thereby derived which explicitly describes the time evolution of the neocortical region in terms of its initial distribution of firings, and expressed in terms of its mesoscopic order parameters which retain a functional form derived from microscopic neuronal interactions. A major benefit derived from this formalism is a variational principle that permits extrema equations to be developed. This also makes it possible to draw analogies to the "orienting field" and "temperature" of other collective systems.

3.3.2. *Further development of mesocolumn model*

As pointed out in this derivation (Ingber, 1982; Ingber, 1983b), this microscopic scale itself represents a high aggregation of submicroscopic scales, aggregating effects of tens of thousands of quanta of chemical transmitters as they influence the 5×10^{-3} μm scale of biomolecular leaflets of phospholipid molecules. This microscopic scale has been aggregated up to the mesoscopic scale, again using the general property of probability distributions, that the aggregated distribution P_q of variable q is calculated from the joint distribution $P_{q_1 q_2}$ of underlying variables q_1 and q_2,

$$P_q(q) = \int dq_1 dq_2 P_{q_1 q_2}(q_1, q_2)\, \delta(q - (q_1 + q_2)) \ . \tag{14.23}$$

To summarize up to this point, the mathematical development of mesocolumns establishes a mesoscopic Lagrangian $\underline{L}$, defining the short-time probability distribution of firings in a minicolumn, composed of ~ 10^2 neurons (Gilbert and Wiesel, 1983; Goldman and Nauta, 1977; Hubel and Wiesel, 1962; Imig and Reale, 1980; Jones *et al*, 1978; Mountcastle, 1978; Szentágothai, 1978), given its just previous interactions with all other neurons in its macrocolumnar surround.

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

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

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

$$P \approx (2\pi\tau)^{-1/2} g^{1/2} \exp(-N\tau \underline{L}) \ , \tag{14.24}$$

where $\underline{L}$ is defined in terms of its drift, diffusion, and potential, all of which depend sensitively on the threshold factor F^G,

$$F^G = \frac{V^G - a_{G'}^{|G|} v_{G'}^{|G|} N^{G'} - \frac{1}{2} A_{G'}^{|G|} v_{G'}^{|G|} M^{G'}}{\{\pi[(v_{G'}^{|G|})^2 + (\phi_{G'}^{|G|})^2](a_{G'}^{|G|} N^{G'} + \frac{1}{2} A_{G'}^{|G|} M^{G'})\}^{1/2}} \ . \tag{14.25}$$

In the first SMNI papers, long-ranged interactions were included in $\underline{L}$ by adding potential terms simulating these constraints, i.e., adding $J_G M^G$ to $\underline{L}$, where J_G was numerically adjusted to account for these interactions.

In order to more properly include long-ranged fibers (i.e., cortico-cortical axons, as discussed in Chapter 2), so that interactions among macrocolumns may be included in the numerical EEG studies, the J_G terms are dropped, and more realistically replaced by a modified threshold factor F^G,

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

$$A_E^{\ddagger I} = A_I^{\ddagger E} = A_I^{\ddagger I} = B_E^{\ddagger I} = B_I^{\ddagger E} = B_I^{\ddagger I} = 0 ,$$

$$a_E^{\ddagger E} = \frac{1}{2} A_E^{\ddagger E} + B_E^{\ddagger E} . \tag{14.26}$$

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^*: Nearly every pyramidal cell has an axon branch that makes a cortico-cortical connection; i.e., the number of cortico-cortical fibers may be as high as 10^{10} (Braitenberg, 1978).

At this point, attention is also drawn to the similar algebraic structure of the threshold factors in Eqs. (14.18) and (14.26), illustrating common forms of interactions between their entities, i.e., neurons and columns of neurons, respectively. The nonlinear threshold factors are defined in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges.

The net short-time probability distribution can be folded over and over (multiplied) in time increments Δt to yield a path-integral algorithm for calculating the long-time probability distribution (Langouche *et al*, 1982). This result depends on the use of the Markov property of our distribution, wherein the short-time evolution of the system at time $t + \tau$ depends only on the state of the system at time t. For example, in a very compacted notation, labeling u intermediate time epochs by s, i.e., $t_s = t_0 + s\Delta t$, in the limits $u \to \infty$ and $\Delta t \to 0$, and assuming $M_{t_0} = M(t_0)$ and $M_t = M(t \equiv t_{u+1})$ are fixed,

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

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

$$P[M_t|M_{t_0}] = \int \cdots \int \underset{\sim}{D}M \exp\left(- \sum_{s=0}^{u} \Delta t \underset{\sim}{L}_s \right) ,$$

$$\underset{\sim}{D}M = (2\pi \hat{g}_0^2 \Delta t)^{-1/2} \prod_{s=1}^{u} (2\pi \hat{g}_s^2 \Delta t)^{-1/2} dM_s \ . \tag{14.27}$$

Similarly, the short-time probability distribution P can be folded over and over at each point r, to give a field-theoretic Lagrangian, $L(r, t)$. The above "prepoint-discretization" representation of $\underset{\sim}{L}$ derived for the neocortex, e.g., $g_s^G = g^G[M^G(t_0 + s\Delta t)]$, disguises the Riemannian geometry induced by the nonconstant metric $g_{GG'}$, discussed further below.

3.4. Macroscopic development

3.4.1. *General description*

Inclusion of all the above microscopic and mesoscopic features of the neocortex permits a true nonphenomenological Gaussian-Markovian formal development for macroscopic regions encompassing ~ 5×10^3 macrocolumns of spatial extent ~ 5×10^9 μm^2, albeit one that is still highly nonlinear and nonequilibrium. The development of mesocolumnar domains presents conditional probability distributions for mesocolumnar firings with spatially coupled nearest-neighbor interactions. The macroscopic spatial folding of these mesoscopic domains and their macroscopic temporal folding of tens to hundreds of τ, with a resolution of at least τ/N (Ingber, 1984b), yields a true path-integral formulation, in terms of a macroscopic Lagrangian possessing a variational principle for most-probable firing states. At this point in formal development, no continuous-time approximation has yet been made; this is done, with clear justification, only for some applications discussed below. Much of this algebra is greatly facilitated by, but does not require, the use of Riemannian geometry to develop the nonlinear means, variances, and "potential" contributions to the Lagrangian (Langouche *et al*, 1982).

This formalism can also be recast in several other representations (Langouche *et al*, 1982), perhaps more familiar to other investigators, and sometimes more useful for particular calculations. For example, a Hamiltonian formulation can be obtained, one that does not permit simple "energy"-type conservation approximations, but one that does permit the usual time-evolution picture. The time-dependent differential macroscopic probability distribution, or "propagator," is found to satisfy a true Fokker-Planck second-order partial-differential equation, expressing the rate of change of the distribution as the sum of contributions from nonlinear drifts and diffusion in the space of E and I firings. With respect to a reference stationary state, a well-defined information, or "entropy," can be formulated. Also, a set of Langevin rate equations for E and I firings can be obtained, expressing the rate of change of firings as the sum of drifts and multiplicative noise. The most-probable firing states derived variationally from the path-integral Lagrangian as the Euler-Lagrange equations represent a reasonable average over the noise in the Langevin system; the noise cannot be indiscriminately neglected. Because of the presence of multiplicative noise, the Langevin system differs in its Itô (prepoint) and Stratonovich (midpoint) discretizations. Furthermore, there exists a midpoint-discretized covariant description, in terms of the Feynman Lagrangian $\underline{L}_F$, which is defined such that (arbitrary) fluctuations occur about solutions to the Euler-Lagrange variational equations. In contrast, the usual Itô and corresponding Stratonovich discretizations are defined such that the path integral reduces to the Fokker-Planck equation in the weak-noise limit.

Using the Lagrangian formulation, a systematic numerical procedure has been developed for fitting parameters in such stochastic nonlinear systems to data using methods of ASA (Ingber, 1989; Ingber, 1993a; Ingber, 1993b), and then integrating the path integral using a non-Monte Carlo technique especially suited for nonlinear systems (Wehner and Wolfer, 1983a; Wehner and Wolfer, 1983b; Wehner and Wolfer, 1987). This numerical methodology has been applied with success to military modeling (Ingber, Fujio, and Wehner, 1991; Ingber and Sworder, 1991) and to financial markets (Ingber, 1990; Ingber, Wehner *et al*, 1991), and, as discussed further below, we are using it in this neocortical system to correlate EEG to behavioral states (Ingber, 1991b; Ingber and Nunez, 1990).

3.4.2. *Riemannian geometry*

A series of papers has recognized that a few of the most popular Riemannian-geometric transformation properties possessed by physics systems might be advantageous for a theory of cortical interactions, i.e., most specifically in the cerebellum, and they have gone further to postulate this geometry as the essential component of their theory (A. Pellionisz and Llinás, 1979; A. Pellionisz and Llinás, 1980; A.J. Pellionisz, 1984).

As developed most notably by Einstein (Einstein *et al*, 1923), Riemannian geometry has been firmly established as a necessary component of the foundations of physics. There are still two viable camps of opinions, considering this geometry itself as a basic foundation (Misner *et al*, 1973), or considering the physical entities on which its transformations operate as the basic foundation (Weinberg, 1972). However, there is unanimous agreement that Riemannian geometry is an essential theoretical construct to explain some observed physical phenomena. The existence of Riemannian geometry also is a natural mathematical consequence of properties possessed by quite general stochastic systems, including those models of neural systems assumed or endorsed by most investigators (Arbib and Amari, 1985). These properties have been stressed in the SMNI series of papers.

It is the purpose here to stress these general properties, and to make the short but important observation that there is indeed mathematical support on which to conjecture possible neural mechanisms that might exist as a result of invariance under Riemannian-geometric transformations. This observation then leads us back to the spirit, if not the essence, of other neuroscience investigators. However, whereas they *hypothesize* a Riemannian metric *between* cortical regions, SMNI *derives* a Riemannian metric *within* each cortical region or *within* an aggregate set of interacting cortical regions, quite a physical distinction.

Corresponding to the differential-operator ordering problem in the Fokker-Plank equation is the discretization problem in the path integral and in the Langevin rate equations, both of which are equivalent mathematical representations of the Fokker-Plank equation (Dekker, 1980; Grabert and Green, 1979; Graham, 1977a; Graham, 1977b; Langouche *et al*, 1982; Schulman, 1981). An overview of these equations is required to at least note where the Riemannian geometry enters. The appendix of an SMNI paper provides a brief derivation (Ingber, 1991b).

It must be emphasized that the output need not be confined to complex algebraic forms or tables of numbers. Because $\underline{L}_F$ possesses a variational principle, sets of contour graphs, at different long-time epochs of the path-integral of P, integrated over all its variables at all intermediate times, give a visually intuitive and accurate decision aid to view the dynamic evolution of the scenario. For example, this Lagrangian approach permits a quantitative assessment of concepts usually only loosely defined.

Concept	Lagrangian equivalent
Momentum	$\Pi^G = \dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)}$
Mass	$g_{GG'} = \dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)\partial(\partial M^{G'}/\partial t)}$
Force	$\dfrac{\partial \underline{L}_F}{\partial M^G}$
$F = ma$	$\delta \underline{L}_F = 0 = \dfrac{\partial \underline{L}_F}{\partial M^G} - \dfrac{\partial}{\partial t} \dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)}$

(14.28)

These physical entities provide another form of intuitive, but quantitatively precise, presentation of these analyses (Ingber, 1991a).

The key issue is that Riemannian geometry is not required to derive the mathematics of multiplicative Gaussian-Markovian systems. More interestingly, after this derivation, it can be demonstrated that the space of random variables actually induces a Riemannian geometry, obtained explicitly by simply (in hindsight) reorganizing terms in their defining equations. Then, the differential and path-integral representations can be rewritten only in terms of functions $f(M)$ of random variables M that are tensor invariant under quite generally nonlinear point transformations, i.e., $M' = M'(M)$.

The derived probability distribution also is invariant under an equivalence class of discretizations. This is not the same as incorporating *bona fide* physical delays, e.g., those that can give rise to EEG wave propagation, in local circuits as emphasized in SMNI, or in long-ranged circuits (Nunez, 1981a).

The possibility of rewriting any theory or model of neural systems, which can be described by multiplicative Gaussian-Markovian dynamics, into an algebraic form invariant under Riemannian-geometric transformations, does not require that neural systems develop or elect mechanisms to take advantage of these transformations. However, the most obvious candidate for a physical consequence of invariance under such transformations is the information Υ, developed in SMNI, sometimes loosely referred to as the "entropy" of the system. The invariance of Υ implies that, although different cortical regions may have different anatomical features and superficially appear to have quite different sets of firing states, they may indeed share, encode, or decode the same information using their own specific anatomy and physiology to develop their own sets of firing states, related to each other by (nonlinear) transformations.

This possibility is in the original spirit of other authors, who were attracted to the use of Riemannian geometry to explain how information in sensory regions might be transmitted to motor regions, albeit that their neural properties differ in many respects. SMNI develops columnar interactions, and here too it has been tempting to conjecture that local and global processing of columnar firing patterns is enhanced, if not primarily effected, by transmitting blocks of information that are invariant under nonlinear transformations of firing states.

