ARTICLE **Communicated by Carl van Vreeswijk**

Dynamics of Encoding in Neuron Populations: Some General Mathematical Features

Bruce W. Knight
Laboratory of Biophysics, Rockefeller University, and Laboratory of Applied Mathematics, Mount Sinai Medical School, New York University, New York, NY 10021, U.S.A.

The use of a population dynamics approach promises efficient simulation of large assemblages of neurons. Depending on the issues addressed and the degree of realism incorporated in the simulated neurons, a wide range of different population dynamics formulations can be appropriate. Here we present a common mathematical structure that these various formulations share and that implies dynamical behaviors that they have in common. This underlying structure serves as a guide toward efficient means of simulation. As an example, we derive the general population firing-rate frequency-response and show how it may be used effectively to address a broad range of interacting-population response and stability problems. A few specific cases will be worked out. A summary of this work appears at the end, before the appendix.

1 Population Dynamics: Preliminaries

A small patch of cortex may contain thousands of similar neurons, each communicating with hundreds or thousands of other neurons in that same patch or in other patches. By the very nature of this situation, an attempt to simulate it directly must be massive. A recently advanced and promising alternative is the simulation of the dynamical rules stating how the statistical status of the whole population evolves with time. Three recent expositions of this approach are Knight, Manin, and Sirovich (1996), Nykamp and Tranchina (1999), and Omurtag, Knight, and Sirovich (1999). Precursors with goals of a more analytic nature are Stein (1965), Johannesma (1969), Wilbur and Rinzel (1982), Kuramoto (1991), Abbott and van Vreeswijk (1993), Gerstner (1995). Here we briefly give one example of the approach and then discuss the generality of its central features.

Models of single neuron dynamics are typified by the equations of Hodgkin and Huxley (1952). A serviceable caricature of these equations, and also of more elaborate dynamical models (see the appendix), is the generalized integrate-and-fire equation:

$$dx/dt = f(x, s) + \eta(x, s), s = s(t), x = 1 \text{ resets to } x = 0. \tag{1.1}$$

Here x may be regarded as a transmembrane voltage, which depends on the accumulation of a transmembrane current $f(x, s)$ common to all neurons in a population, and which current, in turn, depends on the membrane voltage x and also on a synaptic input signal $s(t)$; each neuron also feels a "noise current" (Tuckwell, 1988) $\eta(x, s)$ which is statistically uncorrelated with the noise currents felt by other neurons. We may view the resetting of $x = 1$ to $x = 0$ as the occurrence of a nerve impulse. We will use equation 1.1 to illustrate particular cases of the general features.

The central structural features of the population approach emerge from the consideration of a single population of *noninteracting* neurons. The interesting further features that emerge, for interacting subpopulations, may be addressed with little further overhead in formulation. Our early discussion consequently concerns a population of noninteracting neurons. (Interactions among subpopulations and feedback interaction among the members of one population are discussed in sections 4 and 5).

The neuron population is distributed over the voltage variable x, between 0 and 1, with a distribution function $\rho(x, t)$, which integrates to unity and whose evolution in time must be determined by equation 1.1. The fraction of neurons whose voltage is more than a given value x changes with time in a way that depends on the probability current,

$$\begin{aligned} J(x, t) &= f(x, s)\rho(x, t) - D(x, s)\frac{\partial}{\partial x}\rho(x, t) \\ &= \left(f(x, s) - D(x, s)\frac{\partial}{\partial x}\right)\rho(x, t), \end{aligned} \tag{1.2}$$

which follows from equation 1.1. The first term—the advection term—states that the probability density is swept along by the common velocity. The second term—the diffusion term—states that neurons are more likely to be shaken (by fluctuations) from a voltage region, where more of them reside to an adjacent region, where they are less dense, than their converse likelihood of being shaken up the density gradient. Quantitatively, the diffusion coefficient $D(x, s)$ must be derived from a more detailed specification of the noise $\eta(x, s)$. Equation 1.2 is taken from Knight et al. (1996), where a brief derivation of D is given. Comparable expressions for current (specialized in different ways) are given by Abbott and van Vreeswijk (1993; equation 9.4), by Nykamp and Tranchina (1999; equations A4 and A16), and by Omurtag et al. (1999; equations 36 and 48). We note in equation 1.2 that the current is obtained from the probability density by a linear transformation on $\rho(x, t)$. From the defining property of the current, we easily show that its change with x gives the conservation equation (or continuity equation) for the local rate at which density accumulates:

$$\partial\rho/\partial t = -\partial J/\partial x. \tag{1.3}$$

(In the present context, equation 1.3 was used by Knight et al. 1996, equation 4; Abbott & van Vreeswijk, 1993, equation 3.3; Nykamp & Tranchina, 1999, equation 10.)

The linear transformation from ρ to J, equation 1.2, may be substituted into the conservation equation, 1.3, to obtain the dynamical equation,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(-f(x,s) + D(x,s)\frac{\partial}{\partial x}\right)\rho \tag{1.4}$$

(Knight et al. 1996, equation 8; Abbott & van Vreeswijk, 1993, equation 9.4, Omurtag et al. 1999, equation 49), which describes how $\rho(x,t)$ evolves with the passage of time. (Such a partial differential equation for probability density is technically a Fokker-Planck equation; an extensive reference is Risken, 1996.) Equation 1.4 may be integrated over time to obtain the statistical description of the population's evolution. This also yields $J(x,t)$ by equation 1.2, and the current across threshold yields the per-neuron firing rate:

$$r(t) = J(1,t) \tag{1.5}$$

(Abbott & van Vreeswijk, 1993, equation 9.9; Nykamp & Tranchina, 1999, equation 11; Omurtag et al., equation 39). We note from equation 1.5 that $r(t)$ is linearly related to $J(x,t)$, and so also is related linearly to $\rho(x,t)$.

Particular neurons, or particular questions concerning their dynamical behavior, or other special circumstances may require a model more detailed, or differently expressed, than what followed from equation 1.1. For example, the Hodgkin-Huxley equations describe a neuron's momentary state in terms of four variables (v,m,h,n), including two (m,h) that determine the conductance of sodium channels and one (n) that governs potassium conductance. The population density in this case lives in a four-dimensional state-space, and the corresponding population current has four components that describe the flow of population density in the several coordinate directions in that space. Relay cells in the lateral geniculate nucleus (McCormick & Huguenard, 1992) reveal the presence of 10 conductance channel types, which specify a state-space with 18 dimensions. Some neurons have cell bodies that are not tightly coupled electrically with their extended dendrites, and such cells must be described by a state-space that has separate sets of dimensions for each of their multiple electrical compartments. Synapses that deliver finite increments of synaptic electric charge require a more detailed formulation (Nykamp & Tranchina, (1999); Omurtag et al., 1999). Synapses with a slow response dynamics add synapse dynamics dimensions to the state-space. On the other hand (as briefly discussed in section 6.1), within a state-space the neuronal dynamics places a huge constraint on which probability-density configurations can actually be reached; in a specific case (Sirovich, Knight, & Ormutag, 1999c), this allows one to

replace the population density of equation 1.4 (where the density is a member of the infinite-dimensional space of functions that depend on x between 0 and 1) by a 17-entry vector, and allows the reduction of the differential operator of equation 1.4 to a 17×17 (s-dependent) matrix. The general structural features reviewed below carry over directly to such an efficient matrix representation of the population dynamics.

2 Population Dynamics: Common Structure

In all of the examples cited, we find the following common elements of structure:

$s(t)$: an input signal. (2.1)

ρ: a probability density. (2.2)

In many of the above cases, ρ is a function defined over a state-space; the collection of such functions has the structure of a linear vector space (a weighted sum of two members yields another member). In this sense, in all of the above cases ρ may be regarded as a member of a linear vector space.

J: a probability current. (2.3)

Since in many cases, such as the Hodgkin-Huxley case above, J has several components, convenience dictates that we regard J as in a space different from that which contains ρ. However, there are natural connections:

$C(s)$: a current linear operator such that (2.4)

$$C\rho = J. \tag{2.5}$$

In the case discussed above, equation 1.2 is a particular example of equation 2.5, and the linear operator that appears in equation 1.2 is the corresponding particular example of $C(s)$:

K: a conservation operator such that (2.6)

$$\frac{\partial \rho}{\partial t} = -KJ. \tag{2.7}$$

In the examples above where ρ is represented in a state-space, K is the familiar divergence operator of ordinary vector analysis, which in our one-dimensional example reduces equation 2.7 to equation 1.3.

$Q(s) = -KC(s)$: a dynamical operator. (2.8)

Substitution of equation 2.5 into 2.7 gives

$$\frac{\partial \rho}{\partial t} = Q(s(t))\rho: \qquad \text{a dynamical equation for} \qquad \rho. \tag{2.9}$$

Integration of the dynamical equation yields ρ at all times, and thus gives a statistical description of the evolving population. In our first example, equation 2.9 becomes 1.4.

$$r(t): \qquad \text{the firing rate (per neuron) of the population.} \tag{2.10}$$

$$R[J, s]: \qquad \text{a firing-rate linear functional of } J, \tag{2.11}$$

a linear transformation from J to the scalar firing rate $r(t)$ such that

$$r = R[J], \tag{2.12}$$

which implies the linear relation,

$$r = R\left[(C\rho)\right], \tag{2.13}$$

that yields the momentary firing rate from the momentary population density. In our illustrative case above, equation 2.12 specializes to 1.5. Other cases can be more elaborate. For example, the firing rate of a population of Hodgkin-Huxley neurons is the rate at which individual members achieve the peak values of their action potentials. At the moment of maximum voltage, the right-hand side of the Hodgkin-Huxley voltage equation equals zero, and that constraint defines a 3-surface in the four-dimensional state-space. Integration of the flux of probability current across that surface yields the (per-neuron) firing rate, and that integral is linear in J, as equation 2.11 states. Because the value of an action potential's voltage maximum is influenced by the input current s, the location of the 3-surface is s-dependent, which leads to the general s-dependence of $R[J, s]$ in equation 2.11.

For an additional view of the mathematical structure just presented, we may imagine ourselves endowed with a computer of mythical capability and ask how we might approach the dynamical equation, 2.9. For examples where ρ is naturally represented as a probability density supported by a state-space, we may subdivide that state-space into hypercubes, each sufficiently small to ensure whatever numerical accuracy we demand. If we do this, all the numbered items above may be cast in terms of finite matrices, and several structural features become more clearly visible. For example, we may observe that the dynamical equation, 2.9, describes what technically is a Markov process (Bailey, 1964), and this ensures the existence of some items we use below.

Our informal discretization is introduced at this point not as an approximation but as an aid to uniform presentation. It leads us, by means of finite linear algebra, to some valuable general structural features that may also be obtained directly for each of the several different cases quoted above, but by means that must be customized anew for each case.

