%A L. Ingber
%T Statistical mechanics of neocortical interactions:
Stability and duration of the 7±2 rule of short-term-memory capacity
%J Phys. Rev. A
%V 31
%P 1183-1186
%D 1985

Statistical mechanics of neocortical interactions. Stability and duration of 7 ± 2 rule of short-term-memory capacity

Lester Ingber

Physical Studies Institute, Drawer W, Solana Beach, California 92075 and Institute for Pure and Applied Physical Sciences, University of California at San Diego, La Jolla, California 92093

(Received 23 July 1984)

This paper is an essential addendum to L. Ingber, ‘‘Statistical mechanics of neocortical interactions. Derivation of short-term-memory capacity,’’ Phys. Rev. A 29, 3346-3358 (1984). Calculations are presented here to support the claim there, that there exists an approximate one-dimensional solution to the two-dimensional neocortical Fokker-Planck equation. This solution is extremely useful, not only to obtain a closed algebraic expression for the time of first passage, but also to establish that minima of the associated path-integral stationary Lagrangian are indeed stable points of the transient dynamic system. Also, a relatively nontechnical summary is given of the basic theory.

PACS numbers 1983 87.30.Gy, 87.10.+e, 02.50.+s, 05.50+q

The purpose of this paper is two-fold: First, and most important, calculations and plots are presented to support statements made in Ref. (1) that an approximate one-dimensional solution exists to the two-dimensional neocortical Fokker-Planck equation developed in a series of papers [1-5]. Second, and perhaps most useful to physicists in other specialties who are interested in this system, a relatively nontechnical summary is given of the basic theory.

This theory is geared to explain macroscopic neocortical activity, retaining as much correct description of underlying microscopic synaptic activity as can be carried by modern mathematical physics, which turns out to be sufficient for several important circumstances. In fact, as calculated here, the retained nonlinearities and statistics are essential to understand limitations to short-term memory (STM) capacity. (The mechanisms limiting STM capacity are not necessarily equivalent to the various electrical-chemical processes that are candidates for specific mechanisms of STM).

Microscopic neurons. Granted, 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, it is reasonable to at least respect the incredibly similar modular structure present in all regions of neocortex [6-11], still allowing for the differentiation among the laminar structure of individual modules and among neurons active at different time scales. Although laminar structure and tonic vs. phasic neurons can and will be included in future investigations, in preliminary studies the only differentiation was among E and I activities. For each of the four resulting interactions ($E \rightarrow E$, $I \rightarrow I$, $E \rightarrow I$, $I \rightarrow E$), neuronal interactions were calculated using distributions ranging from Poisson to Gaussian for inter-neuronal chemical-electrical synaptic interactions, in terms of quantal transfers of chemical transmitters. Note that some investigators have shown a Bernoulli distribution to be more accurate in some cases [3,12]. A Gaussian distribution was used for intra-neuronal activity, in terms of average effective voltages contributed from synaptic interactions to axonal polarizations. The mathematics [3,4] is similar to other studies [13], but with different physical interpretations. By extending the validity of an approximation [14] to the ‘‘erf’’ function [3], the firing of an individual neuron is well defined as a sigmoid distribution of synaptic interactions with other neurons. The derived ‘‘threshold factor’’ of synaptic interactions, being the argument of the conditional probability distribution which determines the rate at which the distribution changes from concave to convex, contains a mild nonlinearity induced by the above folding of inter-neuronal and intra-neuronal distributions. However, the following development is capable of accepting any differentiable nonlinear function for the threshold factor, including laminar circuitries and more complex synaptic interactions [4]. No linear rate-equation approximations to these distributed interactions or artificial assumptions of quadratic distributions are necessary, nor will these yield results consistent with empirical STM capacity.

Mesoscopic domains. As is found for most nonequilibrium systems, a mesoscopic scale is required to formulate the statistical mechanics of the microscopic system, from which the macroscopic scale can be developed. 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 $N \approx 100$ neurons (about 200 in visual cortex) of extent $\rho \sim 10^2$ μm comprise modular units vertically oriented relative to the warped and convoluted neocortical surface throughout most, if not all, regions of neocortex [6-11]. Clusters of about 100 neurons have been deduced to be reasonable from other considerations as well, which process information more as a neural ‘‘throng’’ similar to social interactions, than as hard-wired circuitry [15]. In this context, it is noted that the methods used here also have been applied to financial markets [16]. Since the short-ranged interactions between neurons take place within ~1 mm, which is the extent of a ‘‘macrocolumn’’ comprising $\sim 10^3$ 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 empirical situation, it is gratifying to find that $N \approx 10^2$ is just the right order of magnitude to permit a formal analysis using methods of mathematical physics recently developed for statistical systems [17,18]. N is small enough to permit nearest-neighbor (NN) 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 [19], 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 [20]. This is a definite mathematical convenience, else a macrocolumn of minicolumns would have to be described by a system of minicolumns with up to sixteenth order next-nearest neighbors. 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 to 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.