Ultimately, these issues must be decided by experiment. There is presently no evidence, pro or con, to bear on the issue of the explicit Riemannian-geometric nature of information processing of neural firings. In principle, this could be accomplished by numerically fitting neuronal firing data to Lagrangians describing regions behaviorally proven to be processing similar information, similar to fits to data proposed for other artificial intelligence systems (Ingber, 1985b).

3.4.3. *Information, potential, and long-ranged interactions*

There have been attempts to use information as an index of EEG activity (Gersch, 1987; Mars and Lopes da Silva, 1987). However, these attempts have focused on the concept of "mutual information" to find correlations of EEG activity under different electrodes. The SMNI approach at the outset recognizes that, for most brain states of late latency, at least a subset of regions being measured by several electrodes is indeed to be considered as one system, and their interactions are to be explicated by mathematical or physical modeling of the underlying neuronal processes. Then, it is not relevant to compare joint distributions over a set of electrodes with marginal distributions over individual electrodes. The concept of information, as expressed below, may yet prove to be a useful valid measure to compare different subjects within certain categories.

With reference to a steady state $\bar{P}(\tilde{M})$, when it exists, an analytic definition of the information gain $\hat{\Upsilon}$ in state $\tilde{P}(\tilde{M})$ is defined by (Graham, 1978; Haken, 1983)

$$\hat{\Upsilon}[\tilde{P}] = \int \cdots \int \underset{\sim}{D}\tilde{M}\ \tilde{P} \ln(\tilde{P}/\bar{P}), \tag{14.29}$$

where again a path integral is defined such that all intermediate-time values of $\tilde{M}$ appearing in the folded short-time distributions $\tilde{P}$ are integrated over. This is quite general for any system that can be described as Gaussian-Markovian (Haken, 1988), even if only in the short-time limit, e.g., the SMNI theory. (As time evolves, the distribution likely no longer behaves in a Gaussian manner, and the apparent simplicity of the short-time distribution typically must be supplanted by numerical calculations.) Although $\hat{\Upsilon}$ is well defined and useful for discussing macroscopic neocortical activity, it may not be as useful for all applications. Certainly many important local changes of information effected by the neocortical system are a function of the microscopic degrees of freedom already averaged over for the purposes of this study. However, it should also be noted that the path integral represents an enormous number of spatial-temporal degrees of freedom of the mesoscopic system. For example, even neglecting specific coding of presynaptic and postsynaptic membranes, detailed neuronal circuitry, and the dynamics of temporal evolution, in a hypothetical region of 10^9 neurons with 10^{13} synapses: considering each synapse as only conducting or not conducting, there are $\approx \exp(7 \times 10^{12})$ possible synaptic combinations; considering only each neuron as firing or not firing, there are $\approx \exp(7 \times 10^8)$ neuronal combinations; considering only each mesocolumn as having integral firings between -100 and 100, there are $\approx \exp(5 \times 10^7)$ mesocolumnar combinations.

$\underline{T}$, defined as the "kinetic-energy" $\underline{V}$-independent part of $\underline{L}$ in Eq. (14.22), is scale independent of N. Therefore, the small scale of the neocortical system, about which the system fluctuates, is derived to be N^{-1}, the inverse of the number of neurons in a mesocolumn. This is interpreted as the effective "temperature" or inherent noise of the system. Thus, STM defines a rather "hot" and volatile system, wherein the relevant activity takes place on the same order as N^{-1}. By contrast, some long-term-memory calculations described below (Ingber, 1983b) are consistent with the interpretation of transpiring at a much lower temperature taking place in a locally more stable environment.

4. VERIFICATION OF SMNI

It is relevant to the use of the SMNI model to fit experimental data that there is some justification for considering this approach a viable description of the real neocortex. As outlined below, the mesoscopic scale developed by SMNI has been verified by calculations detailing some properties of short-term memory. (Of course, this does not preclude some other theories from possessing similar numerical verification.) The macroscopic scale developed by SMNI has been verified by calculations detailing some properties of phenomena associated with EEG measurements. The aggregate calculations suggest mechanisms for long-term storage of memories via synaptic modifications.

4.1. Short-term memory

4.1.1. *General description*

The most detailed and dramatic application of the theory outlined here is to predict stochastic bounds for the phenomena of human STM capacity during focused selective attention (Ingber, 1972; Ingber, 1981b; Ingber, 1984b; Ingber, 1985a; Ingber, 1985d), transpiring on the order of tenths of a second to seconds, limited to the retention of 7 ± 2 items (Miller, 1956). This is true even for apparently exceptional memory performers who, while they may be capable of more efficient encoding and retrieval of STM, and while they may be more efficient in "chunking" larger patterns of information into single items, nevertheless they also are limited to a STM capacity of 7 ± 2 items (Ericsson and Chase, 1982). This "rule" is verified for acoustical STM, but for visual or semantic STM, which typically require longer times for rehearsal in an hypothesized articulatory loop of individual items, STM capacity appears to be limited to two to four (Zhang and Simon, 1985). This STM capacity-limited chunking phenomenon also has been noted with items requiring varying depths and breadths of processing (Ingber, 1972; Ingber, 1976; Ingber, 1981a; Ingber, 1981b; Ingber, 1985a). Another interesting phenomenon of STM capacity explained by this theory is the primacy versus recency effect in STM serial processing, wherein first-learned items are recalled most error-free, with last-learned items still more error-free than those in the middle (Murdock, 1983).

The basic assumption being made is that a pattern of neuronal firing that persists for many τ cycles is a candidate to store the "memory" of activity that gave rise to this pattern. If several firing patterns can simultaneously exist, then there is the capability of storing several memories. The short-time probability distribution derived for the neocortex is the primary tool to seek such firing patterns. Since this distribution is exponentially sensitive to (minus) the Lagrangian function $\underline{L}$, it is more convenient to deal directly with $\underline{L}$, whereby its minima specify the most likely states that can be sustained at a given time. Then, several important features of these patterned states can be investigated, as is done for other physical systems (Haken, 1983), e.g., the evolution of these states, the "time of first passage" to jump from one state to another state, hysteresis between states that have different depths (values of the Lagrangian at these local minima), the stability of each state under external forces, etc.

We define the "stationary" (sometimes referred to as the "uniform") Lagrangian, $\bar{\underline{L}}$, by setting to zero all temporal and spatial derivatives of M^G, e.g., as appearing in Eq. (14.22). Contour plots of the stationary Lagrangian, $\bar{\underline{L}}$, for typical synaptic parameters balanced between predominately inhibitory and predominately excitatory firing states are examined at many scales when the background synaptic noise is only modestly shifted to cause both efferent and afferent mesocolumnar firing states to have a common most-probable firing, centered at $M*^G = M^G = 0$ (Ingber, 1984b). Within the range of synaptic parameters considered, for values of $\tau\bar{\underline{L}}$~10^{-2}, this "centering" mechanism causes the appearance of from 5 to 10 or 11 extrema for values of $\tau\bar{\underline{L}}$ on the order of ~ 10^{-2}. The centering mechanism is achieved by modestly shifting $B^G_{G'}$ to cause $(V^G - a^{|G|}_{G'} v^{|G|}_{G'} N^{G'})$ to go to zero, thereby driving the threshold factor F^G to zero. (Note that at $F^G = 0$, the mesoscopic derivation of Gaussian $\underline{L}$ breaks down, so that we can only consider a finite region, heavily weighted by N, about this point.) In the absence of external constraints and this centering mechanism, no stable minima are found; i.e., the system either shuts down, with no firings, or it becomes epileptic, with maximal firings at the upper limits of excitatory or of excitatory and inhibitory firings. The appearance of these extrema due to the centering mechanism is clearly dependent on the nonlinearities present in the derived Lagrangian, stressing competition and cooperation among excitatory and inhibitory interactions at columnar as well as at neuronal scales. (See Fig. 14-4.)

Figure 14-4.

It must be stressed that these numbers of minima are determined when the resolution of the contours is commensurate with the resolution of columnar firings, i.e., on the order of five to ten neuronal firings per columnar mesh point. If the resolution is forced to go down to one neuronal firing per columnar mesh point, then typically only about half these minima are found. The coarser resolution, in fact, is the one appropriate for numerical solution of the derived time-dependent path integral: Most important contributions to the probability distribution P come from ranges of the time slice θ and the "action" $N\underline{L}$, such that $\theta N\underline{L} \le 1$. By considering the contributions to the first and second moments of ΔM^G for small time slices θ, conditions on the time and variable meshes can be derived (Wehner and Wolfer, 1983a). The time slice is determined by $\theta \le (N\bar{\underline{L}})^{-1}$ throughout the ranges of M^G giving the most important contributions to the probability distribution P. The variable mesh, a function of M^G, is optimally chosen such that ΔM^G is measured by the covariance $g^{GG'}$ (diagonal in the neocortex due to the independence of E and I chemical interactions) or $\Delta M^G \sim (g^{GG}\theta)^{1/2}$ in the notation of the SMNI papers. For $N \sim 10^2$ and $\bar{\underline{L}} \sim 10^{-2}/\tau$, it is reasonable to pick $\theta \sim \tau$. Then it is calculated that optimal meshes are $\Delta M^E \sim 7$ and $\Delta M^I \sim 4$, essentially the resolutions used in the published SMNI coarse contour plots.

Since the extrema of the Lagrangian appear to lie fairly well along a line in the two-dimensional M^G space, and since coefficients of slowly varying dM^G/dt terms in the nonstationary $\underline{L}$ are noted to be small perturbations on $\bar{\underline{L}}$ (Ingber, 1983b), a solution to the stationary probability distribution was hypothesized to be proportional to $\exp(-\Phi/D)$, where $\Phi = CN^2\bar{\underline{L}}$, the diffusion $D = N/\tau$, and C is a constant. Surprisingly, at least until more recent research has shown the generality of such results (Colet *et al*, 1989), along the line of the extrema, for $C \approx 1$, this is determined to be an accurate solution to the full two-dimensional Fokker-Planck equation (Ingber, 1985d). A weak-noise high-barrier regime defined by $\Delta\Phi/D > 1$, where $\Delta\Phi$ is the difference in Φ from minima to maxima, can be assumed for further analyses (Shenoy and Agarwal, 1984). This is extremely useful, as a linear stability analysis shows that stability with respect to mesocolumnar fluctuations induced by several neurons changing their firings is determined by the second derivatives of $-\Phi$ (Agarwal and Shenoy, 1981); here this just measures the parabolic curvature of $\bar{\underline{L}}$ at the extrema. Thus, all the extrema of the stationary Lagrangian are determined to be stable minima of the time-dependent dynamic system. Note however, that it is unlikely that a true potential exists over all M^G space (Graham, Roekaerts, and Tél, 1985; Graham and Tél, 1984).

This stationary solution is also useful for calculating the time of first passage, t_{vp}, to fluctuate out of a valley in one minimum over a peak to another minimum (Agarwal and Shenoy, 1981). It turns out that the values of $\tau\bar{\underline{L}} \sim 10^{-2}$ for which the minima exist are just right to give t_{vp} on the order of tenths a second for about nine of the minima when the maximum of 10 to 11 are present. The other minima give t_{vp} on the order of many seconds, which is large enough to cause hysteresis to dominate single jumps between other minima (Ingber, 1984b). Thus, 7 ± 2 is the capacity of STM, for memories or new patterns that can be accessed in any order during tenths of a second, all as observed experimentally (Ericsson and Chase, 1982). When the number of neurons per minicolumn is taken to be ~220, modeling the visual neocortex (Ingber, 1984b), then the minima become deeper and sharper, consistent with sharper depth of processing, but several minima become isolated from the main group. This effect might be responsible for the lowering of STM capacity for visual processing, as mentioned above. I.e., the statistical time of passage between clusters becomes many hours, longer than STM, while the time between minima within a cluster, now with only 2 to 4 minima per cluster, is on the order of tenths of a second, as observed. This effect also serves to illustrate that the "practical" number of emergent mesoscopic stable states does not necessarily increase with an increasing number of microscopic units. (See Fig. 14-5.)

Figure 14-5.

This estimate of the number of minima involves a very sensitive calculation. That is, if N were a factor of 10 larger, or if $\tau \bar{\underline{L}} \sim 0.1$ at the minima, then t_{vp} is on the order of hours instead of seconds, becoming unrealistic for STM durations. Alternatively, if t_{vp} were much smaller, i.e., less than $\sim 5\tau$, this case would be inconsistent with observed time scales necessary for formation of any memory trace (Libet, 1982). In this context, it is noted that the threshold factor of the probability distribution scales as $(N^* N)^{1/2}$, demanding that both the macrocolumnar divergence and minicolumnar convergence of mesocolumnar firings be tested by these calculations.