In particular we note that for any specified input s, the dynamical operator Q of equation 2.9 has a set of eigenvectors ϕ_n, which it maps to eigenvalue multiples λ_n of themselves:

$$Q\phi_n = \lambda_n \phi_n. \tag{2.14}$$

All of these depend on the value of the input s:

$$Q = Q(s), \phi_n = \phi_n(s), \lambda_n = \lambda_n(s). \tag{2.15}$$

For time-independent s, equation 2.9 has a time-independent steady-state solution $\rho_0(s)$ for which

$$Q\rho_0 = 0; \tag{2.16}$$

and time independence need not be invoked to observe that this is a distinguished eigenvector of equation 2.14 with eigenvalue $\lambda_0 = 0$. This eigenvalue, of course, is independent of s in spite of equation 2.15. However, the zeroth eigenvector $\rho_0(s)$ *does* depend on the input value s.

The existence of a zero eigenvalue implies that all columns of the matrix Q have a common linear dependence expression satisfied by their elements. To determine this dependence we may exploit the conservation of probability. (Only in equations 2.17, 2.18, and 2.19 will subscripts refer to the compartments of our informal discretization.) Discretized, equation 2.9 is a column of equations, the jth of which reads:

$$\frac{d\rho_j}{dt} = \sum_k Q_{jk}\rho_k. \tag{2.17}$$

The ρ_j's are probabilities that sum to unity, and so their sum has no time dependence. Thus, if equation 2.17 is summed on j, the left side is zero:

$$0 = \sum_k \left(\sum_j Q_{jk} \right) \rho_k. \tag{2.18}$$

Equation 2.18 is satisfied at every moment, including the moment that arbitrary initial data are imposed. Initially, we may choose any one of the

entries ρ_k to be unity, and choose the rest to be zero, so that equation 2.18 implies

$$\sum_j Q_{jk} = 0: \tag{2.19}$$

every column sums to zero.

The discretized case has a natural inner product; if u and v are two column vectors, then

$$(v, u) \tag{2.20}$$

is the sum of products of the entries of v and u if those entries are real numbers; if the entries are complex numbers, we instead take the complex conjugates of the v entries before performing the bilinear sum. (This prescription endows each of our particular cases with an inner product, a bilinear integral over state-space, if we take the small hypercubes, of our discretizing construction, to their infinitesimal limit.) For any Q, the inner product induces

$$Q^*: \qquad \text{the adjoint operator to } Q, \tag{2.21}$$

which, for any u, v satisfies

$$\left(Q^* v, u\right) = (v, Qu). \tag{2.22}$$

The entries of the matrix Q are real numbers, and its matrix transpose Q^* satisfies equation 2.22.

The determinant of a matrix has the same value as the determinant of its transpose, so the eigenvalues of Q, which satisfy

$$\det(\lambda I - Q) = 0, \tag{2.23}$$

are the same as those of its adjoint Q^*. We may regard equation 2.23 as a polynomial in λ and must anticipate that it can have complex roots. However, λ is the only possibly complex item in equation 2.23, and so complex conjugation of that equation simply replaces λ by λ^*. Thus we see λ^* is also a root, and complex eigenvalues occur in conjugate pairs.

The adjoint operator $Q^*(s)$ has eigenvectors $\widehat{\phi}_m(s)$, which satisfy

$$Q^* \widehat{\phi}_n = \lambda_n^* \widehat{\phi}_n. \tag{2.24}$$

Here (for convenience, which soon will be evident), we have chosen a labeling that associates the label n with the eigenvector whose eigenvalue is the complex conjugate of λ_n. We now observe that

$$\begin{aligned} \lambda_n \left(\widehat{\phi}_m, \phi_n\right) &= \left(\widehat{\phi}_m, \lambda_n \phi_n\right) = \left(\widehat{\phi}_m, Q\phi_n\right) = \left(Q^* \widehat{\phi}_m, \phi_n\right) \\ &= \left(\lambda_m^* \widehat{\phi}_m, \phi_n\right) = \left(\widehat{\phi}_m, \lambda_m \phi_n\right) = \lambda_m \left(\widehat{\phi}_m, \phi_n\right). \end{aligned} \tag{2.25}$$

Subtracting the last form from the first gives

$$(\lambda_n - \lambda_m)\left(\widehat{\phi}_m, \phi_n\right) = 0. \tag{2.26}$$

Since one factor or the other must vanish, we see that sets of eigenvectors, and sets of adjoint eigenvectors, which are associated with distinct eigenvalues, form biorthogonal sets. Linear algebra tells us that generically (that is, except for exceptional cases that are "unlikely" in a strict sense) there are as many distinct eigenvectors (and adjoint eigenvectors) as there are dimensions in the vector space. We may impose a normalization on these sets and get

$$\left(\widehat{\phi}_m, \phi_n\right) = \delta_{mn}. \quad \delta_{mn} = 1 \text{ if } m = n, = 0 \text{ otherwise.} \tag{2.27}$$

Moreover, the dimension count says that the eigenvectors are complete in the sense that any vector v may be represented as a sum of eigenvectors:

$$v = \sum_n a_n \phi_n, \tag{2.28}$$

where each a_m may be evaluated by observing that

$$\left(\widehat{\phi}_m, v\right) = \left(\widehat{\phi}_m, \sum_n a_n \phi_n\right) = \sum_n a_n \left(\widehat{\phi}_m, \phi_n\right) = a_m, \tag{2.29}$$

where the last step followed from equation 2.27.

An immediate first application is the following: suppose in the dynamical equation 2.9 that s has a fixed value and that for $\rho(t)$, we have initial data $\rho(0)$; then since equation 2.28 can represent any vector, we know we can represent $\rho(t)$ as

$$\rho(t) = \sum_n a_n(t) \phi_n, \tag{2.30}$$

where

$$a_n(0) = \left(\widehat{\phi}_n, \rho(0)\right). \tag{2.31}$$

The full solution to the dynamical equation, 2.9, is given by

$$a_n(t) = a_n(0) e^{\lambda_n t}, \tag{2.32}$$

whence

$$\rho(t) = \sum_n a_n(0) e^{\lambda_n t} \phi_n, \tag{2.33}$$

as substitution of equation 2.33 into 2.9 shows that the time differentiation, on the left of equation 2.9, produces on each term the same eigenvalue as does the action of Q on the corresponding eigenvector on the right.

In general, if noise entered the formulation of the operator Q in equation equation 2.9 (and if s is constant), then an initial departure from equilibrium will die out, and with the exception of $\lambda_0 = 0$, all the other eigenvalues will have real parts that are negative. (This is a general property of Markov matrices; Bailey, 1964.) At long times only the $n = 0$ term in equation 2.33 will survive, giving (with equation 2.31)

$$\rho\,(t \to \infty) = \left(\widehat{\phi}_0,\, \rho\,(0)\right) \rho_0. \tag{2.34}$$

Clearly the zeroth adjoint eigenfunction is a valuable item to understand. We have seen in the discretized format that each matrix column of $Q(s)$ must sum to zero. As $Q^*(s)$ is its transpose, every row of $Q^*(s)$ will sum to zero. By the rules of matrix action on a column vector, for a vector whose entries are all the same, the action of Q^* will yield the zero vector because every row-sum vanishes. Thus, the column vector whose entries are all the same constant is the eigenvector $\widehat{\phi}_0(s)$ of $Q^*(s)$, which has eigenvalue zero. In fact it is *not* a function of s, but a vector $\widehat{\rho}_0$ that is universal for *all* s (just as is its eigenvalue zero).

We have asserted both that the equilibrium probability function $\rho_0(s)$ accumulates to a total probability of unity and that it is the eigenfunction of $Q(s)$ with eigenvalue zero. We have also chosen the biorthonormal normalization, equation 2.27. In the particular case of $m = n = 0$, equation 2.27 becomes

$$(\widehat{\rho}_0,\, \rho_0(s)) = 1, \tag{2.35}$$

which implies that we choose the value 1 for the constant entry of $\widehat{\rho}_0$.

The constant-entry property of $\widehat{\rho}_0$ shows a special feature of the eigenfunctions of $Q(s)$. If, in the biorthogonality equation, 2.27, we choose $m = 0$ and $n \neq 0$, we obtain

$$(\widehat{\rho}_0,\, \phi_n(s)) = 0. \tag{2.36}$$

Because the entries of $\widehat{\rho}_0$ are constant, this says that the entries of ϕ_n sum to zero. Thus, in the eigenfunction expansion, equation 2.30, the vectors ϕ_n for $n \neq 0$ do not contribute to the probability inventory, and we verify that the decay of those modes in equation 2.33 does not make probability vanish from the system.

If the input s is constant, by equation 2.13 the steady-state firing rate is given by

$$r(s) = R\left[C(s)\rho_0(s),\, s\right], \tag{2.37}$$

which specifies the steady-state input-output relation for the neuron population. We note from equation 2.37 that already in the steady state, r responds to s in a manner that in general is nonlinear. Evidently equation 2.37 also predicts the population response to a time-dependent $s(t)$, which changes slowly enough. Although the estimate is informal, it is fairly clear that "slowly enough" means that the characteristic change rate of $s(t)$ should be small compared to the slowest decay rate determined by the eigenvalues in the exponents of equation 2.33, in mathematical terms,

$$\left|(1/s(t))\left(ds(t)/dt\right)\right| << \left|\mathrm{Re}\left(\lambda_1\left(s(t)\right)\right)\right|, \tag{2.38}$$

where the index 1 has been assigned to the nonzero eigenvalue whose absolute real part is smallest.

The response to a step input also may be stated in exact closed form, because the population response is determined by the linear functional R of equation 2.13. Suppose that the population density has achieved an equilibrium distribution $\rho_0(s_A)$ at input s_A and that the input signal is jumped to a new value, s_B. We set our time origin at the time of the jump. Equations 2.13, 2.31, and 2.33 give the time course of the response:

$$\begin{aligned} r(t) &= \sum_n e^{\lambda_n(s_B)t}\left(\widehat{\phi}_n\left(s_B\right), \rho_0\left(s_A\right)\right) R\left[C\left(s_B\right)\phi_n\left(s_B\right), s_B\right] \\ &= R\left[C\left(s_B\right)\rho_0\left(s_B\right), s_B\right] \\ &\quad + \sum_{n\neq 0} e^{\lambda_n(s_B)t}\left(\widehat{\phi}_n\left(s_B\right), \rho_0\left(s_A\right)\right) R\left[C\left(s_B\right)\phi_n\left(s_B\right), s_B\right], \end{aligned} \tag{2.39}$$

where our special information on the zeroth eigenfunctions (equation 2.35) has been used in the second line, to isolate the new steady state (compare the first term with equation 2.37). The remaining terms contribute a smooth transient to what would have been an abrupt jump if equation 2.37 had been recklessly used and 2.38 ignored. The jump response (equation 2.39) has been confirmed in the case of equation 1.4 by comparison to both a direct integration of that equation and a direct simulation involving 90,000 integrate-and-fire neurons (Knight, Ormutag, & Sirovich, 1999).