Macroscopic regions. Inclusion of all the above microscopic and mesoscopic features of neocortex permits a true nonphenomenological Gaussian-Markovian formal development for macroscopic regions encompassing $\sim 5 \times 10^5$ minicolumns of spatial extent $\Omega \sim 5 \times 10^9$ $\mu\mathrm{m}^2$, albeit one that is still highly nonlinear and nonequilibrium. The development of mesocolumnar domains presents conditional probability distributions for mesocolumnar firings M^G , $G = E$ or I, with spatially coupled NN interactions. The macroscopic spatial folding of these mesoscopic domains and their macroscopic temporal folding of tens to hundreds of τ yields a true path-integral formulation, in terms of a Lagrangian possessing a *bona fide* variational principle for most-probable firing states. 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 [18]. In the context of representing firing patterns by eigenfunctions of a Lagrangian, it is noted that these methods also have been used to study nucleon-nucleon velocity-dependent [21] potential contributions to nuclear-matter binding energy [22,23].

The prepoint-discretized Lagrangian $\underline{L}$, associated with the midpoint-discretized Feynman Lagrangian $\underline{L}_F$ [1], defines the path-integral solution for the conditional probability P of mesocolumnar firings M^G .

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

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

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

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

$$g^G = -\tau^{-1}(M^G + N^G \tanh F^G) , \; g^{GG} = \tau^{-1} N^G \mathrm{sech}^2 F^G ,$$

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

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

where $\bar{G}$ represents contributions from E and I neurons, and minicolumnar-averaged synaptic parameters $\{A^G_{G'}, B^G_{G'}, v^G_{G'}, \phi^G_{G'}\}$ and NN interactions $\underline{V}'^G$ are detailed in Ref. (1). The (excitatory) long-ranged fibers, represented by J_G constraints on M^G, most likely sustain activity over widely separated regions [24], to coordinate information processing ranging between local and global scales [5].

A Hamiltonian formulation can be obtained, one which does not permit simple "energy"-type conservation approximations, but one which does permit the usual time-evolution picture [18]. 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 diffusion in the space of E and I firings. The Fokker-Planck equation for the region Ω corresponding to Eq. (1) is

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

STM capacity. The most detailed and dramatic application of this theory has been to predict a stochastic mechanism underlying the phenomena of human STM capacity [1], transpiring on the order of tenths of a second to seconds, limited to the retention of 7 ± 2 items [25]. This is true even for apparently exceptional memory performers who, while they may be capable of more efficient encoding and retrieval of LTM, 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 [26]. This STM capacity-limited chunking phenomena also has been noted with items requiring varying depths and breadths of processing [27-31].

Contour plots of the stationary Lagrangian $\bar{\underline{L}}$ for model BC′ [1], wherein typical synaptic parameters are 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 $M^G = 0$, where $G = E$ or I, $-N^G \le M^G \le N^G$, and $N = N^E + N^I$ [1]. Within the range of synaptic parameters considered, for values of $\tau \bar{\underline{L}} \sim 10^{-2}$, this ‘‘centering’’ mechanism causes the appearance of from 5 to 10-11 extrema for values of $\tau \bar{\underline{L}}$ on the order of $\sim 10^{-2}$. (See Fig. 1.) 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.

Since the extrema 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 $\bar{\underline{L}}$ [4], a solution to the stationary probability distribution is hypothesized to be proportional to $\exp(-\Phi/D)$, where $\Phi = CN^2 \bar{\underline{L}}$, the diffusion $D = N/\tau$, and C a constant. 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. (See Fig. 2.) 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 $-\Phi$ [32], but 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.

More precisely, 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,

$$P_{\rm stat} \approx N_{\rm stat} g^{1/2} \exp(-CN\tau \bar{\underline{L}}) \ , \tag{3}$$

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

where $N_{\rm stat}$ and C are constant factors. An estimate of the approximation made in Eq. (3) is made by seeking values of constants C, such that the stationary Fokker-Planck equation is satisfied exactly. Contour plots in Fig. 2 of C versus $\bar{M}^G$ demonstrates that there exists real positive C which may only range from $\sim 10^{-1}$ to ~ 1, for which there exists unbroken contours of C which pass through or at least border the line of minima. At each point $\bar{M}^G$, Eq. (3) leaves a quadratic equation for C to be solved. Dropping the $g^{1/2}$ factor in Eq. (3) results in C not being real throughout the domain of $\bar{M}^G$.

Thus, Eq. (3) defines a solution with potential $N^2 \bar{\underline{L}} = \int A dM$, drift A, and diffusion N/τ. Stability of transient solutions, defined for δM^G about a stationary state by $\delta \dot{M}^G \approx -A_{,G} \delta M^G = -N^2 \bar{\underline{L}}_{,GG} \delta M^G$, is therefore equivalent to $\ll \bar{M} \gg$ being minima of $\bar{\underline{L}}$. This stationary solution is also useful for calculating the time of first passage, t_{vp}, to fluctuate out of a valley in one minima over a peak to another minima. 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 9 of the minima when the maximum of 10—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 [1]. Thus, 7 ± 2 is the capacity of STM, for memories or new patterns which can be accessed in any order during tenths of a second, all as observed empirically [26].