These results pose serious problems for other models, such as "mean-field" theories or reductionist doctrines. The mean-field approach essentially sets $N = 1$, and N^* is effectively taken by some investigators to be $\sim 10^5$, the size of a macrocolumn, but others even consider it to be as large as $\sim 10^{10}$, the total number of neurons in the neocortex. The reductionist doctrine claims that only circuitries among a few to several neurons are responsible for a specific pattern of neocortical function, and this effectively sets $N \approx N^*$, on the order of a few neurons. It is hard to understand how both the capacity and duration of STM can be explained by these other models, even assuming they were or could be derived with realistic synaptic interactions and correct statistical dynamics.

The statistical nature of this storage and processing also explains the primacy versus recency effect: The deepest minima of the Lagrangian are more likely accessed than the others of this probability distribution, and these valleys are sharper than the others. I.e., they are more readily accessed and sustain their patterns against fluctuations more accurately than the relatively more shallow minima. The more recent memories or newer patterns may be presumed to be those having synaptic parameters more recently tuned and/or more actively rehearsed. Thus, both the nonlinearities and the statistical nature of this theory are tested by STM capacity. These insights have helped to correct the notions of some experimentalists who claimed they could not find this effect in the visual cortex: Their experimental paradigms were testing the visual cortex using rules of auditory capacity (i.e., they were testing 7 ± 2 instead of 4 ± 2) and therefore they were washing out this effect.

These calculations give experimental support to the derivation of the mesoscopic probability distribution, yielding similar algebraic structures of the threshold factors in Eqs. (14.18) and (14.26), illustrating common forms of interactions between their entities, i.e., neurons and columns of neurons, respectively. The nonlinear threshold factors are defined in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges.

4.1.2. *STM calculation*

Three cases of neuronal firings were considered (Ingber, 1984b). That reference is recommended for several figures that exhibit the minima structures algebraically described below. Since STM duration is still long relative to τ, stationary solutions of $\overline{\underline{L}}$, derived from $\underline{L}$ in Eq. (14.24), were investigated to determine how many stable minima, $\ll \bar{M}^G \gg$, may simultaneously exist within this duration. Also, individual mesocolumns were studied, i.e., taking the uniform limit of $\dot{\bar{M}}^G = 0 = \nabla \bar{M}^G$. Although the $\dot{\bar{M}}^G = 0$ limit should only be taken for the midpoint-discretized Lagrangian $\underline{L}_F$, this is a small difference here (Ingber, 1984b).

A model of dominant inhibition describes how minicolumnar firings are suppressed by their neighboring minicolumns. For example, this could be effected by developing nearest-neighbor (NN) mesocolumnar interactions (Ingber, 1983b), but here the averaged effect is established by inhibitory mesocolumns (IC) by setting $A_E^I = A_I^E = 2A_E^E = 0.01 N^*/N$. Since there appears to be relatively little $I - I$ connectivity, set $A_I^I = 0.0001 N^*/N$. The background synaptic noise is taken to be $B_I^E = B_E^I = 2B_E^E = 10B_I^I = 0.002 N^*/N$. As minicolumns are observed to have ~110 neurons (the visual cortex appears to have approximately twice this density) (Mountcastle, 1978), and as there appear to be a predominance of E over I neurons (Nunez, 1981a), 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" $F_{\rm IC}^G$ for this IC model are then

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

$$F^I_{\rm IC} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E - 45.8}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 11.2)^{1/2}} \;. \tag{14.30}$$

In the prepoint-discretized deterministic limit, the threshold factors determine when and how smoothly the "step functions" $\tanh F^G_{\rm IC}$ in $g^G(t)$ change $M^G(t)$ to $M^G(t+\theta)$. $F^I_{\rm IC}$ will cause afferent $\bar{M}^I$ to fire for most of its values, as $\bar{M}^I \sim -N^I \tanh F^I_{\rm IC}$ will be positive for most values of $\bar{M}^G$ in $F^I_{\rm IC}$, which is already weighted heavily with a term -45.8. Looking at $F^E_{\rm IC}$, it is seen that the relatively high positive values of efferent $\bar{M}^I$ require at least moderate values of positive efferent $\bar{M}^E$ to cause firings of afferent $\bar{M}^E$.

The calculations presented here support the contention that the neocortex functions at multiple hierarchies. While specific long-term memory (LTM) information is most likely coded at the microscopic neuronal level, the mesoscopic scale most likely provides the context for multiple most-probable firing patterns which process STM and which facilitate plastic synaptic encoding of LTM (Ingber, 1983b). E.g., $\tau\bar{\underline{L}}$ can range from 0 to values greater than 10^3 (Ingber, 1982; Ingber, 1983b). However, realistic constraints on STM duration dictate that only values of $\tau\bar{\underline{L}} \le 0.04$ are of interest here. Detailed mesoscalar calculations demonstrate that only this range exhibits sufficient nonlinear structure to support STM phenomena.

It is discovered that more minima of $\bar{\underline{L}}$ are created, or "restored," if the numerator of F^G contains terms only in $\bar{M}^G$, tending to center $\bar{\underline{L}}$ about $\bar{M}^G = 0$. Of course, any mechanism producing more as well as deeper minima is statistically favored. However, this particular "centering" mechanism has plausible support: $M^G(t+\tau) = 0$ is the state of afferent firing with highest statistical weight. I.e., there are more combinations of neuronal firings, $\sigma_j = \pm 1$, yielding this state than any other $M^G(t+\tau)$; e.g., $\sim 2^{N^G + 1/2}(\pi N^G)^{-1/2}$ relative to the states $M^G = \pm N^G$. Similarly, $M^{*G}(t)$ is the state of efferent firing with highest statistical weight. Therefore, it is natural to explore mechanisms that favor common highly weighted efferent and afferent firings in ranges consistent with favorable firing threshold factors $F^G \approx 0$.

Detailed calculations demonstrate that either $\bar{\underline{L}}^E$ or $\bar{\underline{L}}^I$ separately typically give rise to more multiple minima, $\approx$10, than permitted by their sum $\bar{\underline{L}}$ at this resolution. This "loss" of minima apparently is an interesting consequence of $E - I$ competition at the mesoscopic scale. On one hand, since $\overline{\underline{L}}^G$ scales as N^G/N for relatively large $\bar{M}^G$, $\overline{\underline{L}}^E$ dominates due to the larger $\bar{M}^E$ in its mean g^E. On the other hand, for relatively small $\bar{M}^G$, g^G typically is small if there are several multiple minima in $\overline{\underline{L}}^G$, since most of the minima are found to cluster about the origin. Therefore, $\overline{\underline{L}}^G$ scales as $(N^G)^{-1}$ from the variances $(g^{GG})^{-1}$, and $\overline{\underline{L}}^I$ dominates for small $\bar{M}^G$.

The centering effect of the IC model of dominant inhibition, labeled here as the IC′ model, 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,

$$B'^G_E = \frac{V^G - (\frac{1}{2} A^G_I + B^G_I) v^G_I N^I - \frac{1}{2} A^G_E v^G_E N^E}{v^G_E N^G} \tag{14.31}$$

for both $G = E$ and $G = I$. This is modified straightforwardly when regional influences from $M^{\ddagger E}$ are included, as in in Eq. (14.26). 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. (If B'^G_E would be negative, this leads to unphysical results in the square-root denominator of F^G. Here, in all examples where this occurs, it is possible to instead find positive B'^G_I to appropriately shift the numerator of F^G.) 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).

By this centering mechanism, $B'^E_E = 1.38$ and $B'^I_I = 15.3$, and $F^G_{\rm IC}$ is transformed to $F^G_{\rm IC'}$,

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

$$F^I_{\rm IC'} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 20.4)^{1/2}} . \tag{14.32}$$

Note that, aside from the enforced vanishing of the constant terms in the numerators of $F^G_{\rm IC'}$, the only other change in $F^G_{\rm IC'}$ relative to $F^G_{\rm IC}$ is to moderately affect the constant terms in the denominators. This increases the number of minima of $\tau \underline{L}_{\mathrm{IC}'}$ to 4. The two minima clustered close to the origin are no longer discernible for $\tau \underline{L}_{\mathrm{IC}'} > 0.03$.

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 $A^E_E = 2A^I_E = 2A^E_I = 0.01 N^*/N$. This yields

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

$$F^I_{\mathrm{EC}} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.05\bar{M}^E + 7.24)^{1/2}} . \qquad (14.33)$$

The negative constant in the numerator of F^I_{EC} inhibits afferent $\bar{M}^I$ firings. Although there is also a negative constant in the numerator of F^E_{EC}, the increased coefficient of $\bar{M}^E$ (relative to its corresponding value in F^E_{IC}), and the fact that $\bar{M}^E$ can range up to $N^E = 80$, readily permits excitatory firings throughout most of the range of $\bar{M}^E$. This permits three minima.

Applying the centering mechanism to EC, $B'^E_I = 10.2$ and $B'^I_I = 8.62$. The net effect in $F^G_{\mathrm{EC}'}$, in addition to removing the constant terms in the numerators of F^G_{EC}, is to change the constant terms in the denominators: 12.3 in F^E_{EC} is changed to 17.2 in $F^E_{\mathrm{EC}'}$, and 7.24 in F^I_{EC} is changed to 12.4 in $F^I_{\mathrm{EC}'}$. Now six prominent minima are possible along a line through $\bar{M}^G = 0$, and two others are at $\bar{M}^G = \pm N^G$. Each pair of minima above and below the $\bar{M}^I = 0$ axis merge into single minima for $\tau \underline{L}_{\mathrm{EC}'} > 0.02$, and these lose resolution for $\tau \underline{L}_{\mathrm{EC}'} > 0.03$.

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

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

$$F^I_{\mathrm{BC}} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.050\bar{M}^E + 7.24)^{1/2}} . \qquad (14.34)$$

Three minima are possible, on the boundaries of $\bar{M}^G$ space.

Applying the centering mechanism to BC, $B'^E_E = 0.438$ and $B'^I_I = 8.62$. The net effect in $F^G_{\mathrm{BC}'}$, in addition to removing the constant terms in the numerators of F^G_{BC}, is to change the constant terms in the denominators: 8.30 in F^E_{BC} is changed to 7.40 in $F^E_{\mathrm{BC}'}$, and 7.24 in F^I_{BC} is changed to 12.4 in $F^I_{\mathrm{BC}'}$. Now ten minima are possible. The nine minima along the diagonal line lose resolution for $\tau \underline{\overline{L}}_{\mathrm{BC}'} > 0.01$ above $\bar{M}^I = 0$ and for $\tau \underline{\overline{L}}_{\mathrm{BC}'} > 0.02$ below $\bar{M}^I = 0$.

The effects of using the full Feynman Lagrangian $\underline{\overline{L}}_F$ were considered, including all the Riemannian and other nonlinear corrections, discussed below. The net effect is to slightly raise the threshold at which minima dissipate, to about $\tau \underline{\overline{L}}_{\mathrm{BC}'} \geq 0.03$, which is relevant for the duration of STM, discussed subsequently. However, note that the minima structure is essentially the same. (See Fig. 14-4.)

To demonstrate that multiple minima are not an effect of nonlinearities induced by the denominators of F^G, the net effect in $\underline{\overline{L}}_{\mathrm{BC}'}$ by dropping the $\bar{M}^G$ terms in the denominators of $F^G_{\mathrm{BC}'}$ is such that the valleys of minima are only slightly increased. However, these denominators are still important contributions derived from synaptic interactions. E.g., even with the $\bar{M}^G$ terms dropped, the denominators contribute factors of ~1/5 to $F^G_{\mathrm{BC}'}$.

If N^* is scaled larger or smaller, this effectively scales $A^G_{G'} = A^{*G}_{G'} N^*/N$ and $B^G_{G'} = B^{*G}_{G'} N^*/N$, disturbing the relatively sensitive balance that permits a few percent of efferent firings to affect their afferents. Then, the number of possible minima is typically reduced to one or two. If N is scaled larger or smaller, the number of minima is altered and the duration of STM is affected. However, for N still in the range of a few hundred, the number of possible minima is not severely reduced. The case $N = 220$, e.g., the visual cortex was considered: For model BC′, the number of prominent minima found is 11, but they form clusters, with higher peaks between clusters than between minima within a cluster. Note that the larger N sharpens the minima and therefore the resolution of visual information processing. (See Fig. 14-5.)

Note that the sharpness of the tanh F^G step-function contribution to the mean firing is sensitive to a factor of $N^{\frac{1}{2}}$ in F^G. Additionally, the strength of coupling between mesocolumns scales as $N^{3/2}$. Thus the neuronal size of mesocolumns directly affects the breadth and depth of the information processing capability of the neocortex. It is interesting to note that the human visual cortex, which may be assumed to require the finest tuning in the neocortex, is unique in having twice the number of neurons per minicolumn than other regions of the neocortex (Mountcastle, 1978).

These results are unchanged qualitatively for modest changes of any neocortical parameters. However, it is reasonable to conjecture that more drastic abnormal changes in the neocortical parameters might severely reduce the number of minima. This conjecture is based on calculations wherein F^G do not possess the relatively sensitive balances allowing a few percent of efferent neurons to control firings in their afferent neurons. In calculations using these unrealistic or abnormal parameters only one or two minima survive.