The effect of our informal discretization above is to make the sum in equation 2.39 large and finite rather than infinite. This distinction is not of practical importance; Knight et al. (1999) show that the sum converges with a few terms.

An understanding of the eigenvalues clearly is valuable in the interpretation of results such as equation 2.33. While standard algorithms are conveniently available that essentially solve equation 2.23, there are several approximate procedures, and some informal arguments as well, that give valuable systematic information concerning the eigenvalues. A case that we may expect to occur frequently is that of a neuron model that incorporates

stochastic effects but quantitatively resembles a counterpart nonstochastic model, with the further feature that the steady input s_0 is fixed at a value where the nonstochastic counterpart fires repetitively on a stable limit cycle. A simple example is equation equation 1.1, with s_0 so chosen that $f(x, s_0)$ is always positive and does not go to zero for any choice of x between 0 and 1; in the nonstochastic counterpart, the noise η is set to zero. In the more detailed example of the Hodgkin-Huxley equations, a steady input current in the most interesting range gives a limit cycle that is a one-dimensional closed curve in 4-space, around which a system point circulates at fixed frequency. Even in most neuron models with many electrical compartments and hence a state-space with many dimensions, periodic firing under steady input implies a system point that pursues a closed one-dimensional limit cycle curve embedded in that many-dimensional space.

For the nonstochastic counterpart, we may construct a time-dependent probability density $\rho(\mathbf{x}, t)$ that is degenerate in the sense that it is nonzero only on the limit cycle curve. But along that curve, at some initial moment a variable density of individual systems may be freely chosen. After one period, any cluster of neurons in the population will return to the same point from which they started, so that the probability density $\rho(\mathbf{x}, t)$ is periodic in time and may be represented as a Fourier series in time:

$$\rho\,(\mathbf{x}, t) = \sum_n e^{in\omega_0 t} \rho_n\,(\mathbf{x})\,, \tag{2.40}$$

where $\omega_0(s_0)$ is 2π times the fundamental firing frequency. We observe that equation 2.40 is entirely consistent with equation 2.33, with a set of eigenvalues,

$$\lambda_n(s) = in\omega_0(s), \tag{2.41}$$

that are pure imaginary and come in complex-conjugate pairs except for the value zero. Here the integer index ranges over $-\infty < n < \infty$. In equation 2.40, the $\rho_n(\mathbf{x})$ may be regarded (aside from normalization) as the corresponding eigenfunctions.

The stable limit cycle is an "attractor" in the sense that a system point near it is drawn to the limit cycle with the passage of time. The introduction of stochastic effects will superimpose a random walk on the trajectory of a system point, but the excursions of that random walk will be limited by the greater attraction exerted on trajectories further from the limit cycle. In consequence, stochastic effects will produce an equilibrium probability density that will have appreciable size in a region of state-space that we might describe as a "fattened limit cycle." We may expect that other eigenfunctions whose eigenvalues lie nearest zero (counterparts of the $\rho_n(\mathbf{x})$ in 2.40, with integer n near zero) will be confined to this region as well. A probability density on that fattened limit cycle feels both advection and dif-

fusion around the cycle. More detailed exploration suggests that with some generality, eigenvalues whose indices lie near zero approximate the form

$$\lambda_n(s) = -n^2\Gamma(s) + in\omega_0(s), \tag{2.42}$$

where $\Gamma(s)$ is positive. This form was found by Abbott and van Vreeswijk (1993, equation 9.15) for the case of equation 1.4 with f and D constant; by Knight et al. (1996, equation 36) for D small and constant and with

$$f(x, s) = s - \gamma x \tag{2.43}$$

(case of "forgetful" or "leaky" integrate-and-fire encoder); by Sirovich, Knight, and Omurtag (1999a) for a case with finite synaptic jumps (both analytic and numerical confirmation); and further analytic evidence concerning the generality of this approximate eigenvalue rule, equation 2.42, is given in the appendix. (See also equation 3.24 and the comments following it.) The presence of diffusion tends to make the fundamental frequency ω_0 more rapid than it is in the nonstochastic counterpart case (Knight et al., 1996, equation 36).

Finally, there are neurons that, under steady input, cyclically fire bursts of nerve impulses. For such a neuron, the stable attractor is a two-torus embedded in the higher-dimensional state-space. The system motion exhibits two fundamental frequencies ("around and along the doughnut," which, respectively, characterize the mean impulse rate and the burst rate), and instead of equation 2.40, we will have a sum that features the various combination frequencies of the two fundamentals. A similar but more elaborate eigenvalue theory will follow from these considerations.

3 The Frequency Response

Much valuable structural information may be learned by the study of how the neuron population responds to an input that consists of a large, steady part plus a small, superimposed part that is sinusoidal in time. A valuable slight generalization is a superimposed part that was small throughout its past history and consists of a sinusoid multiplied by an exponentially growing envelope function.

The general strategy comes in two steps. The first step is to observe that a small time-variable addition to steady input will cause a small time-dependent addition to the steady value of any system variable that ultimately depends on the input. We will regard the relatively easy (though nonlinear) problem, of determining the steady values of the system variables, as already solved. Large, steady values we will give the index zero; small, time-varying ones will be given the index unity. In algebraic manipulations any product of terms with index unity will be dropped as small to

second order. Following our numbered expressions of section 2, the input signal,

$$s = s_0 + s_1(t), \tag{3.1}$$

gives rise to the probability density,

$$\rho = \rho_0 + \rho_1(t), \tag{3.2}$$

and the probability current,

$$J = J_0 + J_1(t). \tag{3.3}$$

The current linear operator, through first order, is

$$C_0 + C_1 = C\,(s_0 + s_1(t)) = C\,(s_0) + s_1(t)C_s, \tag{3.4}$$

where the subscript s implies differentiation with respect to s of the s-dependent linear operator C. In the illustrative case of the operator C in equation 1.2 (also see the particular example, equation 2.43), both f and D are to be differentiated with respect to s. We note in equation 3.4 that the zeroth-order terms on the left and right sides are equal and may be subtracted, leaving only the evaluation of the first-order term ($C_1 = s_1 C_s$). Below, the same sort of steps will be taken repeatedly. The mapping from density to current, equation 2.5 now gives a first-order part,

$$C_0\rho_1 + s_1 C_s \rho_0 = J_1, \tag{3.5}$$

so that the perturbed time-dependent current arises from two pieces, the second and less obvious of which depends directly on small changes in the input.

The conservation operator K is *not* s-dependent; as it is linear, we have

$$K\,(J_0 + J_1) = KJ_0 + KJ_1, \tag{3.6}$$

and by equation 2.7, probability conservation at first order is

$$\frac{\partial \rho_1}{\partial t} = -KJ_1. \tag{3.7}$$

By equation 2.8, the first-order departure (from equilibrium) of the dynamical operator Q is

$$Q_1 = -KC_1 = -s_1 KC_s = s_1 Q_s, \tag{3.8}$$

where the last form establishes a more convenient notation. From equation 2.9 the first-order dynamical equation is

$$\frac{\partial \rho_1}{\partial t} = Q_0 \rho_1 + s_1 Q_s \rho_0. \tag{3.9}$$

The final term, which depends directly on the input, involves only known quantities and makes the equation linear inhomogeneous in the first-order probability density and also linear in the first-order input. In principle these observations provide the means to solve equation 3.9 analytically for ρ_1 (for example by the construction of a Green's function solution). By equation 2.11 the firing rate satisfies

$$r_0 + r_1(t) = R[J_0, s_0] + R[J_1, s_0] + s_1 R_s[J_0, s_0]. \tag{3.10}$$

With the two terms from the J_1 equation, 3.5, this gives the first-order response,

$$\begin{aligned} r_1 &= R[C_0 \rho_1, s_0] + s_1 (R[C_s \rho_0, s_0] + R_s[C_0 \rho_0, s_0]) \\ &= R[C_0 \rho_1] + s_1 \widehat{R}[\rho_0], \end{aligned} \tag{3.11}$$

where the second line simplifies notation a bit and reminds us that the departure in firing rate can show a direct dependence on input, as well as a response through the departure in population density ρ_1.

The second step in constructing a frequency response is to assume that the small time-dependent input is sinusoidal:

$$s_1(t) = s_1(0) e^{i\omega t}. \tag{3.12}$$

Instead we at once make the slightly more general specialized assumption that s_1 is a sinusoid multiplied by an exponentially growing envelope:

$$s_1(t) = s_1(0) e^{\sigma t} \text{ where } \sigma = \alpha + i\omega; \tag{3.13}$$

here α determines the growth rate of the growing envelope $\exp(\alpha t)$. Because all of the first-order relations above are both linear and time invariant, the other first-order dynamical quantities generically will respond sympathetically with a similar time dependence:

$$J_1(t) = J_1(0) e^{\sigma t}, \rho_1(t) = \rho_1(0) e^{\sigma t}, r_1(t) = r_1(0) e^{\sigma t}. \tag{3.14}$$

Here the coefficients at time zero will depend on σ and are to be determined by the first-order equations. Since σ is generally complex, they will be complex (and, in fact, analytic functions of σ). Their complex nature will

indicate how their phases in the sinusoidal cycle differ from the phase of the input.

Since equation 3.13 is a specialization of the general time-dependent input perturbation above, all the linear first-order equations remain intact. However the dynamical equation, 3.9, reduces to a simpler form:

$$(\sigma I - Q_0)\,\rho_1 = s_1 Q_s \rho_0, \tag{3.15}$$

and this equation has the formal solution

$$\rho_1 = s_1\,(\sigma I - Q_0)^{-1}\,Q_s \rho_0. \tag{3.16}$$

Now the right-hand expression $Q_s\rho_0$ is to be regarded as a vector, which can be represented as a sum of eigenvectors of the operator Q_0, as in equation 2.28. We further observe that

$$(\sigma I - Q_0)^{-1}\,\phi_n = \frac{1}{\sigma - \lambda_n}\phi_n. \tag{3.17}$$

(Multiply both sides by $\sigma I - Q_0$ and use the eigenvector equation 2.14; equation 3.17 becomes an identity.) This enables us explicitly to solve equation 3.16. With equations 2.28 and 2.29, we have

$$Q_s\rho_0 = \sum_n \left(\widehat{\phi}_n, Q_s\rho_0\right)\phi_n, \tag{3.18}$$

which brings equation 3.16 to the form

$$\rho_1 = s_1 \sum_n \frac{1}{\sigma - \lambda_n}\left(\widehat{\phi}_n, Q_s\rho_0\right)\phi_n. \tag{3.19}$$

If we let $\sigma = i\omega$ here, this expression tells us the frequency response of the population density to a small sinusoidal oscillation in input. If our eigenvalues are of the form expected commonly (see equation 2.42), we see that as we move the driving frequency ω, every time it passes a value $n\omega_0$, there will be a resonant response, and the eigenfunction ϕ_n will become prominent in the population density function.