This is a very sensitive calculation. If N were a factor of 10 larger, or if $\tau \bar{\underline{L}} < 0.1$ at the minima, then t_{vp} is on the order of hours instead of seconds, becoming unrealistic for STM durations. Oppositely, if t_{vp} were much smaller, i.e., less than $\sim 5\tau$, this would be inconsistent with empirical time scales necessary for formation of any memory trace [33]. In this context, it is noted that the threshold factor of the probability distribution scales as $(N^* N)^{1/2}$, demanding that both macrocolumnar divergence and minicolumnar convergence of mesocolumnar firings be tested by these calculations.

The statistical nature of this storage and processing also explains the primacy vs. 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 [34]. The deepest minima 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 others. The more recent memories or newer patterns may be presumed to be those having synaptic parameters more recently tuned and/or more actively rehearsed.

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

[2] L. Ingber, ‘‘Towards a unified brain theory,’’ *J. Social Biol. Struct.* **4**, 211-224 (1981).

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

[4] L. Ingber, ‘‘Statistical mechanics of neocortical interactions. Dynamics of synaptic modification,’’ *Phys. Rev. A* **28**, 395-416 (1983).

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

[6] C.D. Gilbert and T.N. Wiesel, ‘‘Functional organization of the visual cortex,’’ *Prog. Brain Res.* **58**, 209-218 (1983).

[7] P.S. Goldman and W.J.H. Nauta, ‘‘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).

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

[9] T.J. Imig and R.A. Reale, ‘‘Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex,’’ *J. Comp. Neurol.* **192**, 293-332 (1980).

[10] E.G. Jones, J.D. Coulter, and S.H.C. Hendry, ‘‘Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys,’’ *J. Comp. Neurol.* **181**, 291-348 (1978).

[11] V.B. Mountcastle, ‘‘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.

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

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

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

[15] T.H. Bullock, ‘‘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).

[16] L. Ingber, ‘‘Statistical mechanics of nonlinear nonequilibrium financial markets,’’ *Math. Modelling* **5**, 343-361 (1984).

[17] R. Graham, ‘‘Covariant formulation of non-equilibrium statistical thermodynamics,’’ *Z. Physik* **B26**, 397-405 (1977).

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

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

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

[21] L. Ingber, ‘‘Nuclear forces,’’ *Phys. Rev.* **174**, 1250-1263 (1968).

[22] L. Ingber, ‘‘Riemannian corrections to velocity-dependent nuclear forces,’’ *Phys. Rev. C* **28**, 2536-2539 (1983).

[23] L. Ingber, ‘‘Path-integral Riemannian contributions to nuclear Schrödinger equation,’’ *Phys. Rev. D* **29**, 1171-1174 (1984).

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

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

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

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

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

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

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

[31] L. Ingber, *Elements of Advanced Karate* (Ohara, Burbank, CA, 1985).

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

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

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

FIG. 1. From Ref. (1). Contours of $\tau \bar{L}_{\mathrm{BC}'}$ over the $(\bar{M}^E, \bar{M}^I)$ plane for values less than 0.04 are drawn, invoking the ''centering'' mechanism for model BC′, where $-80 \le \bar{M}^E \le 80$ and $-30 \le \bar{M}^I \le 30$. Ref. (1) calculates that the results are similar for the Feynman Lagrangian.

FIG. 2. Using the ansatz that a solution to the two-dimensional M^G-space Fokker-Planck equation for the stationary probability distribution of $\bar{M}^G$ is proportional to $\exp(CN\tau\bar{L})$, contours of positive C are plotted over $\bar{M}^G$ for which this ansatz does indeed satisfy this equation. Plots are presented for model BC′ of Fig. 1. (a) C-contours at 7 equally spaced values between the minimum value of 0, and the maximum value of 600, give rise to a peak close to the origin and a solitary contour of 100 at $\bar{M}^G = (\bar{M}^E, \bar{M}^I) = (-80, -30)$. (b) C-contours cut off at 10 give rise to contours of disconnected regions from 0 to 10 along the line of minima of Fig. 1. (c) C-contours cut off at 1.0 give rise to contours ranging from 1.0 at $\bar{M}^E = -80$ towards 0.2 on the LHS, from 0.2 to 1.0 on the RHS at $\bar{M}^E = 80$, and from 0.2 on the outer contours along the line of minima of Fig. 1 to 1.0 on the inner contours. Note that the contours at 1.0 are closest to the center of the minima. (d) C-contours cut off at 0.1 give rise to a valley at 0 at $(\bar{M}^E, \bar{M}^I) = (-80, -30)$, and contours at 0.1. Note that no new contours appear in addition to those already found in Fig. (c).