4.1.3. *STM stability and duration*

The calculation of stability and time of duration in most likely states of firing starts by using the differential-equation Hamiltonian formulation of the path-integral Lagrangian, called the Fokker-Planck equation. For future reference, when EEG's are discussed below in the context of considering a given local minimum, note that the time-dependent differential macroscopic probability distribution $\tilde{P} = \prod_r P$, or "propagator," is found to satisfy a true Fokker-Planck equation, but one with nonlinear drifts and diffusions in the space of E and I firings. The Fokker-Planck equation for the region Ω is

$$\frac{\partial \tilde{P}}{\partial t} \approx \Omega^{-1} \int d^2 r [\frac{1}{2} (g^{GG'} \tilde{P})_{,GG'} - (g^G \tilde{P})_{,G} + N \underline{V}' \tilde{P}] ,$$

$$(\cdots)_{,G} \equiv \partial(\cdots)/\partial M^G . \tag{14.35}$$

The true Fokker-Planck equation is actually more general, e.g., if long-ranged spatial structures are included, where the independent variables M^G are fields which themselves may depend on space and time coordinates. The above equation is derived in the nearest-neighbor approximation from the general equation using functional derivatives (Ingber, 1984b),

$$\partial(\cdots)/\partial M^G \rightarrow \delta(\cdots)/\delta M^G ,$$

$$\delta(\cdots)/\delta M^G = (\cdots)_{,G} - \nabla_i (\cdots)_{,\nabla_i G} + \nabla_i^2 (\cdots)_{,\nabla_i^2 G} , \tag{14.36}$$

where we have used the compacted notation introduced previously (Ingber, 1984b).

An estimate of a stationary solution $P_{\rm stat}$ to the Fokker-Planck differential equation for the probability distribution P of M^G firings for an uncoupled mesocolumn, i.e., $\underline{V}' = 0$, is given by the stationary limit of the short-time propagator,

$$P_{\rm stat} \approx N_{\rm stat} g^{1/2} \exp(-CN\tau \underline{\bar{L}}) ,$$

$$g = \det(g^{GG'})^{-1} \equiv \det(g_{GG'}) = g_{EE} g_{II} , \tag{14.37}$$

where $N_{\rm stat}$ and C are constant factors. An estimate of the approximation made is estimated by seeking values of constants C, such that the stationary Fokker-Planck equation is satisfied exactly. Contour plots of C versus $\bar{M}^G$ demonstrate that there exists real positive C which may only range from ~ 10^{-1} to ~1, for which there exists unbroken contours of C which pass through or at least border the line of minima (Ingber, 1985d). At each point $\bar{M}^G$, this leaves a quadratic equation for C to be solved. Dropping the $g^{1/2}$ factor results in C not being real throughout the domain of $\bar{M}^G$.

Thus we have defined a solution with potential $N^2 \underline{\bar{L}} = \int A \, dM$, drift A, and diffusion N/τ. Stability of transient solutions, defined for $\underline{\delta} M^G$ about a stationary state by

$$\underline{\delta} \dot{M}^G \approx - A_{,G} \underline{\delta} M^G = -N^2 \underline{\bar{L}}_{,GG} \underline{\delta} M^G , \tag{14.38}$$

is therefore equivalent to $\ll \bar{M} \gg$ being a minimum of $\underline{\bar{L}}$.

Since the minima of the Lagrangian lie deep in a valley along a line, the time for first passage, t_{vp}, is estimated in analogy to a one-dimensional system as (Agarwal and Shenoy, 1981)

$$t_{vp} \approx \pi N^{-2} [|\underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_p)| \underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_v)]^{-1/2}$$

$$\times \exp \{ CN\tau [\underline{\bar{L}}(\ll \bar{M} \gg_p) - \underline{\bar{L}}(\ll \bar{M} \gg_v)] \} , \tag{14.39}$$

where $\ll \bar{M} \gg_v$ is the minimum at the valley of $\underline{\bar{L}}$ in question, and $\ll \bar{M} \gg_p$ is the maximum at a peak separating two minima. These equations are reasonable but crude estimates, and future numerical work must be done to detail the extent of their validity.

The exponential factor can be quite large in some instances, and quite small in others. As noted previously (Ingber, 1983b), differences in $\bar{\underline{L}}$ from valleys to peaks are still large relative to the Riemannian correction terms and relative to differential spatial-temporal contributions, thereby permitting this simpler analysis. However, values of $\tau \bar{\underline{L}}$ at maxima separating the far minima may be greater than 1, thereby yielding a very large t_{vp}, typical of many physical systems undergoing hysteresis (Ingber, 1983b). Relaxation times t_r about this stationary state are estimated by $|g^G_{,G}|^{-1}$ (Agarwal and Shenoy, 1981), and are on the order of τ.

It is possible for hysteresis to be highly more probable than simple jump behavior to another firing state. This provides a mechanism whereby an extended temporal firing pattern of information can be processed beyond the time scale of relaxation periods, e.g., reverberation among several local minima. It is to be expected that the effects of long-ranged influences on mesoscopic synaptic parameters, e.g., as calculated in previous SMNI papers (Ingber, 1983b), can create more complex examples of spatial-temporal hysteresis. These sustaining mechanisms may serve to permit other biochemical processes to store information for longer time periods as stable synaptic modifications, e.g., LTM. As detailed previously (Ingber, 1983b), changes in synaptic parameters may duplicate the effects of long-ranged interactions, providing a mechanism whereby columnar firings encode long-range firing constraints. If this encoding of firing patterns can establish itself on short enough time scales, then columnar coding of long-range firings could be a precursor mechanism initiating the centering mechanism above, especially across large regions of the neocortex. Then, there would be a more uniform gradation of mechanism(s) establishing STM and LTM.

However, to address the issue of limited capacity of STM, it is reasonable to require that within time spans of tenths of a second to tens of seconds, simple jumps among minima are more probable than hysteresis. This permits all minima to be readily accessible during STM duration, in any ordering (Ericsson and Chase, 1982), at least more so than if hysteresis were more probable. In agreement with this empirical requirement, as detailed in the previous studies, it is found that $\tau[\bar{\underline{L}}(\ll \bar{M} \gg_p) - \bar{\underline{L}}(\ll \bar{M} \gg_v)] \sim 0.01$—$0.03$ for these models using empirical values for synaptic parameters. Then for $|\tau \bar{\underline{L}}_{,GG'}| \sim 10^{-3}$, $t_{vp} \sim 10\tau$—100τ, on the order of several tenths of a second to a second. Use of the full Feynman Lagrangian $\bar{\underline{L}}_F$ increases t_{vp} slightly. For these relatively short t_{vp} the second inequality above is violated, and simple jumps are more probable than hysteresis, as required for STM.

Under conditions of serial processing, the deeper valleys of $\underline{\bar{L}}$ representing the more likely firing states will be occupied first. In all cases considered here, some valleys are deeper than the others. This implies that the last several items in STM should be harder to encode (learn) and retain, with the possible exception of the last one or two items, which represent the most recent shifting of firing patterns $\bar{M}^G$ to these minima $\ll \bar{M} \gg_v$ of $\underline{\bar{L}}$. These conclusions are consistent with empirical observations, and are obtained independent of any other rehearsal mechanisms that may exist.

Calculations in these models establish that the prefactor most often is $\sim \tau$. However, points close to the corners $\bar{M}^G = \pm(N^E, N^I)$ have much more rapid variations. Therefore, minima at these corners, even when $\tau\underline{\bar{L}}(\ll \bar{M} \gg_p)$~$0.01$—$0.03$, because of their sharp peaks, typically have t_{vp} on the order of tens of seconds to jump to minima clustered on the diagonal. This is within the range where hysteresis is more probable for these minima. Therefore, minima at the corners of $\bar{M}^G$ space most likely do not contribute to STM, bringing the number of available minima down to 7 ± 2 as empirically observed.

4.2. EEG dispersion relations

4.2.1. *General description*

Linear expansions about specific extrema, specified by the Euler-Lagrange variational equations, permit the development of stability analyses and dispersion relations in frequency-wave-number space (Ingber, 1982; Ingber, 1983b; Ingber, 1985c). Of course, such linear expansions are justified only after the nonlinear problem, e.g., such as that encountered for STM, is solved for locations of minima. It is noted in this regard that the corresponding wave propagation velocities pace interactions over several minicolumns, in order to be of magnitude sufficient to permit simultaneous information processing within $\sim 10^{-1}$ sec with interactions mediated by long-ranged fibers possessing much greater propagation velocities ~600−900 cm/sec (Ingber, 1985c). E.g., detailed auditory and visual processing can feed information to the association cortex where it can be processed simultaneously, possibly giving feedback to the primary sensory regions. The propagation velocities calculated by SMNI, ~1 cm/sec, also are consistent with observed movements of attention (Tsal, 1983) and of hallucinations (Cowan, 1982) across the visual field. This strongly suggests that nearest-neighbor mesocolumnar interactions as developed here are an important mechanism in these movements. These velocities scale strongly with the values of $\bar{M}^G$

minima, increasing with their distance from $\bar{M}^G \sim 0$, the range of maximal firing combinations. This effect remains to be further investigated; the appropriate calculations should test the nearest-neighbor spatial dependence of the SMNI theory.

These mesoscopic dispersion relations also are consistent with global macroscopic dispersion relations derived and fitted to EEG data (Nunez, 1981a), yielding oscillatory solutions consistent with the alpha rhythm, i.e., $\omega \approx 10^2$ sec^{-1}, equivalent to $\nu = \omega/(2\pi) \approx 16$ cps (Hz). This suggests that these complementary local and global theories may be confluent, considered as a joint set of dispersion relations evolving from the deterministic limit of a joint Lagrangian, referred to as the "equations of motion," but linearly simplified in neighborhoods of minima of the stationary Lagrangian.

Other researchers have developed quite different approaches to investigating macroscopic neocortical activity, e.g., stressing that systematics of rhythmic EEG (alpha rhythm, sleep delta, etc.) can be modeled by resonant modes of macroscopic dipole-layered firing patterns of the neocortex (Nunez, 1974; Nunez, 1981a; Nunez, 1981b; Nunez, 1981c), as discussed in Chapter 10. These resonances, in linearized coupled excitatory-inhibitory spatial-temporal integral equations describing dipole-layered sources, give rise to a macroscopic dispersion relation relating firing frequencies to spatial wave vectors, consistent with experimental observations. While many other investigators also accept dipole layers to model EEG activity, at least to the extent of recognizing activity perpendicular to laminae, they also demonstrate that there are respectable candidates for mechanisms that might fundamentally be responsible for macroscopic activity, other than those proposed here which detail synaptic dynamics of mesocolumnar interactions (Amari, 1983; Anninos *et al*, 1983; Basar, 1980; Freeman, 1978; Klemm and Sherry, 1981; Steriade, 1981; Traub and Llinás, 1979). For example, given the present lack of experimental knowledge, it is possible to formulate macroscopic neocortical activity in terms of statistics of either membrane or synaptic microscopic neuronal activities, albeit that these two are obviously dependent on each other (Dudek *et al*, 1980). Therefore, the results of this statistical theory derived earlier (Ingber, 1985c) might be interpreted either as suggesting that mesocolumnar activity instigates macroscopic activity, or rather as suggesting that mesocolumnar activity strongly interacts with ongoing macroscopic activity that is instigated or sustained by other mechanisms.

The two approaches outlined above, i.e., local mesocolumnar versus global non-mesocolumnar, give rise to the important alternative conjectures suggested previously in this paper. Other studies also have proposed that EEG may be due to a combination of short-ranged and long-ranged interactions, which combine to form a single dispersion relation with multiple branches (Nunez, 1989), as discussed in Chapter 10.

It is plausible that studies of the origin of rhythmic EEG will give direct insight into related mechanisms underlying evoked potentials. However, in contrast to the alpha rhythm and other gross EEG phenomena being gauges of general alertness to process information, the time-locked averaged evoked potentials appear to be a gauge of more selective attention to information being processed. Therefore, to derive a plausible picture of the nature of evoked potentials, it is more likely that more details of local interaction among columnar interactions must be included, such as those given below.

The first SMNI approach to scalp EEG assumed that the Euler-Lagrange variational limit of the stochastic Lagrangian was a suitable averaging procedure over masses of neurons contributing to this relatively coarse spatial phenomenon (Ingber, 1985c).

It should be noted that at this point in the development of our ‘bottom-up” description we have overlapped with our initial “top-down” description, and therefore have provided a relatively first-principles approach to better understand these issues. We also show that most likely trajectories of the mesoscopic probability distribution, representing averages over columnar domains, give a description of the systematics of macroscopic EEG in accordance with experimental observations.