In equation 3.19 there is no resonance for $\sigma = 0$ even though $\lambda_0 = 0$: the $n = 0$ term in the numerator vanishes, as we now show. By equation 2.16, the equilibrium distribution satisfies

$$0 = \frac{\partial}{\partial s}\,(Q\rho_0) = Q_s\rho_0 + Q\frac{\partial \rho_0}{\partial s}. \tag{3.20}$$

Thus the $n = 0$ numerator in equation 3.19 is

$$(\widehat{\rho}_0, Q_s\rho_0) = \left(\widehat{\rho}_0, -Q\frac{\partial \rho_0}{\partial s}\right) = \left(Q^*\widehat{\rho}_0, -\frac{\partial \rho_0}{\partial s}\right) = 0, \tag{3.21}$$

the last because $\widehat{\rho}_0$ is the zeroth eigenfunction of Q^*. Consequently we can reduce the sum in equation 3.19 to one over $n \neq 0$.

The population firing-rate frequency response, or transfer function, is the (typically complex) ratio of output departure to input departure:

$$T_0(\sigma) = r_1/s_1. \tag{3.22}$$

Because r_1 and s_1 are proportional to the same time course, equations 3.13 and 3.14, their ratio is time independent; because they are linearly related, their ratio is amplitude independent: the transfer function depends on input only through the value of σ. If we substitute equation 3.19 for ρ_1 into 3.11 for r_1, we see that r_1 is indeed proportional to s_1, and the transfer function, equation 3.22, becomes

$$T_0(\sigma) = \widehat{R}[\rho_0] + \sum_{n\neq 0} \frac{\left(\widehat{\phi}_n, Q_s\rho_0\right)}{\sigma - \lambda_n} R[C_0\phi_n]. \tag{3.23}$$

We will loosely call this the population rate frequency response, which more properly is $T_0(i\omega)$, for the more special case of purely oscillatory input. We see again, in the output response, equation 3.23, the appearance of resonance when imaginary σ is brought close to a complex eigenvalue λ_n. The leading term $\widehat{R}$ of equation 3.23 is independent of frequency, and represents accurate tracking even at high frequencies. (However, in a detailed model whose state-space includes synaptic dynamics, the $\widehat{R}$ term goes to zero, as explicated by Gerstner, 1999, and its effect is taken over by large-eigenvalue terms in the sum.)

An example of the frequency response, equation 3.23, was evaluated by a quite different route some years ago (Knight, 1972a). In the case of equation 2.43 ("forgetful integrate-and-fire" model) let the "capacitor discharge rate" γ be small and also let stochastic effects be small. Then encoders, which have simultaneously reset to zero, will disperse only slightly before reaching threshold, and stochastic input current may be replaced by an effective stochastic threshold, which slightly disperses firing times with a small standard deviation τ. The frequency response in this model (Knight, 1972a, equation 8.22, adapted to present notation) is

$$T_0(\sigma) = \left(\frac{\omega_0}{2\pi s_0}\right) + \sum_{n\neq 0} \frac{\gamma\,(\omega_0/2\pi s_0)}{\sigma - \left\{-n^2\left(\frac{\omega_0^3\tau^2}{4\pi}\right) + in\omega_0\right\}}. \tag{3.24}$$

Comparison with equation 3.23 identifies the various terms and relates those of equation 3.23 to "back of the envelope" model parameters. Amplitude and phase dependence on $\sigma = i\omega$ were in agreement with a population frequency response measured (and ω_0 and τ also measured) in an experiment

(Knight, 1972b). The form of the bracketed expression in the denominator of equation 3.24 is in agreement with the expected form of the eigenvalues in equation 2.42. (For this model, as γ and τ go to zero, the radian frequency is $\omega_0 = 2\pi s_0$.) For subsequent convenience, we notationally simplify equation 3.23 to

$$T_0(\sigma) = \widehat{R} + \sum_{n \neq 0} \frac{b_n}{\sigma - \lambda_n}. \tag{3.25}$$

For a multidimensional neuron model, the expedient evaluation of the various constants in equation 3.25 (defined in more detail in equation 3.23) remains a challenge for the future. However, experience to date (Omurtag et al., 1999; Knight et al., 1999; Sirovich et al. 1999) indicates that for one-dimensional neuron models, such evaluations are quite straightforward. In the first section of the appendix, a procedure is outlined to reduce multidimensional neuron models approximately to fairly realistic one-dimensional generalized integrate-and-fire counterparts. Also in section 6.2 a method (so far not implemented) is outlined for the efficient direct numerical evaluation of the coefficients in equations 6.14, which express the population density dynamics in a "moving basis" representation. If these equations are specialized to the sinusoidal perturbation input of equations 3.1 and 3.12, equation 3.25 emerges together with a prescription for numerical evaluation of its various constants.

4 Applications of the Frequency Response: Stability of Interacting Subpopulations and of Feedback

The deliberations above are major tools for the investigation of how interacting subpopulations of neurons behave in consort. In time-independent equilibrium the kth subpopulation will satisfy a generalized form of the population-response relation, equation 2.37:

$$r_k(\mathbf{s}_0) = R_k\left[C_k(\mathbf{s}_0)\,\rho_{k_0}(\mathbf{s}_0), \mathbf{s}_0\right]. \tag{4.1}$$

Here $\mathbf{s}_0$ stands for a list of inputs that may be derived from external sources or from the outputs of other subpopulations. Equation 4.1 has been stated in a way that leaves room for differences not only in the detailed dynamics of each subpopulation but also in the way different input channels deliver synaptic signals to any particular subpopulation. Clearly the dynamical equation, 2.9, may be generalized in a manner similar to equation 4.1:

$$\frac{\partial \rho_k}{\partial t} = Q_k(\mathbf{s}(t))\,\rho_k. \tag{4.2}$$

Again the subscript k refers to the kth subpopulation. The procedures of section 3 may be applied to this more general case, to find the frequency

response of the kth population to the mth, input, as in equation 3.22:

$$r_{k1} = T_{km}(\sigma)\, s_{m1}. \tag{4.3}$$

(Looking back to equations 3.16 and 3.17, we note that the resonant eigenvalues in this expression sensibly depend on the kth dynamical equation but not on the mth input source.)

We still have to specify how the response of one population is to affect the input to another. We will work through some informative simple cases, but initially let us leave this relation very general and write

$$s_m(t) = s_{(ext)m}(t) + u_m\left[\mathbf{r}\left(t'\right), t\right]. \tag{4.4}$$

The first right-hand term is to be regarded as specified external input from other neuron subpopulations, which can be as numerous as the populations under our direct investigation. The second term is a general nonlinear functional, where the prime on t' indicates that this functional can look back across past times t' at earlier outputs of all of the different subpopulations. Thus u_m is general enough to include, for example, distributed propagation time lags from other subpopulations.

In the special state of time-independent equilibrium, equation 4.4 reads

$$s_{m0} = s_{(ext)m0} + u_m\left[\mathbf{r}_0\right], \tag{4.5}$$

where the $s_{(ext)m0}$ are given. These equations simply state how, through back coupling, the time-independent total inputs must depend in a specified way on the time-independent outputs. They are to be solved together with equation 4.1 for the two sets of constants $\mathbf{s}_0, \mathbf{r}_0$. Simple particular cases suggest particular simple strategies, but we note that one standard method for solving equations 4.5 and 4.1 is to replace u_m by βu_m and repeatedly solve while advancing β from $\beta = 0$, by small steps to $\beta = 1$. There remains the question of the stability of the feedback-coupled solution.

The technical means for evaluating a transfer function, discussed in the previous section, are broadly applicable. In particular if we are given a specified $u_m[\mathbf{r}(t'), t]$, and if we assume that the nth input component has the time dependence

$$r_n(t) = r_{n0} + r_{n1}(t) = r_{n0} + r_{n1}(0)\, e^{\sigma t}, \tag{4.6}$$

with r_{n1} small, and we further assume that the remaining components of $\mathbf{r}(t)$ are at their equilibrium values, the general steps of the previous section yield

$$u_m\left[\mathbf{r}\left(t'\right), t\right] = u_m\left[\mathbf{r}_0\right] + U_{mn}(\sigma)\, r_{n1}, \tag{4.7}$$

and the transfer function $U_{mn}(\sigma)$ is straightforward to evaluate in particular cases. If instead we assume that *all* the $r_n(t)$ are of the form of equation 4.6, then equation 4.7 generalizes to

$$u_m\left[\mathbf{r}\left(t'\right), t\right] = u_m\left[\mathbf{r}_0\right] + \sum_n U_{mn}(\sigma)\, r_{n1}, \tag{4.8}$$

which is a component of the matrix equation

$$\mathbf{u}\left[\mathbf{r}\left(t'\right), t\right] = \mathbf{u}\left[\mathbf{r}_0\right] + \mathbf{U}(\sigma)\, \mathbf{r}_1. \tag{4.9}$$

Again in matrix notation, if we assume that the external driving in equation 4.4 is of the form

$$\mathbf{s}_{(ext)}(t) = \mathbf{s}_{(ext)0} + \mathbf{s}_{(ext)1}(0)\, e^{\sigma t}, \tag{4.10}$$

then the first-order part of the feedback equation, 4.4, becomes

$$\mathbf{s}_1 = \mathbf{s}_{(ext)1} + \mathbf{U}(\sigma)\, \mathbf{r}_1. \tag{4.11}$$

The subpopulation input-output relation, equation 4.3, in matrix form becomes

$$\mathbf{r}_1 = \mathbf{T}_0(\sigma)\, \mathbf{s}_1. \tag{4.12}$$

If we now act on the previous equation with the population-response transfer function matrix of equation 4.12, we may express the first-order firing rates of all the subpopulations in terms of those of the external inputs:

$$\mathbf{r}_1 = \mathbf{T}_0(\sigma)\, \mathbf{s}_{(ext)1} + \mathbf{T}_0(\sigma)\, \mathbf{U}(\sigma)\, \mathbf{r}_1. \tag{4.13}$$

This may be solved for the departures from equilibrium of the population rates by a matrix inversion:

$$\mathbf{r}_1 = \left(\mathbf{I} - \mathbf{T}_0(\sigma)\, \mathbf{U}(\sigma)\right)^{-1} \mathbf{T}_0(\sigma)\, \mathbf{s}_{(ext)1}. \tag{4.14}$$

When this expression is written out in components, it tells us the transfer function that expresses the modulation of firing rate, near equilibrium, of any subpopulation, in response to modulation of any one of the external inputs, and includes the effects of cross-talk communicated between different subpopulations as well as the effect of any self-feedback.