4.2.2. *Euler-Lagrange variational equations*

This calculation begins by considering the Lagrangian $\underline{L}_F$, the Feynman midpoint-discretized Lagrangian. The Euler-Lagrange variational equation associated with $\underline{L}_F$ leads to a set of 12 coupled first-order differential equations, with coefficients nonlinear in M^G, in the 12 variables $\{ M^G, \dot{M}^G, \ddot{M}^G, \nabla M^G, \nabla^2 M^G \}$ in $(r; t)$ space. In the neighborhood of extrema $\ll \bar{M}^G \gg$, $\underline{L}_F$ can be expanded as a Ginzburg-Landau polynomial, i.e., in powers of M^E and M^I. To investigate first-order linear oscillatory states, only powers up to 2 in each variable are kept, and from this the variational principle leads to a relatively simple set of coupled linear differential equations with constant coefficients:

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

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

$$(\cdots)_{,\dot{G}:t} = (\cdots)_{,\dot{G}G'}\dot{M}^{G'} + (\cdots)_{,\dot{G}\dot{G}'}\ddot{M}^{G'}\ ,$$

$$\underline{M}^G = M^G - \ll \bar{M}^G \gg\ ,\ \underline{f}^1_E = -\underline{f}^1_I \equiv \underline{f}\ . \tag{14.40}$$

These equations are then Fourier transformed and the resulting dispersion relation is examined to determine for which values of the synaptic parameters and of the normalized wave-number $\underline{\xi}$, the conjugate variable to r, can oscillatory states, $\omega(\underline{\xi})$, persist. E.g., solutions are sought of the form

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

$$\underline{M}^G_{\mathrm{osc}}(r,t) = \int d^2\underline{\xi}\, d\omega\, \hat{\underline{M}}^G_{\mathrm{osc}}(\underline{\xi},\omega)\exp[i(\underline{\xi}\cdot r - \omega t)]\ . \tag{14.41}$$

For instance, a typical example is specified by extrinsic sources $J_E = -2.63$ and $J_I = 4.94$, $N^E = 125$, $N^I = 25$, $V^G = 10$ mV, $A^G_E = 1.75$, $A^G_I = 1.25$, $B^G_{G'} = 0.25$, and $v^G_{G'} = \phi^G_{G'} = 0.1$ mV. The synaptic parameters are within observed ranges (Shepherd, 1979), and the J_G's are just those values required to solve the Euler-Lagrange equations at the selected values of M^G. The global minimum is at $\bar{M}^E = 25$ and $\bar{M}^I = 5$. This set of conditions yields (dispersive) dispersion relations

$$\omega\tau = \pm\{\ -1.86 + 2.38(\xi\rho)^2; -1.25i + 1.51i(\xi\rho)^2\ \}\ , \tag{14.42}$$

where $\xi = |\underline{\xi}|$. The propagation velocity defined by $d\omega/d\xi$ is ~1 cm/sec, taking typical wavenumbers ξ to correspond to macrocolumnar distances ~ 30ρ. Calculated frequencies ω are on the order of EEG frequencies ~ 10^2 sec^{-1}, equivalent to $\nu = \omega/(2\pi)$= 16 cps (Hz). These mesoscopic propagation velocities permit processing over several minicolumns ~ 10^{-1} cm, simultaneous with the processing of mesoscopic interactions over tens of centimeters via association fibers with propagation velocities ~600–900 cm/sec. I.e., both intraregional and interregional information processing can occur within ~ 10^{-1} sec. Note that this propagation velocity is not “slow”: Visual selective attention moves at ~8 msec/deg (Tsal, 1983), which is ~ 1/2 mm/sec, if a macrocolumn of ~mm^2 is assumed to span 180 deg. This suggests that nearest-neighbor interactions play some part in disengaging and orienting selective attention.

5. DIRECT FIT OF SMNI TO EEG

SMNI can give an even more detailed description of EEG than that using the Euler-Lagrange limit of the Lagrangian. The full SMNI probability distribution can be brought to bear on raw stochastic EEG data. A pilot project gave explicit details on how this can be accomplished (Ingber, 1991b). We used the collection of EEG spontaneous and averaged evoked potential (AEP) data from a multi-electrode array under a variety of conditions. We fit data collected under a project sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) project (Holden, 1991; Porjesz and Begleiter, 1990).

5.1. Algebraic development

We take Eq. (14.6) as the basic probability distribution to fit this data. We also take advantage of and extend the results gained for the STM analysis discussed previously, e.g., the nature of the parabolic well of the Lagrangian. Accordingly, we assume a linear relationship (about minima to be fit to data) between the M^G firing states and the measured scalp potential Φ_ν, at a given electrode site ν representing a macroscopic region of neuronal activity:

$$\Phi_\nu - \phi = aM^E + bM^I \ , \tag{14.43}$$

where $\{ \phi, a, b \}$ are constants determined for each electrode site. In the prepoint discretization, the postpoint $M^G(t + \Delta t)$ moments are given by

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

$$= ag^E + bg^I \ ,$$

$$\sigma^2 \equiv < (\Phi_\nu - \phi)^2 > - < \Phi_\nu - \phi >^2 = a^2 g^{EE} + b^2 g^{II} \ , \tag{14.44}$$

where the M^G-space drifts g^G, and diffusions $g^{GG'}$, have been derived above. Note that the macroscopic drifts and diffusions of the Φ's are simply linearly related to the nonlinear mesoscopic drifts and diffusions of the M^G's. For the prepoint $M^G(t)$ firings, we assume the same linear relationship in terms of $\{ \phi, a, b \}$.

The data we are fitting are consistent with invoking the "centering" mechanism discussed above. Therefore, for the prepoint $M^E(t)$ firings, we also take advantage of the parabolic trough derived for the STM Lagrangian, and take

$$M^I(t) = cM^E(t) , \tag{14.45}$$

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 given in Eq. (14.26), each regional influence from electrode site μ acting at electrode site ν, given by afferent firings $M^{\ddagger E}$, is taken as

$$M^{\ddagger E}_{\mu\to\nu} = d_\nu M^E_\mu (t - T_{\mu\to\nu}) , \tag{14.46}$$

where d_ν are constants to be fitted at each electrode site, and $T_{\mu\to\nu}$ is the delay time estimated for inter-electrode signal propagation, based on current anatomical knowledge of the neocortex and of velocities of propagation of action potentials of long-ranged fibers, typically on the order of one to several multiples of τ= 5 msec. Some terms in which d directly affects the shifts of synaptic parameters $B^G_{G'}$ when calculating the centering mechanism also contain long-ranged efficacies (inverse conductivities) $B^{*E}_{E'}$. Therefore, the latter were kept fixed with the other electrical-chemical synaptic parameters during these fits. In future fits, we will experiment taking the T's as parameters.

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

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

$$L_\nu = \frac{1}{2\sigma^2} (\dot{\Phi}_\nu - m)^2 , \tag{14.47}$$

where m and σ have been derived just above. As discussed above in defining macroscopic regions, the probability distribution for all electrodes is taken to be the product of all these distributions:

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

$$L = \sum_{\nu} L_{\nu} \ . \tag{14.48}$$

Note that we are also strongly invoking the current belief in the dipole or nonlinear-string model. The model SMNI, derived for $P[M^G(t+\Delta t)|M^{\bar{G}}(t)]$, is for a macrocolumnar-averaged minicolumn; hence we expect it to be a reasonable approximation to represent a macrocolumn, scaled to its contribution to Φ_ν. Hence we use L to represent this macroscopic regional Lagrangian, scaled from its mesoscopic mesocolumnar counterpart $\underline{L}$. However, the above expression for P_ν uses the dipole assumption to also use this expression to represent several to many macrocolumns present in a region under an electrode of extent on the order of cm: A macrocolumn has a spatial extent of about a millimeter. A scalp electrode has been shown, under extremely favorable experimental circumstances, to have a resolution as small as several millimeters, directly competing with the spatial resolutions under similar circumstances attributed to magnetoencephalography (Cohen *et al*, 1990), but often most data collected under typical conditions represents a resolution more on the order of up to several centimeters, many macrocolumns. Still, it is often argued that typically at least several tens of macrocolumns firing coherently account for the electric potentials measured by one scalp electrode (Nunez, 1990). Then, we are testing this model to see if the potential will scale to a representative region of many macrocolumns. The results presented here seem to confirm that this approximation is in fact quite reasonable.

6. MESOSCOPIC NEURAL NETS

As noted early in this development (Ingber, 1981c; Ingber, 1985b), basing information processing at macroscopic scales on processes at mesoscopic scales leads to new kinds of computational algorithms, such as proposed here. This approach is consistent with the view that many complex nonlinear nonequilibrium systems develop mesoscopic intermediate structures, e.g., to develop a Gaussian-Markovian statistics to maximize the flow of information (Graham, 1978; Haken, 1983). Viewing the neocortex as the prototypical information processor, a mesoscopic-neural-network (MNN) algorithm can be developed to store and predict patterns of activity inherent in many such nonlinear stochastic multivariate systems.

6.1. SMNI-MNN

While the development of nearest-neighbor interactions into a potential term V' was useful to explore EEG dispersion relations (Ingber, 1984b; Ingber, 1985c), for present purposes this is not necessary and, as permitted in the development of SMNI, we simply incorporate nearest-neighbor interactions with firings $M^{\dagger G}$ by again redefining F^G in Eq. (14.26).

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

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

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

$$A^{\ddagger I}_{E} = A^{\ddagger E}_{I} = A^{\ddagger I}_{I} = B^{\ddagger I}_{E} = B^{\ddagger E}_{I} = B^{\ddagger I}_{I} = 0 \ , \ a^{\ddagger E}_{E} = \frac{1}{2} A^{\ddagger E}_{E} + B^{\ddagger E}_{E} . \tag{14.49}$$

This result presents an SMNI-MNN as a set of nodes, each described by a short-time probability distribution interacting with the other nodes. A set of 1000 such nodes represents a macrocolumn, scaled to represent a “dipole” current source (Ingber, 1991b). A circuitry among patches of macrocolumns represents a typical circuit of activity correlated to specific behavioral states as recorded by EEG under specific experimental or clinical conditions. We are applying SMNI-MNN to learn patterns of an individual’s EEG, and then to predict possible future states of brain activities or anomalies, etc. The methodology used is quite general, and will be described generically.

6.2. Generic MNN

We now generalize this SMNI-MNN, generated from Eq. (14.22), 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,

$$\frac{\partial P}{\partial t} = \frac{\partial(-g^G P)}{\partial M^G} + \frac{1}{2} \frac{\partial^2 (g^{GG'} P)}{\partial M^G \partial M^{G'}} . \tag{14.50}$$

The topology, geometry, and connectivity of MNN can of course be generalized. E.g., we need not be restricted to nearest-neighbor interactions, although this is simpler to implement especially on parallel processors. Also, we can include "hidden layers" to increase the complexity of MNN, although the inclusion of nonlinear structure in the drifts g^G and diffusions $g^{GG'}$ may make this unnecessary for many systems. A detailed presentation of using "learning" and "prediction" of information has been given for these systems (Ingber, 1992).

7. CONCLUSION

7.1. Extracting signal from noise

There are several factors in the SMNI development that support optimism for extracting more signal from noise than is currently possible.

7.1.1. *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.

7.1.2. *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.

7.1.3. *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. A new C-language code, PATHINT, has been developed to handle n-dimensional Fokker-Planck-type systems, using the algorithm mentioned previously. 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.

7.1.4. *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.

7.1.5. *Riemannian invariants*

Yet to explore are the ramifications of using the derived (not hypothesized) Riemannian metric induced by multivariate Fokker-Plank-type systems (Ingber, 1982; Ingber, 1983b; Ingber, 1988a). This seems to form a natural invariant measure of information, that could/should be used to explore flows of information between neocortical regions.

7.1.6. *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 of scalp measurement, wherein there is attenuation of ranges of wave numbers (Nunez, 1981a). While SMNI is most logically tested using data collected from brain surface recordings, the necessity and utility of performing non-invasive EEG scalp recordings argues strongly for further developing SMNI to extract better signal out of noise from scalp recordings.

7.2. Importance of modeling at multiple scales

To learn more about complex systems, we inevitably must form functional models to represent huge sets of data. Indeed, modeling phenomena is as much a cornerstone of 20th century science as is collection of empirical data (Jammer, 1974). The ability to fit data to these particular SMNI functional forms goes beyond nonlinear statistical modeling of data. The plausibility of the SMNI model, as emphasized in this and previous SMNI papers, as spanning several important neuroscientific phenomena, suggests that the fitted functional forms may yet help to explicate some underlying biophysical mechanisms responsible for the normal and abnormal behavioral states being investigated, e.g., excitatory and/or inhibitory influences, background electromagnetic influences from nearby firing states (by using SMNI synaptic conductivity parameters in the fits).

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 physics and chemistry 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.

For example, both artificial neural networks (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; only the latter has a chance for any statistical justification. 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 can 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.

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; 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. 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.

7.3. Modelling EEG with SMNI

In order to detail such a model of EEG phenomena we found it useful to seek guidance from "top-down" models; e.g., the nonlinear string model representing nonlinear dipoles of neuronal columnar activity. In order to construct a more detailed "bottom-up" model that could give us reasonable algebraic functions with physical parameters to be fit by data, we then needed to bring together a wealth of empirical data and modern techniques of mathematical physics across multiple scales of neocortical activity. At each of these scales, we had to derive and establish reasonable procedures and submodels for climbing from scale to scale. Each of these submodels could then be tested against some experimental data to see if we were on the right track. For example, at the mesoscopic scale we checked the consistency of SMNI with known aspects of visual and auditory short-term memory; at the macroscopic scale we checked consistency with known aspects of EEG and the propagation of information across the neocortex. Here, we have demonstrated that the currently accepted dipole EEG model can be derived as the Euler-Lagrange equations of an electric-potential Lagrangian.

The theoretical and experimental importance of specific scaling of interactions in the neocortex has been quantitatively demonstrated: We have shown that the explicit algebraic form of the probability distribution for mesoscopic columnar interactions is driven by a nonlinear threshold factor of the same form taken to describe microscopic neuronal interactions, in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges; these threshold factors largely determine the nature of the drifts and diffusions of the system. This mesoscopic probability distribution has successfully described STM phenomena and, when used as a basis to derive the most likely trajectories using the Euler-Lagrange variational equations, it also has described the systematics of EEG phenomena. More recently, we have taken the mesoscopic form of the full probability distribution more seriously for macroscopic interactions, deriving macroscopic drifts and diffusions linearly related to sums of their (nonlinear) mesoscopic counterparts, scaling its variables to describe regional interactions correlated with observed electrical activities measured by electrode recordings of scalp EEG, with apparent success. These results give strong quantitative support for an accurate intuitive picture, portraying neocortical interactions as having common algebraic or physics mechanisms that scale across quite disparate spatial scales and functional or behavioral phenomena, i.e., describing interactions among neurons, columns of neurons, and regional masses of neurons.

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. 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, very fast simulated reannealing (ASA), and parallel-processing computation.

7.4. Work in progress

There is much more work to be done. We have not yet addressed the "renormalization" issues raised, which are based on the nature of EEG data collection and which are amenable to this framework. While the fitting of these distributions certainly compacts the experimental data into a reasonable algebraic model, a prime task of most physical theory, in order to be useful to clinicians (and therefore to give more feedback to theory) even more data reduction must be performed. We must experiment with path-integral calculations and some methods of "scientific visualization" to determine what minimal, or at least small, set of "signatures" might suffice, which would be faithful to the data yet useful to clinicians. We also must examine the gains that might be made by putting these codes onto a parallel processor, which might enable real-time diagnoses based on noninvasive EEG recordings.

It seems reasonable to speculate on the evolutionary desirability of developing Gaussian-Markovian statistics at the mesoscopic columnar scale from microscopic neuronal interactions, and maintaining this type of system up to the macroscopic regional scale. I.e., this permits maximal processing of information (Graham, 1978), an entity which can be codified as invariants in a Riemannian geometry induced by the stochastic system. This concept must be tested against experimental data.

There is much work to be done, but we believe that modern methods of statistical mechanics have helped to point the way to promising approaches.

REFERENCES

Adey, W.R. Long-range electromagnetic field interactions at brain cell surfaces, in *Magnetic Field Effect on Biological Systems*, (Plenum, New York, NY, 1978).

Afifi, A.K. and Bergman, R.A. *Basic Neuroscience* (Urban and Schwarzenberg, Baltimore, MD, 1978).

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

Amari, S.-I. Field theory of self-organizing neural nets, *IEEE Trans. Syst. Man Cybern.* **13**, 741-748 (1983).

Anninos, P., Zenone, S., and Elul, R. Artificial neural nets: Dependence of the EEG amplitude's probability distribution on statistical parameters, *J. Theor. Biol.* **103**, 339-348 (1983).

Arbib, M.A. and Amari, S.-I. Sensori-motor transformations in the brain (with a critique of the tensor theory of cerebellum), *J. Theor. Biol.* **112**, 123-155 (1985).

Babloyantz, A. and Kaczmarek, L.K. Self-organization in biological systems with multiple cellular contacts, *Bull. Math. Biol.* **41**, 193-201 (1979).

Basar, E. *EEG-Brain Dynamics* (Elsevier, Amsterdam, 1980).

Boyd, I.A. and Martin, A.R. The end-plate potential in mammalian muscle, *J. Physiol. (London)* **132**, 74-91 (1956).

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

Bullock, T.H. Reassessment of neural connectivity and its specification, in *Information Processing in the Nervous System*, ed. by H.M. Pinsker and W.D. Willis, Jr. (Raven Press, New York, NY, 1980).

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.* **58**, 1723-1726 (1980).

Cartling, B. From short-time molecular dynamics to long-time stochastic dynamics of proteins, *J. Chem. Phys.* **91**, 427-438 (1989).

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.* **28**, 811-817 (1990).

Colet, P., Wio, H.S., and San Miguel, M. Colored noise: A perspective from a path-integral formalism, *Phys. Rev. A* **39**, 6094-6097 (1989).

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

Cragg, B.G. and Temperley, H.N.V. The organization of neurons: A cooperative analogy, *Electroencephal. clin. Neurophysiol.* **6**, 85-92 (1954).

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

DeWitt, B.S. Dynamical theory in curved spaces. I. A review of the classical and quantum action principles, *Rev. Mod. Phys.* **29**, 377-397 (1957).

Dismukes, R.K. New concepts of molecular communication among neurons, *Behav. Brain Sci.* **2**, 409-448 (1979).

Dudek, F.E., Hatton, G.I., and Macvicar, B.A. Intracellular recordings from the paraventricular nucleus in slices of rat hypothalamus, *J. Physiol.* **301**, 101-114 (1980).

Dykes, R.W. The anatomy and physiology of the somatic sensory cortical regions, *Prog. Neurobiol.* **10**, 33-88 (1978).

Dykes, R.W. Parallel processing of somatosensory information: A theory, *Brain Res. Rev.* **6**, 47-115 (1983).

Einstein, A., Lorentz, A., Minkowski, H., and Weyl, H. *The Principle of Relativity* (Dover, New York, 1923).

Erickson, R.P The across-fiber pattern theory: An organizing principle for molar neural function, *Sensory Physiol.* **6**, 79-110 (1982).

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

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.* **192**, 589-601 (1980).

Freeman, W.J. *Mass Action in the Nervous System* (Academic Press, New York, NY, 1975).

Freeman, W.J. Spatial properties of an EEG event in the olfactory bulb and cortex, *Electroencephal. clin. Neurophysiol.* **44**, 586-605 (1978).

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 *Methods of Brain Electrical and Magnetic Signals. EEG Handbook*, ed. by A.S. Gevins and A. Remond (Elsevier, New York, NY, 1987; p. 261-296).

Gilbert, C.D. and Wiesel, T.N. Functional organization of the visual cortex, *Prog. Brain Res.* **58**, 209-218 (1983).

Goel, N.S. and Richter-Dyn, N. *Stochastic Models in Biology* (Academic Press, New York, NY, 1974).

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.* **122**, 393-413 (1977).

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

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

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

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

Graham, R., Roekaerts, D., and Tél, T. Integrability of Hamiltonians associated with Fokker-Planck equations, *Phys. Rev. A* **31**, 3364-3375 (1985).

Graham, R. and Tél, T. Existence of a potential for dissipative dynamical systems, *Phys. Rev. Lett.* **52**, 9-12 (1984).

Griffith, J.S. *Mathematical Neurobiology* (Academic Press, New York, NY, 1971).

Haken, H. *Synergetics*, 3rd ed. (Springer, New York, 1983).

Haken, H. *Information and Self-Organization: A Macroscopic Approach to Complex Systems* (Springer, Berlin, 1988).

Hebb, D.O. *The Organisation of Behavior* (Wiley, New York, NY, 1949).

Holden, C. Probing the complex genetics of alcoholism, *Science* **251**, 163-164 (1991).

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

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

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

Ingber, L. Editorial: Learning to learn, *Explore* **7**, 5-8 (1972).

Ingber, L. *The Karate Instructor's Handbook* (Physical Studies Institute-Institute for the Study of Attention, Solana Beach, CA, 1976).

Ingber, L. Attention, physics and teaching, *J. Social Biol. Struct.* **4**, 225-235 (1981a).

Ingber, L. *Karate: Kinematics and Dynamics* (Unique, Hollywood, CA, 1981b).

Ingber, L. Towards a unified brain theory, *J. Social Biol. Struct.* **4**, 211-224 (1981c).

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

Ingber, L. Riemannian corrections to velocity-dependent nuclear forces, *Phys. Rev. C* **28**, 2536-2539 (1983a).

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

Ingber, L. Path-integral Riemannian contributions to nuclear Schrödinger equation, *Phys. Rev. D* **29**, 1171-1174 (1984a).

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

Ingber, L. Statistical mechanics of nonlinear nonequilibrium financial markets, *Math. Modelling* **5**, 343-361 (1984c).

Ingber, L. *Elements of Advanced Karate* (Ohara, Burbank, CA, 1985a).

Ingber, L. Statistical mechanics algorithm for response to targets (SMART), in *Workshop on Uncertainty and Probability in Artificial Intelligence: UC Los Angeles, 14-16 August 1985*, ed. by P. Cheeseman (American Association for Artificial Intelligence, Menlo Park, CA, 1985b; p. 258-264).

Ingber, L. Statistical mechanics of neocortical interactions. EEG dispersion relations, *IEEE Trans. Biomed. Eng.* **32**, 91-94 (1985c).

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

Ingber, L. Towards clinical applications of statistical mechanics of neocortical interactions, *Innov. Tech. Biol. Med.* **6**, 753-758 (1985e).

Ingber, L. Riemannian contributions to short-ranged velocity-dependent nucleon-nucleon interactions, *Phys. Rev. D* **33**, 3781-3784 (1986a).

Ingber, L. Statistical mechanics of neocortical interactions, *Bull. Am. Phys. Soc.* **31**, 868 (1986b).

Ingber, L. Applications of biological intelligence to Command, Control and Communications, in *Computer Simulation in Brain Science: Proceedings, University of Copenhagen, 20-22 August 1986*, ed. by R. Cotterill (Cambridge University Press, London, 1988a; p. 513-533).

Ingber, L. Mesoscales in neocortex and in command, control and communications (C^3) systems, in *Systems with Learning and Memory Abilities: Proceedings, University of Paris 15-19 June 1987*, ed. by J. Delacour and J.C.S. Levy (Elsevier, Amsterdam, 1988b; p. 387-409).

Ingber, L. Statistical mechanics of mesoscales in neocortex and in command, control and communications (C^3): Proceedings, Sixth International Conference, St. Louis, MO, 4-7 August 1987, *Mathl. Comput. Modelling* **11**, 457-463 (1988c).

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

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

Ingber, L. Statistical mechanical measures of performance of combat, in *Proceedings of the 1991 Summer Computer Simulation Conference 22-24 July 1991, Baltimore, MD*, ed. by D. Pace (Society for Computer Simulation, San Diego, CA, 1991a; p. 940-945).

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

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

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, (1993a).

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

Ingber, L. Statistical mechanics of combat and extensions, in *Toward a Science of Command, Control, and Communications*, ed. by C. Jones (American Institute of Aeronautics and Astronautics, Washington, D.C., 1993c; p. 117-149).

Ingber, L. Statistical mechanics of neocortical interactions: Multiple scales of EEG, *Electroencephal. clin. Neurophysiol.* (to be published) (1996).

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

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

Ingber, L. and Rosen, B. Very Fast Simulated Reannealing (VFSR), [ringer.cs.utsa.edu: /pub/rosen/vfsr.Z], University of Texas, San Antonio, TX, (1992).

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

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* **15**, 77-98 (1991).

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

John, E.R. Switchboard versus statistical theories of learning and memory, *Science* **177**, 850-864 (1972).

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.* **181**, 291-348 (1978).

Kaczmarek, L.K. and Babloyantz, A. Spatiotemporal patterns in epileptic seizures, *Biol. Cybernetics* **26**, 199-208 (1977).

Katchalsky, A.K., Rowland, V., and Blumenthal, R. Dynamic patterns of brain assemblies, *Neurosci. Res. Prog. Bull.* **12**, 1-187 (1974).

Katz, B. *Nerve, Muscle, and Synapse* (McGraw-Hill, new York, NY, 1966).

Klemm, W.R. and Sherry, C.J. Serial ordering in spike trains: What's it 'trying to tell us'?, *Int. J. Neurosci.* **14**, 15-33 (1981).

Korn, H. and Mallet, A. Transformation of binomial input by the postsynaptic membrane at a central synapse, *Science* **225**, 1157-1159 (1984).

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* **213**, 898-900 (1981).

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

Landau, L.D., Lifshitz, E.M., and Pitaevskii, L.P. *Statistical Physics* (Pergamon, New York, NY, 1980).