It is time to observe the connection between the analytic properties of the transfer function and the question of stability. In general, a modulated output r_1 is related to a modulated input $\widehat{s}_1$ by a (generally complex) ratio

$T(\sigma)$ (as in equation equation 3.22), which allows us to find what *input* is required to achieve a specified *output*:

$$\widehat{s_1} = (T(\sigma))^{-1} r_1. \tag{4.15}$$

If we move σ on the complex plane, the required value of $\widehat{s_1}$ changes. In particular if we move σ so that

$$(T(\sigma))^{-1} = 0, \tag{4.16}$$

we have an output in the absence of input, which is to say that a root of equation 4.16 gives us a free-running behavior of our system, which has a time course proportional to $\exp(\sigma t)$. It may be technically simpler to observe that the free-running exponents of the system occur where the transfer function itself becomes infinite. A case in point is the population firing-rate response transfer function of equation equation 3.25, which goes to infinity at the eigenvalues $\sigma = \lambda_n$. We note from equation 2.39 that the eigenvalues λ_n are the free-running exponents of the system.

We see from equation 4.14 that the free-running exponents for the coupled subpopulations are given by the values of σ, where the inverse matrix goes to infinity, and these are the roots of the equation,

$$\det(\mathbf{I} - \mathbf{T}_0(\sigma)\,\mathbf{U}(\sigma)) = 0. \tag{4.17}$$

This expression is a polynomial in the known transfer functions of the various component pieces of the overall coupled system. If any root of equation 4.17 has a real part that is positive, then the equilibrium determined by equations 4.1 and 4.5 is unstable.

In typical cases, the features of the specific problem may be exploited to find the informative roots of equation 4.17 by means more efficient than a direct determinental expansion. The principal virtue of equation 4.17 is to show that in general, there is a well-defined route to a well-defined goal that answers the question about stability.

5 Stability: Specific Cases

5.1 One Population, Self-Feedback with Fixed Lag. In this case we choose the feedback functional in equation 4.4 to be the linear relation

$$u\left[r\left(t'\right), t\right] = gr\left(t - \widehat{t}\right), \tag{5.1}$$

where g is the gain from output to feedback input, and $\widehat{t}$ is the fixed lag time, so that the feedback equation, 4.4, becomes

$$s(t) = s_{(ext)}(t) + gr(t - \widehat{t}). \tag{5.2}$$

For time-independent equilibrium, equation 4.5 becomes

$$s_0 = s_{(ext)0} + gr_0. \tag{5.3}$$

This is to be solved together with the steady-state population-response equation, 2.37:

$$r_0 = R\left[C(s_0)\rho_0(s_0), s_0\right] \equiv R_0(s_0), \tag{5.4}$$

which is a nonlinear relation between r_0 and s_0. Thus the determination of s_0 in terms of $s_{(ext)0}$ cannot be done in a simple and exact closed form. But to understand the basic facts about the static behavior of our population response, we should already have on hand the evaluation of the function $R_0(s_0)$ over a range of s_0 values, and so we can solve equations 5.3 and 5.4 together to express the s_0 versus $s_{(ext)0}$ relation as

$$s_{(ext)0} = s_0 - gR_0(s_0). \tag{5.5}$$

Values of $s_{(ext)0}$ that are outside the range of positive values given by equation 5.5 will not allow solution of the steady-state equations 5.3 and 5.4. We make the cautionary observation that in equation 5.5, any steady-state total input s_0 may be excluded from the solvable region by making the positive feedback g strong enough. Clearly our discussion applies to the remaining broad range of cases where the time-independent solution exists.

If we let

$$r(t) = r_0 + r_1(t) = r_0 + r_1(0)\, e^{\sigma t}, \tag{5.6}$$

our specialization equation, 5.1, becomes

$$u\left[r_0 + r_1\left(t'\right), t\right] = gr_0 + gr_1(0)\, e^{\sigma\left(\widehat{t-t}\right)} = gr_0 + ge^{-\sigma\widehat{t}}r_1(t), \tag{5.7}$$

so in this case the transfer function of equation 4.7 is

$$U(\sigma) = ge^{-\sigma\widehat{t}}. \tag{5.8}$$

The first-order part of equation 5.2 is

$$s_1 = s_{(ext)1} + U(\sigma)\, r_1, \tag{5.9}$$

while from equation 3.22,

$$T_0(\sigma)\, s_1 = r_1. \tag{5.10}$$

We multiply equation 5.9 by $T_0(\sigma)$ and find

$$r_1 = \frac{1}{1 - T_0(\sigma)\, U(\sigma)} T_0(\sigma)\, s_{(ext)1} \tag{5.11}$$

for the population rate frequency response to external input. (Note that this is the 1×1 matrix case of the matrix equation, 4.14.) By equation 4.16, the free-running response exponents of the system will be those values of σ for which the denominator vanishes, so that

$$T_0(\sigma)\, U(\sigma) = 1, \tag{5.12}$$

or in our particular case (equation 5.8),

$$e^{-\sigma \widehat{t}} T_0(\sigma) = \frac{1}{g}, \tag{5.13}$$

a transcendental equation to be solved for σ, where $T_0(\sigma)$ is the explicit expression (equation 3.25).

A specialized case will show some qualitative features. In the simple case of self-feedback without delay ($\widehat{t} = 0$), equation 5.13 becomes

$$T_0(\sigma) = \frac{1}{g}. \tag{5.14}$$

If the feedback gain g is small, then a root σ must make $T_0(\sigma)$ large; this will happen if σ lies near a pole λ_n of $T_0(\sigma)$. As a first approximation we may truncate the sum in equation 3.25 to the term that is large because σ lies near its pole:

$$T_0(\sigma) \cong \frac{b_n}{\sigma - \lambda_n}, \tag{5.15}$$

and equation 5.14 may now be solved to give

$$\sigma = \lambda_n + g b_n. \tag{5.16}$$

If b_n is real and positive (as in the explicit example of $T_0(\sigma)$ given by equation 3.24) then, for positive feedback g, equation 5.16 shows that the real part of σ is less negative than that of λ_n, and also that increasing g moves the real part of σ in the positive direction and eventually past zero and beyond the threshold of instability. More generally, to find the border of instability, we may set the real part of σ to zero and take the real part of equation 5.16 to find the critical feedback value,

$$g_{crit} = |\mathrm{Re}\, \lambda_n| \,/\, \mathrm{Re}\, b_n, \tag{5.17}$$

to within the approximation of equation 5.15. The system will be unstable if *any* root of equation 5.14 has a positive real part, so in equation 5.17, we must choose the index n that makes g_{crit} smallest. Equation 2.42 suggests that commonly this will occur for $n = 1$.

Returning to equation 5.13, for finite lag $\widehat{t}$ the "small g-single pole" approximation, equation 5.15, still makes good sense; instead of equation 5.16, it now gives us the implicit relation

$$\sigma = \lambda_n + e^{-\sigma\widehat{t}} b_n g. \tag{5.18}$$

Again let us observe that a critical value of g yields a purely imaginary σ at the edge of instability. At that value take the real part of equation 5.18, and instead of equation 5.17 we find

$$g_{crit}\left(\widehat{t}\right) = \frac{1}{\cos\left(|\sigma|\,\widehat{t}\right)} \frac{|\mathrm{Re}\,\lambda_n|}{\mathrm{Re}\,b_n}, \tag{5.19}$$

and (because the cosine cannot exceed unity) the presence of the lag tends to stabilize positive feedback. Equation 5.18 may be used in this way if $g_{crit}(\widehat{t})$ is still small, which should hold if $|\,\mathrm{Re}\,\lambda_n|$ is small, which should be the case if the influence of stochastic effects is modest.

5.2 One Population, Distributed Feedback Lags. Suppose that instead of a single fixed lag $\widehat{t}$ as in equation 5.1, we have a distribution of lags weighted according to a distribution function $P(t')$, which integrates to unity. In place of equation 5.1, our feedback functional now is

$$u\left[r\left(t'\right)\right] = g\int_0^\infty dt' P\left(t'\right) r\left(t - t'\right), \tag{5.20}$$

which is still linear in r, and where g again is the feedback gain. Substitution of a constant r_0 in equation 5.20 retrieves exactly the same results as in the previous exercise, since $P(t')$ integrates to unity. If we assume a superimposed small modulation,

$$r_1(t) = r_1\,(0)\, e^{\sigma t}, \tag{5.21}$$

then equation 5.20 gives

$$\begin{aligned} u\left[r_1\left(t'\right), t\right] &= g\int_0^\infty dt' P\left(t'\right) r_1\,(0)\, e^{\sigma\left(t-t'\right)} \\ &= g r_1(t)\int_0^\infty dt' e^{-\sigma t'} P\left(t'\right) = U\,(\sigma)\, r_1(t), \end{aligned} \tag{5.22}$$

and the explicit feedback transfer function,

$$U(\sigma) = g \int_0^\infty dt' e^{-\sigma t'} P\left(t'\right) = g\widehat{P}(\sigma)\,, \tag{5.23}$$

may be substituted into equation 5.12, whose roots give us the free-running frequencies. An example of such a situation is given by Abbott and van Vreeswijk (1993).

If the feedback coupling g is small enough, we may again use the "small g-single pole" approximation, equation 5.15. Since in this case σ lies near λ_n, to lowest order in g we may replace $U(\sigma)$ by $U(\lambda_n)$, and equation 5.12 may be solved for the free-running exponents,

$$\sigma_n = \lambda_n + b_n U(\lambda_n) = \lambda_n + g b_n \widehat{P}(\lambda_n)\,. \tag{5.24}$$

Here the first right-hand form makes no assumption concerning the nature of the feedback transfer function $U(\sigma)$, except that its effect should be small enough to permit the single-pole approximation. In the second right-hand form, we may exploit the fact that $P(t')$ is positive in equation 5.23. If, further, $P(t')$ is concentrated around small values of t', then for small integer $|n|$, equation 5.23 tells us that $U(\lambda_n)$ will have a real part that is positive. Equation 3.24 suggests (as do some detailed analytic arguments, as well as numerical evidence; Sirovich et al., 1999a) that the real part of b_n (for small $|n|$) should be positive as well. Consequently the second right-hand form of equation 5.24 indicates that, again in this case, more positive feedback leads to real parts of the σ_n that are less negative, and so leads to more persistent free-running modes of motion.

Again, the borderline of instability may be estimated. If we assume that σ_1 is pure imaginary and take the real part of equation 5.24, we obtain

$$g_{crit} = \frac{|\mathrm{Re}\,\lambda_1|}{\mathrm{Re}\left(b_1\widehat{P}(\lambda_1)\right)}\,, \tag{5.25}$$

which should have quantitative accuracy if g_{crit} is small enough to fulfill the single-pole approximation.