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

Libet, B. Brain stimulation in the study of neuronal functions for conscious sensory experience, *Human Neurobiol.* **1**, 235-242 (1982).

Little, W.A. The existence of persistent states in the brain, *Math. Biosci.* **19**, 101-120 (1974).

Little, W.A. and Shaw, G.L. Analytic study of the memory storage capacity of a neural network, *Math. Biosci.* **39**, 281-290 (1978).

Ma, S.-K. *Modern Theory of Critical Phenomena* (Benjamin/Cummings, Reading, MA, 1976).

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 *Methods of Brain Electrical and Magnetic Signals. EEG Handbook*, ed. by A.S. Gevins and A. Remond (Elsevier, New York, NY, 1987; p. 297-307).

Mathews, J. and Walker, R.L. *Mathematical Methods of Physics, 2nd ed.* (Benjamin, New York, NY, 1970).

McCulloch, W.S. and Pitts, W. A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* **12**, 207-219 (1943).

McGeer, P.C., Eccles, J.C., and McGeer, E.G. *Molecular Neurobiology of the Mammalian Brain* (Plenum, New York, NY, 1978).

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

Misner, C.W., Thorne, K.S., and Wheeler, J.A. *Gravitation* (Freeman, San Francisco, 1973).

Mountcastle, V.B. An organizing principle for cerebral function: The unit module and the distributed system, in *The Mindful Brain*, ed. by G.M. Edelman and V.B. Mountcastle (Massachusetts Institute of Technology, Cambridge, 1978; p. 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.* **1**, 1218-1235

(1981).

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

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

Nunez, P.L. The brain wave equation: A model for the EEG, *Math. Biosci.* **21**, 279-297 (1974).

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

Nunez, P.L. A study of origins of the time dependencies of scalp EEG: II—Experimental support of theory, *IEEE Trans. Biomed. Eng.* **28**, 281-288 (1981b).

Nunez, P.L. A study of origins of the time dependencies of scalp EEG: I—Theoretical basis, *IEEE Trans. Biomed. Eng.* **28**, 271-280 (1981c).

Nunez, P.L. Generation of human EEG rhythms by a combination of long and short-range neocortical interactions, *Brain Topography* **1**, 199-215 (1989).

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

Pellionisz, A. and Llinás, R. Brain modeling by tensor network theory and computer simulation. The cerebellum: distributed processor for predictive coordination, *Neuroscience* **4**, 323-348 (1979).

Pellionisz, A. and Llinás, R. Tensorial approach to the geometry of brain function: cerebellar coordination via a metric tensor, *Neuroscience* **5**, 1125-1136 (1980).

Pellionisz, A.J. Coordination: a vector-matrix description of transformations of overcomplete CNS coordinates and a tensorial solution using the Moore-Penrose generalized inverse, *J. Theor. Biol.* **110**, 353-375 (1984).

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

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

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* **2**, 99-118 (1989).

Schmitt, F.O., Dev, P., and Smith, B.H. Electrotonic processing of information by brain cells, *Science* **193**, 114-120 (1976).

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.* **47**, 487-533 (1975).

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

Shenoy, S.R. and Agarwal, G.S. First-passage times and hysteresis in multivariable stochastic processes: The two-mode ring laser, *Phys. Rev. A* **29**, 1315-1325 (1984).

Shepherd, G.M. *The Synaptic Organization of the Brain*, 2nd ed. (Oxford University Press, New York, NY, 1979).

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

Steriade, M. Mechanisms underlying cortical activation: Neuronal organization and properties of the midbrain reticular core and intralaminar thalamic nuclei, in *Brain Mechanisms and Perceptual Awareness*, ed. by O. Pompeiano and C.A. Marsan (Raven, New York, 1981; p. 327-377).

Szentágothai, J. The 'module-concept' in cerebral cortex architecture, *Brain Res.* **95**, 475-496 (1975).

Szentágothai, J. The neural network of the cerebral cortex: A functional interpretation, *Proc. Roy. Soc. Lon.* **201**, 219-248 (1978).

Szentágothai, J. and Arbib, M.A. Conceptual models of neural organization, *Neurosci. Res. Bull.* **12**, 307-510 (1974).

Takeuchi, A. and Amari, S. Formation of topographic maps and columnar microstructures in nerve fields, *Biol. Cybernetics* **35**, 63-72 (1979).

Traub, R.D. and Llinás, R. Hippocampal pyramidal cells: Significance of dendritic ionic conductances for neuronal function and epileptogenesis, *J. Neurophysiol.* **42**, 476-496 (1979).

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

van Kampen, N.G. Fluctuations in closed and open non-linear systems, in *Statistical Physics*, ed. by L. Pál and P. Szépfalusy (North-Holland, Amsterdam, The Netherlands, 1976; p. 29-48).

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* **41**, 153-164 (1981).

von der Marlsburg, Ch. Development of ocularity domains and growth behavior of axon terminals, *Biol. Cybernetics* **32**, 49-62 (1979).

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

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

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

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* **35**, 1795-1801 (1987).

Weidlich, W. and Haag, G. *Concepts and Models of a Quantitative Sociology* (Springer, Berlin, 1983).

Weinberg, S. *Gravitational Cosmology* (Wiley, New York, 1972).

Williamson, S.J., Kaufman, L., and Brenner, D. Evoked neuromagnetic fields of the human brain, *J. Appl. Phys.* **50**, 2418-2421 (1979).

Wilson, H.R. Hysteresis in binocular grating perception: Contrast effects, *Vision Res.* **17**, 843-851 (1977).

Wilson, H.R. and Cowan, J.D. Excitatory and inhibitory interactions in localized populations of model neurons, *Biophys. J.* **12**, 1-23 (1972).

Wilson, H.R. and Cowan, J.D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue, *Kybernetik* **13**, 55-80 (1973).

Wilson, K.G. Problems in physics with many scales of length, *Sci. Am.* **241**, 158-179 (1979).

Wilson, K.G. and Kogurt, J. The renormalization group and the ϵ expansion, *Phys. Rep. C* **12**, 75-200 (1974).

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

LESTER INGBER

Professor Lester Ingber was born in Brooklyn, New York on 26 March 1941. He received: his diploma from Brooklyn Technical High School in 1958; his B.S. in physics from Caltech in 1962; his Ph.D. in theoretical nuclear physics from UC San Diego in 1966, having studied at the Niels Bohr Institute in 1964, and having been a consultant at RAND in 1965-1966. He was a National Science Foundation Postdoctoral Fellow at UC Berkeley in 1967-1968 and at UC Los Angeles in 1968-1969, an Assistant Professor in physics at SUNY at Stony Brook from 1969-1970, and a research physicist in the Physics department and in the Institute for Pure and Applied Physical Sciences (IPAPS) at UC San Diego from 1970-1972. From 1970-1986 he was President of Physical Studies Institute (PSI), a nonprofit corporation he founded in 1970, which was an agency account in IPAPS from 1980-1986. He was a Research Associate at UC San Diego in the Music department from 1972-1974 and in IPAPS from 1980-1986. He was awarded a Senior Research Associateship for 1985-1986 by the National Research Council (NRC) of the National Academy of Sciences, taken at the Naval Postgraduate School (NPS) in Monterey, CA. From 1986-1989 he was Professor of Physics at NPS at a GS-15 Step 10 equivalent position. In March 1988 he was officially offered an ES-4 Senior Executive Service (SES) appointment as Assistant Director, Washington Operations, of the Joint Tactical C^3 (command, control and communications) Agency (JTC3A); he declined in order to complete his projects. From February through June 1989 he was on extended temporary duty at US Army Concepts Analysis Agency (CAA) in Bethesda, MD. In 1989 He won a second NRC Senior Research Associateship, taken at the Naval Ocean Systems Center (NOSC) in San Diego. From 1989-1990 he was Research Professor of Mathematics at The George Washington University (GWU), D.C.

He has published over sixty papers and books, approximately a dozen in each category of: nuclear physics, education and karate, combat analysis, neuroscience, and general optimization and finance. His publications in neuroscience in 1983 and in finance in 1990 were the first papers accepted on these subjects in the conservative premier physics journal, *Physical Review*. He is the creator of Adaptive Simulated Annealing (ASA), one of the most powerful optimization algorithms for nonlinear and stochastic systems, originally published as Very Fast Simulated Reannealing (VFSR). He has devoted over thirty years and his personal resources to motivating and developing children and adults, in disciplines spanning karate, academics, and technical projects.

From 1958-1988 he founded and instructed karate classes at: Caltech, UC Berkeley, UC Los Angeles, SU New York at Stony Brook, UC San Diego, PSI, and NPS. He has developed and published in several textbooks techniques promoting the learning of attentional skills in parallel with the learning of traditional physical skills. He received his black belt in karate in 1961, and became the first Westerner to receive the Instructor's (Sensei) degree from the Japan Karate Association in 1968. From 1989-1991 he was Director of Scientific Studies of the American JKA Karate Associations (AJKA). Now he is a 7th Dan karate Sensei. From 1970-1972 he developed teaching methodologies for academics and fine arts, instructing in and administrating a six-course program through UC San Diego Extension. From 1972-1978 he founded, funded, and instructed in an experimental alternative high school offering 30+ courses in academics, fine arts, and physical disciplines.

From 1965-1972 he published papers in atomic, nuclear, astro-, and elementary particle physics. His major work was to develop a nucleon-nucleon interaction described by exchanged mesons, and to apply this interaction to calculate properties of nucleon-nucleon scattering, the deuteron, nuclear matter, and neutron stars. In 1983-1986 he used modern methods of nonlinear functional analysis to discover contributions induced by velocity-dependent potentials to nuclear matter binding energies. Since 1978 he has developed a statistical mechanics of neocortical interactions applicable to a broad range of spatial and temporal scales, using modern methods of nonlinear nonequilibrium statistical mechanics to calculate brain 'observables' from neuronal dynamics. A present project is extending this theory to develop mathematical signatures of behavioral states correlated to EEG and evoked potential data. From 1986-1989 he applied these methods of mathematical physics to problems of governmental concern. As principal investigator (PI) of an Army contract, he led a team of scientists and officers to develop algorithms for mathematical comparisons of Janus computer combat simulations with exercise data from the National Training Center (NTC), developing a testable theory of combat successfully baselined to empirical data. These algorithms are being applied to financial instruments.

FIGURE CAPTIONS

FIG. 14-1. *Scales illustrated*. Illustrated are three biophysical scales of neocortical interactions: (a), (a^*), (a') microscopic neurons; (b) and (b') mesocolumnar domains; (c) and (c') macroscopic regions. In (a^*) synaptic interneuronal interactions, averaged over by mesocolumns, are phenomenologically described by the mean and variance of a distribution Ψ. 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 of the neocortex are depicted 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.

FIG. 14-2. *Minima structure of nonlinear Lagrangian*. Examination of the minima structure of the spatially averaged and temporally averaged Lagrangian provides some quick intuitive details about the most likely states of the system. This is supported by further analysis detailing the actual spatial-temporal minima structure. Illustrated is the surface of the static (time-independent) mesoscopic neocortical Lagrangian $\bar{\underline{L}}$ 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.

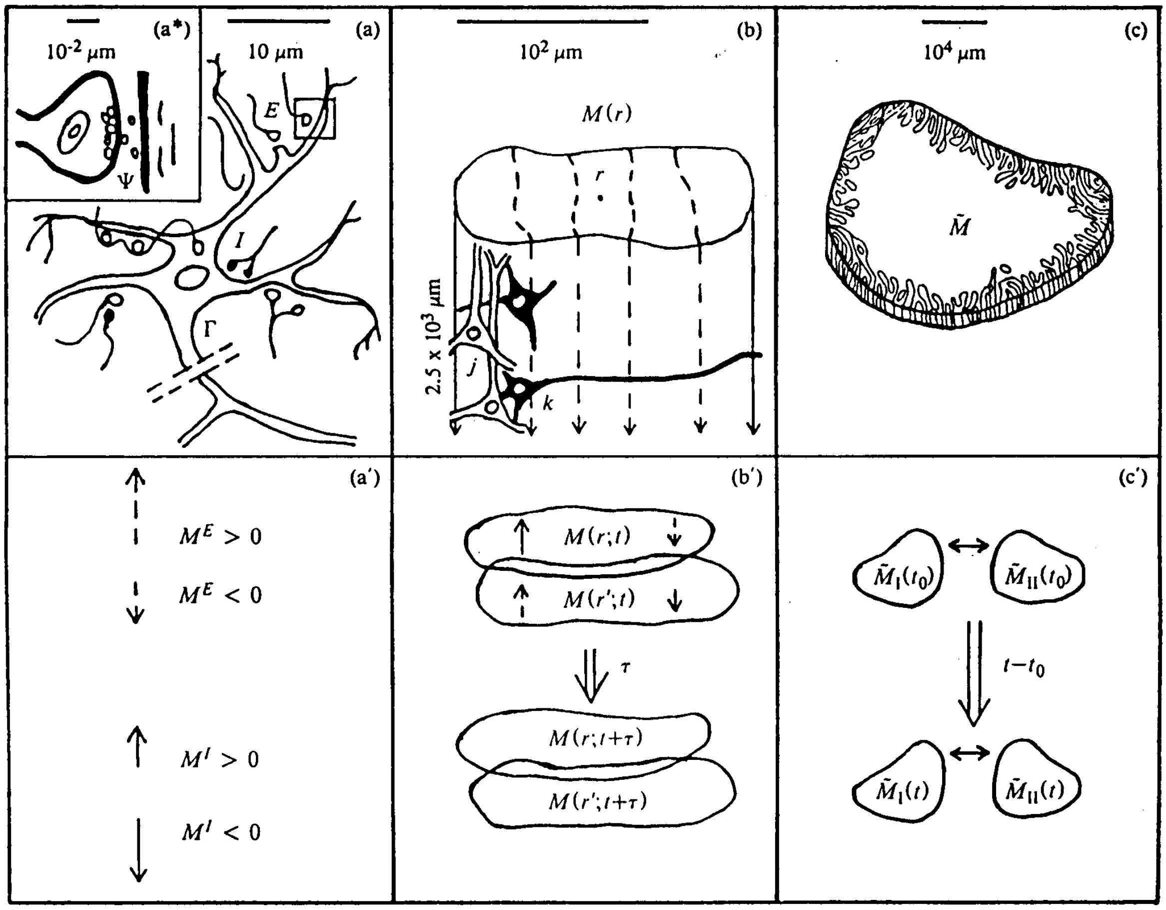
FIG. 14-3. *Nearest neighbors*. Nearest-neighbor interactions between mesocolumns are illustrated. Afferent minicolumns of $\sim 10^2$ neurons are represented by the inner circles, and efferent macrocolumns of $\sim 10^5$ neurons by the outer circles. Illustrated are the nearest-neighbor interactions 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. I.e., these neurons outside the macrocolumnar area of influence of the central minicolumn are contacted through interactions with neurons in the central macrocolumn.

FIG. 14-4. *Valleys of STM*. Contours of the Lagrangian illustrate "valleys" that trap firing-states of mesocolumns. ($\tau \bar{\underline{L}}_{\mathrm{BC}}$ can be as large as 10^3.) These valleys are candidates for short-term memories. Detailed calculations support the identification of the inner valleys with stable short-term-memory states having durations on the order of tenths of a second. (a) Contours for values less than 0.04 are drawn for

$\tau \underline{\overline{L}}_{\mathrm{BC}}$, where BC designates the balanced case of firing states being at a moderate level of excitatory and inhibitory firings. 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. Note the absence of any closed contours in the interior space. (b) Contours for values less than 0.04 are drawn for $\tau \underline{\overline{L}}_{\mathrm{BC}'}$, where BC′ designates that the "centering mechanism" has been turned on. A right brace } signifies enclosure of other nested closed contours above and to the left of this brace.

FIG. 14-5. *Modeling the visual cortex STM*. When $N = 220$, modeling the number of neurons per minicolumn in the visual neocortex, then only clusters containing 2−4 up to 5−6 minima are found, consistent with visual STM. These minima are narrower, consistent with the sharpness required to store visual patterns.

Figure 14-1



(a*)
10^{-2} μm
Ψ
(a)
10 μm
E
I
Γ
(b)
10^2 μm
M(r)
r
2.5×10^3 μm
j
k
(c)
10^4 μm
$\tilde{M}$
(a′)
$M^E > 0$
$M^E < 0$
$M^I > 0$
$M^I < 0$
(b′)
M(r;t)
M(r′;t)
τ
M(r;t+τ)
M(r′;t+τ)
(c′)
$\tilde{M}_I(t_0)$
$\tilde{M}_{II}(t_0)$
$t-t_0$
$\tilde{M}_I(t)$
$\tilde{M}_{II}(t)$

Figure 14-2

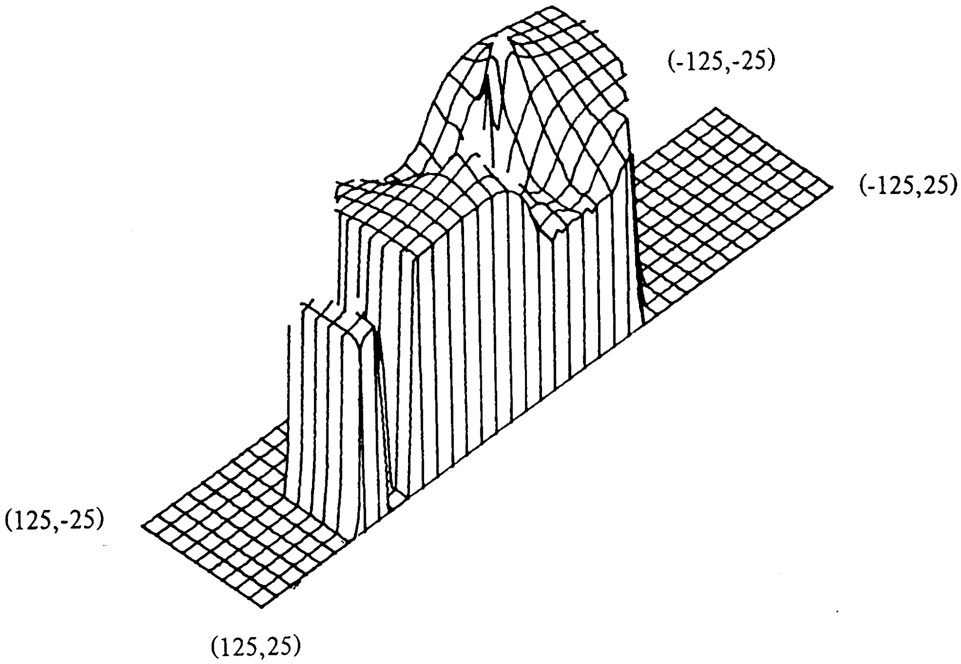


Figure 14-3

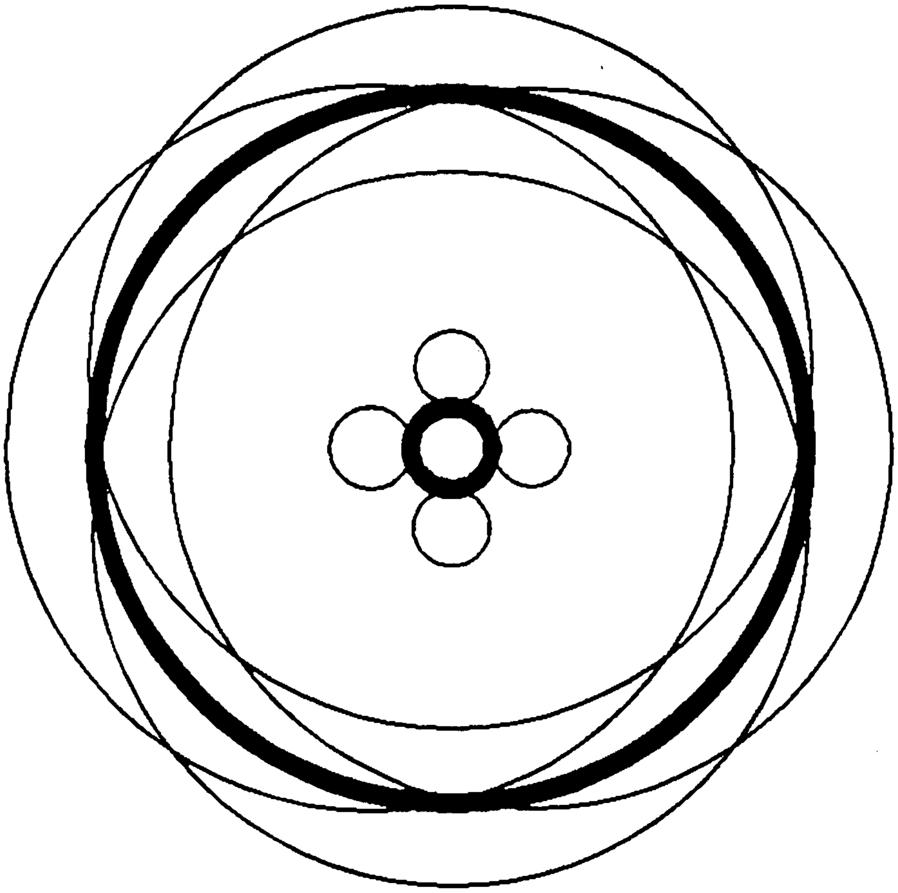
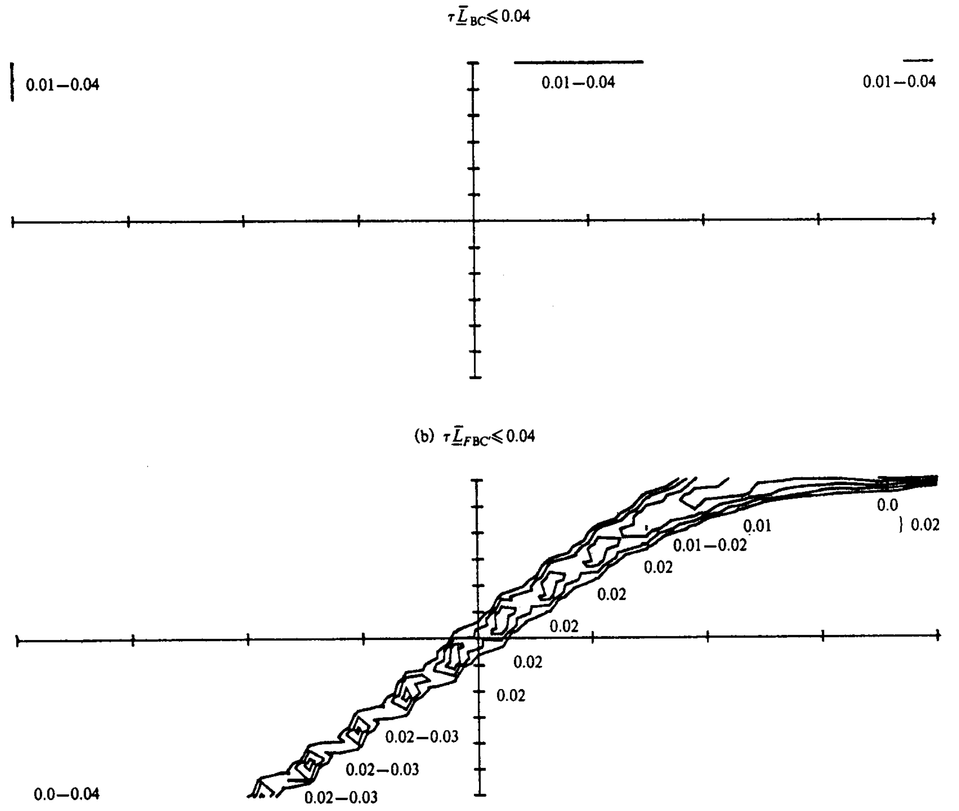
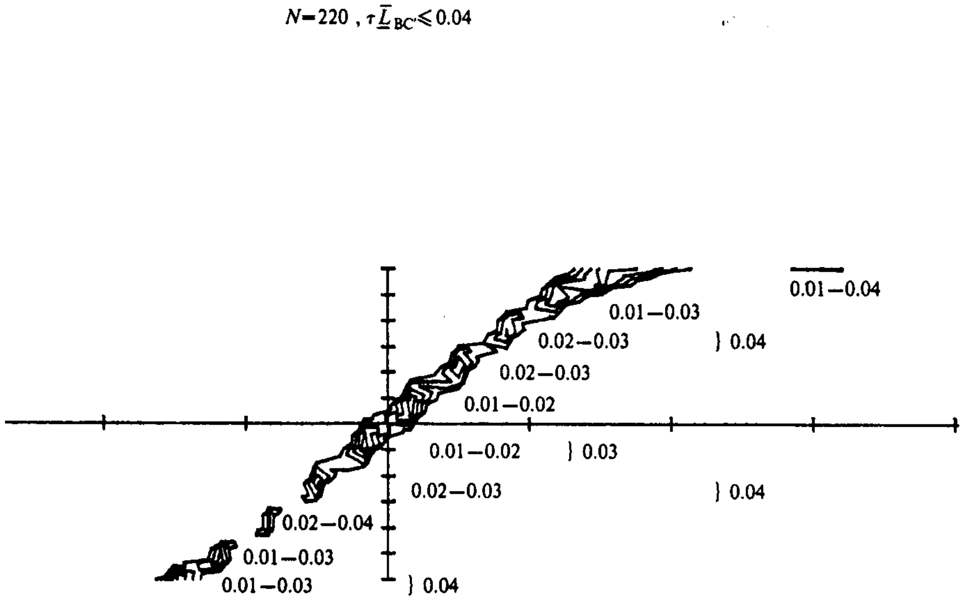


Figure 14-4



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

Figure 14-5



Author Photo