5.3 A Multipopulation Generalization for Small Feedback. If feedbacks are small among interacting subpopulations of neurons, the free-running exponent equation, 4.17, has an instructive and valuable approximate analytic solution that can be obtained by the single-pole approximation. The determinant of a matrix vanishes if that matrix has a zero eigenvalue, and so equation 4.17 will be solved by any value σ that solves the matrix equation,

$$\left(\mathbf{I} - \mathbf{T}_0(\sigma)\,\mathbf{U}(\sigma)\right)\mathbf{w} = \mathbf{0}. \tag{5.26}$$

A typical matrix element of $\mathbf{T}_0(\sigma)$ is

$$T_{km}(\sigma) = \sum_{n \neq 0} \frac{b_n^{(km)}}{\sigma - \lambda_n^{(k)}}, \tag{5.27}$$

where the row k indexes the kth subpopulation and the column m the mth input, while n indexes the nth free-running mode of one subpopulation in isolation. (The fact that a subpopulation's free-running eigenvalues depend on which subpopulation, but not on which input, is reflected in equation 5.27.) There is a pole for every eigenvalue of every subpopulation. If the matrix elements of $\mathbf{U}$ are small, then $\mathbf{T}_0$ must have large elements to compensate, and so, in equation 5.27, σ must lie near some $\lambda_n^{(k)}$, let us say for the qth subpopulation, and all other poles may be ignored. The matrix elements (see equation 5.27) become

$$T_{km}(\sigma) = \frac{b_n^{(qm)}}{\sigma - \lambda_n^{(q)}} = \frac{v_m}{\sigma - \lambda_n^{(q)}} \quad \text{if } k = q, = 0 \text{ otherwise,} \tag{5.28}$$

and the second right-hand form defines the elements of the specified vector:

$$\mathbf{v}\text{: } v_m = b_n^{(qm)}. \tag{5.29}$$

Let us specify another vector:

$$\widehat{\boldsymbol{q}}\text{: } \widehat{q}_k = 1 \quad \text{if } k = q, = 0 \text{ otherwise.} \tag{5.30}$$

Then (with the superscript T for transpose) equation 5.28 may be expressed as the outer-product matrix,

$$\mathbf{T}_0(\sigma) = \frac{1}{\sigma - \lambda_n^{(q)}} \left(\widehat{\mathbf{q}} \otimes \mathbf{v}^T \right). \tag{5.31}$$

If σ lies close to $\lambda_n^{(q)}$, then, in equation 5.26, to lowest order $\mathbf{U}(\sigma)$ may be replaced by $\mathbf{U}(\lambda_n^{(q)})$, and equation 5.26 becomes

$$\mathbf{w} - \frac{1}{\sigma - \lambda_n^{(q)}} \left(\widehat{\mathbf{q}} \otimes \mathbf{v}^T \right) \left(\mathbf{U} \left(\lambda_n^{(q)} \right) \right) \mathbf{w} = 0. \tag{5.32}$$

Performing the right-hand operations and multiplying by the denominator:

$$\left(\sigma - \lambda_n^{(q)} \right) \mathbf{w} = \widehat{\mathbf{q}} \left(\mathbf{v}, \mathbf{U} \left(\lambda_n^{(q)} \right) \mathbf{w} \right), \tag{5.33}$$

where the right-hand parenthesis is the scalar product of $\mathbf{v}$ with $\mathbf{Uw}$. Equation 5.33 shows us that the components of the unknown vector $\mathbf{w}$ are proportional to the components of the known vector $\widehat{\mathbf{q}}$, so the eigenvector of equation 5.26 is

$$\mathbf{w} = \widehat{\mathbf{q}}, \tag{5.34}$$

and we may equate the coefficients of equation 5.33 to obtain the exponent σ:

$$\sigma = \lambda_n^{(q)} + \left(\mathbf{v}, \mathbf{U}\left(\lambda_n^{(q)}\right)\widehat{\mathbf{q}}\right) = \lambda_n^{(q)} + \sum_m b_n^{(qm)} U_{mq}\left(\lambda_n^{(q)}\right), \tag{5.35}$$

where equations 5.29 and 5.30 were used, in the last step, to evaluate the scalar product. Equation 5.35 tells us how far the exponent σ is moved from the nearby eigenvalue $\lambda_n^{(q)}$, by the interpopulation coupling U_{mq}.

This subsection has shown how the single-population results of the previous subsection generalize naturally to a situation with interacting subpopulations. In a single small patch of cerebral cortex, we find subpopulations of neurons of different dynamical design, and we also find, in separated patches, subpopulations of neurons with the same dynamical design. This subsection was intended, together with the next subsection, as a guide to show how a model that is a caricature of cortex, with subpopulations that interact and have separate inputs, but are of only a single dynamical type, may be generalized to the more realistic situation, which involves several different dynamical types.

5.4 Caricature of an Orientation Hypercolumn. In primary visual cortex, nearby patches of neurons typically respond best to linear visual stimuli that have similar orientations. There is neural interaction between the patches, and this interaction is weaker between pairs of patches whose preferred orientations are more dissimilar. Approximately, patches for all preferred orientations are on an equal footing, with no preferred orientation exciting a larger subpopulation of neurons than others do or differing from others in design characteristics. A set of nearby patches that includes all orientations is referred to as a hypercolumn. If a hypercolumn contained neurons of only one dynamical type, and contained N subpopulations with equally spaced preferred orientations, then the following caricature would be accurate:

A common dynamical equation,

$$\frac{\partial \rho_k}{\partial t} = Q\left(s_k(t)\right)\rho_k, \tag{5.36}$$

for the kth subpopulation.

A common firing-rate rule

$$r_k(t) = R\left[C\left(s_k(t)\right)\rho_k, s_k(t)\right]. \tag{5.37}$$

An even-handed feedback rule,

$$s_k(t) = s_{(ext)k}(t) + \sum_n \beta_n u\left[r_{k-n}\left(t'\right), t\right]. \tag{5.38}$$

Here the summed index is understood to be modulo N, so no subpopulation has preference. Since u has been left general, we may assume that the weightings β_n sum to unity:

$$\sum_n \beta_n = 1. \tag{5.39}$$

Let us also reasonably assume that β_n is always positive and is even ($= \beta_{-n}$). For simplicity we assume that u is a linear functional, as in the earlier subsections.

If we assume a time-independent baseline external input that also is even-handedly independent of k, then it is consistent to assume that the resulting total inputs and the subpopulation responses are independent of k as well, and we retrieve, for the N subpopulations the equilibrium equations, 5.3, 5.4, and 5.5, which held for a single population with feedback.

Proceeding as before with the small-departure analysis, from our earlier results we easily derive from equations 5.36 and 5.37 that

$$r_{k,1} = T_0(\sigma)\, s_{k,1}, \tag{5.40}$$

while equation 5.38 gives us

$$s_{k,1} = s_{(ext)k,1} + U(\sigma) \sum_n \beta_n r_{k-n,1}. \tag{5.41}$$

To relate population response to external input, we may multiply equation 5.41 by $T_0(\sigma)$ and use equation 5.40:

$$r_{k,1} = T_0(\sigma)\, s_{(ext)k,1} + T_0(\sigma)\, U(\sigma) \sum_n \beta_n r_{k-n,1}. \tag{5.42}$$

This is a special case of the matrix equation, 4.13. It is much simplified because the single population dynamics results in a population transfer function matrix, which is a single scalar transfer function that multiplies the identity matrix, while the feedback matrix is a single scalar function of σ which multiplies a discrete convolution.

The presence of a discrete convolution suggests that a discrete-Fourier-transform strategy should yield major simplifications, and this is indeed

the case. In fact discrete Fourier transformation of our nonlinear starting equations, 5.36–5.38, gives a reduction in numerical work by isolating a collection of "nonlinear dynamical modes," which can be much less than N in number and also carry a valuable physiological interpretation.

The stability question, which we address here, is expedited by a small bit of the discrete-Fourier machinery. We have seen that the free-running exponents of the coupled system may be found as the values of σ that produce a zero eigenvalue and so solve the matrix equation, 5.26. Corresponding to equation 5.42, that matrix eigenvalue equation specializes to

$$w_k - T_0(\sigma) U(\sigma) \sum_n \beta_n w_{k-n} = 0. \tag{5.43}$$

Because of the discrete convolution, we can anticipate that the eigenvector $\mathbf{w}$ will be a discrete-Fourier basis vector:

$$w_k = e^{iqk} \quad \text{where } q = 2\pi j/N, \tag{5.44}$$

and where j can be any integer in the range that includes zero to $N-1$. To verify this, substitute equation 5.44 into 5.43 and at once reduce that equation to

$$1 - T_0(\sigma) U(\sigma) \widetilde{\beta}(q) = 0, \tag{5.45}$$

where

$$\widetilde{\beta}(q) = \sum_n e^{-iqn} \beta_n \tag{5.46}$$

is the discrete Fourier transform of β_n. We note for $j = 0$ that

$$\widetilde{\beta}(0) = 1, \tag{5.47}$$

by equation 5.39. Because we chose the β_n real, positive, and even in n, we see that $\widetilde{\beta}$ is a real number and

$$\widetilde{\beta}(q) \leqq 1. \tag{5.48}$$

Thus equation 5.45, which is to be solved for σ, is identical in form to equation 5.12, which was for a single-neuron population with the same population dynamics and the same feedback dynamics. For $q \neq 0$, the presence of $\widetilde{\beta}(q)$ in equation 5.45 is simply to reduce the strength of the effective feedback coupling according to equation 5.48. Thus, we anticipate that the

least stable exponent will go with $q = 0$, which gives us exactly the single-population case, equation 5.12), discussed earlier.

Three further points should be made here. The first is that discrete subpopulations appeared in equations 5.36–5.38 in anticipation of numerical work to come, while a more realistic hypercolumn model would feature a continuous gradation in best-orientation sensitivity. But a free-running exponent equation, 5.45, has emerged as an effective single-population relation whose form is independent of how finely we discretize the system. We could have started equally well with a continuously graded system, and the same effective one-dimensional result would have followed. The simplification of equation 5.45 was a consequence of an underlying symmetry of the system. We could have started with the symmetry of the circle, which is what the problem gave to us, rather than reverting to the less natural symmetry of an N-sided regular polygon, which is what the *discrete* Fourier transform addresses.

The second point is that we could have undertaken a more realistic hypercolumn model involving several different cell types, each with a population dynamics differing in its details. All the steps above would have their counterparts in this more realistic case. At the final step we would have, instead of the single equation 5.45, a matrix eigenvalue equation like 5.26. But the number of entries in the eigenvector would be *not the large number of subpopulations but rather the small number of dynamically different cell types*. Thus, the stability problem still is very tractable for this case.

The third point is in a more speculative vein. Recent evidence (Everson et al. 1998) suggests that primary visual cortex has a single processing strategy and even-handed allocation of resources for the processing of patches of retinal image that differ not only in orientation but also in magnification scale over a fair range of magnifications, to reasonable approximation. The combined two-parameter ("twist" and "zoom") symmetry has a set of two-parameter eigenvectors analogous to the "discrete twist" eigenvector (see equation 5.44). These mathematical constructs might not only assist more realistically detailed modeling of visual cortex, but perhaps also might provide some new insight into the nature of cortical processing.

6 Possibilities for Efficient Simulation

6.1 Principal Components. Section 1 concluded with the mention of a means to reduce the dynamical equation, 2.9, and the firing-rate equation, 2.13, to matrix equations of fairly modest size. This may be achieved, once we have on hand a sample solution of equation 2.9 in response to a span of "typical input" $s(t)$, by principal component analysis (Sirovich & Everson, 1992) (also called "singular value decomposition" or "representation by empirical eigenfunctions"), which now will be briefly described. For rea-

sonable input $s(t)$, some dynamical equations of the form 2.9 prove to be almost unable to carry the population density function $\rho(t)$ outside a subspace of modest dimension that is embedded in its full function space. If at a time t^* we have a snapshot $\rho(t^*)$ of the density function, we may create from it the singular operator

$$P_{t^*} = \rho\left(t^*\right) \otimes \rho\left(t^*\right)^T, \tag{6.1}$$

which is a projection operator in the sense that its action on any vector u gives

$$P_{t^*} u = \rho\left(t^*\right)\left(\rho\left(t^*\right), u\right), \tag{6.2}$$

which is a vector in the one-dimensional manifold defined by the snapshot vector $\rho(t^*)$. From equation 6.1 we immediately show that P_{t^*} is self-adjoint. Because

$$P_{t^*}\rho\left(t^*\right) = \rho\left(t^*\right)\left(\rho\left(t^*\right), \rho\left(t^*\right)\right), \tag{6.3}$$

it has only nonnegative eigenvalues and is positive-semidefinite. If we have a collection of snapshots, taken at times t^*, we may form the operator

$$A = \sum_{t^*} P_{t^*}, \tag{6.4}$$

which, according to equation 6.2, maps any vector into the modest-dimensional manifold accessible to the dynamical equation.

The self-adjoint property, and the positive-semidefinite property, are inherited by A, and this ensures a set of orthonormal eigenvectors in the eigenvalue problem

$$Aw_n = \lambda_n w_n, \tag{6.5}$$

whose eigenvectors we can number in the order of their decreasing nonnegative eigenvalues. Because the eigenvectors are orthonormal, we have

$$(w_m, w_n) = \delta_{mn}. \tag{6.6}$$

Because the effectively nonnull subspace of A is of modest dimension, to good approximation we may span it with a modest subset of the eigenvectors to equation 6.5.

Now we can expand a solution $\rho(t)$ as

$$\rho(t) = \sum_n a_n(t) w_n, \tag{6.7}$$

where the sum extends over a modest number of terms (17 terms in the quoted example). This we now can substitute into the dynamical equation, 2.9. We take the inner product of that equation with w_m and (employing equation 6.6 on the left-hand side) directly conclude that

$$\frac{da_m}{dt} = \sum_n (w_m, Q(s(t))\, w_n)\, a_n, \tag{6.8}$$

which may be integrated in time for the $a_m(t)$. With these $a_m(t)$ on hand, substitution of equation 6.7 into the firing-rate equation, 2.13, explicitly evaluates the firing rate as

$$r = \sum_n R[Cw_n]\, a_n. \tag{6.9}$$

We note that the a_n may be regarded as the components of a vector, and this endows equations 6.8 and 6.9 with an inherited structure identical to that of equations 2.9 and 2.13. Equation 6.9 is an explicitly specified linear functional of its input vector. Equation 6.8 expresses the evolution of that vector in terms of the action on it, of an s-dependent operator—here an s-dependent matrix. If the operator in equation 2.9 can be expressed as a polynomial in s (as, for example, by substituting equation 2.43 into equation 1.4), then the matrix in equation 6.8 can be expressed as a corresponding polynomial in which the same set of powers of s multiplies corresponding component constant matrices.

These considerations translate directly into a practical and efficient computational scheme. However, this procedure requires caution at one point: the operator Q has conservation of probability built into it, but its representation by a truncated matrix in equation 6.8 may slightly miss perfect conservation because of the truncation; the formerly zero eigenvalue may be slightly perturbed, and this will give to the total probability a spurious slow exponential drift as time progresses.

Conservation of probability is easily restored to the matrix of equation 6.8. In discretized format, conservation for the original matrix Q was assured by the fact that every column of Q was orthogonal to the zeroth adjoint eigenvector $\widehat{\rho}_0$, which had constant entries. In the present representation the nth component of this vector is given by

$$\widehat{a}_n = (w_n, \widehat{\rho}_0)\,. \tag{6.10}$$

Without truncation, the admixture of this vector in each column of the matrix in equation 6.8 would be zero. Let us inner-product the truncated equation 6.10, with a column of the truncated matrix, to obtain the (very small) coefficient of actual admixture; we then multiply the vector (equation 6.10) by this admixture coefficient and subtract the result from the matrix column.

Conservation of probability is now fulfilled. In cases tested (Sirovich et al., 1999c) the resulting matrix applied in equations 6.8 and 6.9 gives excellent agreement with the much longer direct calculation of equations 2.9 and 2.13. The choice of 17 dimensions rendered graphical results indistinguishable from those of the longer direct calculation. (This was not quite achieved with 15 dimensions. The dimension count jumps by twos because the natural subspaces relate to consecutive pairs of complex conjugate eigenvalues, as the next section will elucidate.)

6.2 Moving Basis. We conclude with a brief comment on a set of possibilities that might lead to extremely efficient simulation and give an additional viewpoint on mathematical structure. Once the efficient principal components representation has been constructed, it is not a heavy computation to solve the eigenvalue problem and tabulate the results, parametrically in the input s, for the eigenvalues and eigenfunctions that are within the subspace that the dynamical equation is able to access. In cases examined, these have proved to be in good practical agreement with their counterparts derived from the more laborious direct calculation (Sirovich et al., 1999c). Thus, it is reasonable to envision a representation of the dynamics expressed in terms of the input-dependent dynamical eigenvectors of equations 2.14 and 2.15:

$$\rho(t) = \sum_n \widetilde{a}_n(t)\phi_n\left(s(t)\right). \tag{6.11}$$

This is a representation in terms of a "moving basis" that at each time t remains optimized for the moving present value of the input $s(t)$. (The overhead mark is to distinguish these coefficients from those of equation 6.7.) Effective convergence thus should be achieved by a sum of fewer terms in equation 6.11 than were required in equation 6.7, where the fixed basis was a compromise that had to span the first several eigenvectors of $Q(s)$ over the entire range of $s(t)$.

Once the $\widetilde{a}_n(t)$ in equation 6.11 have been determined, the population firing rate is given by the short sum

$$r = \sum_n \widetilde{a}_n R\left[C(s)\phi_n\left(s\right), s\right] = \sum_n \widehat{R}_n(s)\widetilde{a}_n(t). \tag{6.12}$$

To determine the $\widetilde{a}_n(t)$, we may work from the differential equations that they satisfy. To do this we may inner-product $\rho(t)$ in equation 6.11 with the adjoint eigenvector $\widehat{\phi}_m(s)$, to obtain

$$\widetilde{a}_m(t) = \left(\widehat{\phi}_m\left(s\left(t\right)\right), \rho(t)\right). \tag{6.13}$$

Differentiation now yields

$$\begin{aligned}\frac{d\widetilde{a}_m}{dt} &= \left(\widehat{\phi}_m, \frac{\partial \rho}{\partial t}\right) + \left(\widehat{\phi}_{m,s}\frac{ds}{dt}, \rho\right) \\ &= \left(Q^*\widehat{\phi}_m, \rho\right) + \frac{ds}{dt}\left(\widehat{\phi}_{m,s}, \rho\right) \\ &= \sum_n \widetilde{a}_n \left\{\left(\lambda_m^*\widehat{\phi}_m, \phi_n\right) + \frac{ds}{dt}\left(\widehat{\phi}_{m,s}, \phi_n\right)\right\} \\ &= \lambda_m\left(s(t)\right)\widetilde{a}_m + \frac{ds}{dt}\sum_n M_{mn}\left(s(t)\right)\widetilde{a}_n. \end{aligned} \tag{6.14}$$

In the second line we used the dynamical equation and the adjoint operator relation; in the third we substituted equation 6.11 for ρ; in the fourth line we abbreviated by defining the matrix

$$M_{mn}(s) = \left(\partial\widehat{\phi}_m(s)/\partial s, \phi_n(s)\right) \tag{6.15}$$

(whose $m = 0$ entries are all zero, because of the constant-entry property of $\widehat{\rho}_0$, as discussed at equations 2.35 and 2.36). The fourth line of equation 6.14 gives an explicit system of ordinary differential equations to be integrated for the $\widetilde{a}_m(t)$. A check on the equations 6.14 is that if s is constant, they may be solved analytically, and their solution is equation 2.32. Because for $m \neq 0$ the eigenvectors ϕ_m do not carry any probability (see the discussion at equation 2.36) we must have

$$\widetilde{a}_0(t) \equiv 1. \tag{6.16}$$

(by equation 6.11) and this is consistent with equation 6.14 even when $s(t)$ is time dependent, since, when $m = 0$, both λ_m and M_{mn} are zero.

If $s(t)$ changes very slowly (see equation 2.38), then $\rho(t)$ will pass through a progression of steady states, and we will have

$$\rho(t) = \rho_0\left(s(t)\right). \tag{6.17}$$

The remaining members of equation 6.14 furnish a sequence of corrections to this result. We can estimate how many terms we will need in the sums in equations 6.11 and 6.14. If our eigenvalues are of the "damped limit-cycle" form (see equation 2.42), then the equations 6.14 are those for a sequence of "tuned filters," each tuned for a radian frequency given by the imaginary part of the eigenvalue, and each driven by the right-hand-most term in the equation. We see that the pair of equations for m and $-m$ are needed to furnish the part of the response spectrum in a frequency range near m times the limit cycle frequency. If we wish to compute an output response that

is accurate through frequency components that are N times as high as the highest single-neuron firing rate, then the estimated number of terms in the sums in the firing rate and dynamical equations 6.12 and 6.14 will be $2N+1$. Thus, for example, seven terms should yield a computed output with a short-term-estimated spectrum that remains accurate through frequencies about three times the momentary single-neuron firing rate.

Direct application of the moving basis, as just discussed, to a simulation in which the value of the input $s(t)$ passes from the limit-cycle regime of an individual neuron to the regime of the stable equilibrium point, will encounter a technical nuisance. (This situation will arise in the simulation of primary visual cortex.) In the region of passage between the two regimes, each pair of early eigenvalues $\lambda_{\pm m}(s)$, at some critical value of s, will merge by having their imaginary parts move to zero, and further change in s will give rise to two negative real eigenvalues that correspond to a faster and a slower pure relaxation. The members of equation 6.14 for $\pm m$ together may be written

$$d\widetilde{a}_{+m}/dt = (\lambda_{+m} \cdot \widetilde{a}_{+m} + 0 \cdot \widetilde{a}_{-m}) + \left(ds/dt\right) L_{+m}\left(\widetilde{a}_{+}, \widetilde{a}_{-}, s\right)$$

$$d\widetilde{a}_{-m}/dt = (0 \cdot \widetilde{a}_{+m} + \lambda_{-m} \cdot \widetilde{a}_{-m}) + \left(ds/dt\right) L_{-m}\left(\widetilde{a}_{+}, \widetilde{a}_{-}, s\right) \tag{6.18}$$

to emphasize that here we are dealing with a 2×2 diagonal matrix. The $L_{\pm}$ terms are the remaining linear sums in equation 6.14. Clearly we can pursue equation 6.18 into the pure-relaxation regime by assigning $+m$ and $-m$, respectively, to the less and to the more negative real eigenvalues there. But the transition is not smooth at the merger value, as the motion of each eigenvalue (for example) on the complex plane takes a right-angle turn at that value of s. The two invariants of the matrix, its determinant $\triangle$ and its trace T, are related to the eigenvalues by

$$\triangle_m(s) = \lambda_{+m}(s) \cdot \lambda_{-m}(s) \tag{6.19}$$

$$T_m(s) = \lambda_{+m}(s) + \lambda_{-m}(s) \tag{6.20}$$

$$\lambda_{\pm m}(s) = \frac{1}{2}\left\{T_m(s) \pm \sqrt{(T_m(s))^2 - 4\triangle_m(s)}\right\}. \tag{6.21}$$

The critical condition occurs at the value of s where the argument of the root passes through zero and then changes sign. We observe that the right-angle-turn behavior of the eigenvalues is not shared by either of the two invariants, and this indicates that the passage may be made smooth by a

linear transformation upon equation 6.18 that carries it to the form

$$dx_m/dt = \left(0 \cdot x_m + 1 \cdot y_m\right) + \left(ds/dt\right) X_m\left(x, y, s\right) \tag{6.22}$$

$$dy_m/dt = \left(-\triangle_m \cdot x_m + T_m \cdot y_m\right) + \left(ds/dt\right) Y_m\left(x, y, s\right) \tag{6.23}$$

whose 2×2 matrix has the same determinant and trace as in equation 6.18. Equation 6.22 may be solved for y_m and substituted into equation 6.23 to remove explicit dependence on y_m from that equation's right-hand side. An additional differentiation of equation 6.22 then carries the form of the dynamics to that of the fairly familiar damped and forced harmonic oscillator. The critical value of s corresponds to critical damping in the oscillator, which does not introduce singular behavior into the oscillator's forced motion. The moving-basis equations in the real-component form, 6.22 and 6.23, may be forward integrated across the critical damping condition without encountering embarrassment.

The introduction of the moving basis in terms of sharpening the results of principal component analysis at the start of this subsection was convenient but inessential. The evaluated pieces that go into equations 6.12 and 6.14, which are $\widehat{R}_n(s)$, $\lambda_m(s)$, and $M_{mn}(s)$, are all to be evaluated at constant input s. If state-space is divided up into small compartments, then the only compartments visited by a neuron at constant s will be those that lie within a "fattened limit cycle" or in the case of an s so chosen that the neuron without noise would be quiescent, will lie analogously within the "stochastically fattened neighborhood of a stable equilibrium point." These small compartment subsets of the state-space offer a numerically tractable eigenvalue problem and direct evaluation of the pieces just mentioned.

The standard sinusoidal small perturbation assumption, as stated in equations 3.1 and 3.12, can be substituted into the moving-basis dynamical equations 6.14. Since equation 6.14 contains no prior approximation, the transfer function (see equation 3.25) must follow, and does so with a little thoughtful manipulation. Hence the numerically tractable eigenvalue problem will in principle lead to a similarly tractable numerical evaluation of the various constants that appear in equation 3.25.

Equations 6.12 and 6.14 summarize the dynamics of a neuron population in a sort of a compact canonical form, in the sense that in principle, two dynamically distinct neuron designs can lead to numerically identical examples of equations 6.12 and 6.14, but if they do then both alternative designs would play dynamically identical roles as subpopulations in a larger array of interacting neuron subpopulations. If the effect on a specific neuron population by an elsewhere-directed pharmacological agent is to be minimal, or if the effect of a base-pair change in a mouse gene is to be maximal, then this more properly could be measured by minimal or maximal numerical changes in equations 6.12 and 6.14 rather than in the underlying dynamical equations. Regarding efficient simulation, we can

envision an elaborate neuron model with multiple electrical compartments, and a highly simplified neuron of the general integrate-and-fire variety, giving rise to numerically almost identical instances of equations 6.12 and 6.14.

7 Summary

We may anticipate that a fairly broad range of useful neuron-population simulations will exhibit a common mathematical structure. The main ingredient of that structure is a dynamical equation that describes the momentary evolution of a population density. That density exists in a state-space whose points represent the possible states of a single model neuron. The population density dynamical equation follows from the postulated dynamics of the single neuron, including the consequences of an input noise uncorrelated among different neurons of the population. The population dynamics equation is nonlinear in the common inputs from other neuron populations (which nonlinearity reflects the nonlinear dynamics of the identical individual neurons in the population) but is linear in the population density. This linearity endows the population dynamics with a structure whose broad outlines are familiar from applied mathematics and which gives insight regarding the ways the neuron population will respond to different input situations.

The second major ingredient of a neuron population dynamics model is an expression that specifies its momentary per-neuron output rate of impulses. This output rate expression depends linearly on the momentary disposition of the population density. (In the sense that the impulse train from each neuron is a point process in time, the merged impulse trains from a large population of noninteracting neurons will be a time-dependent Poisson point process.)

Because the population dynamical equation is linear in the population density, it may be viewed in terms of an input-dependent population dynamical linear operator that acts on the population density. At a specified input, that operator has a set of eigenfunctions in terms of which the population density can be expanded as a linear weighted sum and a corresponding set of eigenvalues that include zero and complex conjugate pairs with negative real parts. Under common circumstances, the eigenvalue pairs nearest zero may be approximated by a simple orderly progression, and we may anticipate that the corresponding eigenfunctions will dominate in the expansion of the population density.

We may examine how such a population dynamics model responds to input that is a small sinusoidal oscillation around an operating point. This leads to an input-output frequency response that is a simple analytic function of frequency with "resonance denominators" at the complex eigenvalues that, under common circumstances, lie near integer multiples of the population's mean firing rate.

The output of one population can be used as input to another. If several neuron subpopulations are thus coupled together, commonly a set of equilibrium firing rates may be found. A frequency response function can be assigned to each input-output pair; from these we may derive an analytic function that is a polynomial combination of those frequency responses, and whose complex roots specify the possible free-running behaviors near equilibrium of the coupled set of subpopulations. A root whose real part is positive indicates a free-running behavior that is a sinusoidal oscillation times an exponentially growing envelope; hence the equilibrium is unstable in this case.

In particular cases this stability analysis is quite tractable. The question of the stability of a model of an orientation hypercolumn of neurons in primary visual cortex may be reduced by symmetry to the tractable stability analysis of a much smaller system, which contains only one subpopulation of each cortical cell type.

Finally, because only fairly few early eigenfunctions proved important in exploring the analytic methods discussed above, we are led to a preliminary examination of how this feature might enable extremely efficient simulations of neuron population dynamics. One such method involves examination of the solution of the full population dynamics equation in response to a sample of typical variable input. The small subspace of function space, in which the population density always resides, may be determined by principal component analysis. The dynamical equation may be partially diagonalized into this subspace, where it becomes a modest system of linear ordinary differential equations coupled by a matrix that changes in time with the changing input.

A second approach, which currently is speculative but promising, is to recast the dynamical equation in terms of coefficients of a "moving basis" of the input-dependent eigenfunctions of the dynamical operator. This approach has the potentially useful additional feature that it reduces neuron population models to a single canonical form. This should lead, for example, to a quantitative estimate of how well an elaborate neuron model may be approximated by a simpler model.

We certainly can envision models of interacting cortical neuron subpopulations whose mathematical features will include some that lie beyond those presented here. The eventual accurately descriptive model may well fall in the broader set. Still, what is presented here may help furnish a bit of guidance for theoretical developments to come.

Appendix: Estimate of the Eigenvalues

A.1 The Model. The generalized integrate-and-fire model, of equation 1.1, can be made to mimic a more detailed model fairly well, for driving inputs $s(t)$ that do not change with unnatural abruptness. In the more detailed case of the Hodgkin-Huxley neuron, for example, the voltage equation is of

the form

$$dv/dt = F(v, m, h, n, s), \tag{A.1}$$

where m, h, n satisfy differential equations of their own. For a typical fixed value of s, these variables follow a limit cycle, and on that s-dependent limit cycle, the other variables can be expressed as functions of the concurrent value of v, whence

$$F(v, m, h, n, s) = F(v, m(v, s), h(v, s), n(v, s), s) = \widehat{F}(v, s). \tag{A.2}$$

The model equation,

$$dv/dt = \widehat{F}(v, s), \tag{A.3}$$

when s is held constant, will integrate to yield the exact voltage time course on the limit cycle, for each s-value. More generally, for time-dependent s, the momentary exact m, h, n values will not depart far from their limit-cycle values for present v and s, so long as s does not change by a major fraction of its value during a single cycle of v. The other variables relate to v in different ways on the fast descending side of the voltage cycle and on the much slower ascending side. To avoid double-valuedness, we may define our variable x in equation 1.1 as

$$x = \frac{1}{2\,(v_{\max} - v_{\min})} \times \begin{cases} v_{\max} - v & \text{(for downward motion)} \\ v_{\max} - v_{\min} + (v - v_{\min}) & \text{(for upward motion).} \end{cases} \tag{A.4}$$

If x is taken modulo unity, it is periodic on the limit cycle. Substitution of v in terms of x in equation A.3 now gives the $f(x, s)$ of equation 1.1.

If we proceed in this way, the boundary condition induced on the consequent dynamical equation, 1.4, for the population density $\rho(x, t)$, is that $\rho(x, t)$ should be smoothly periodic in x, with period 1. The more familiar boundary condition, for integrate-and-fire dynamics with diffusion, is an absorbing (or "no-return") upper boundary. We note the same effect is built in here, as on the final fast voltage rise of the nerve impulse, forward advection overwhelms backward diffusion.

The form of equation 1.4, as a good approximation to more elaborate and realistic dynamics, may be reached by an entirely different path. Suppose that in a state-space of higher dimension, we have properly formulated the dynamical operator in diffusion approximation. Suppose further that stochastic effects are modest enough so that the "fattened limit cycle" (discussed in section 2, equations 2.40 and 2.41), which supports the population

density function, is reasonably thin. Then our partial differential equation eigenvalue problem may be separated into a preliminary transverse eigenvalue problem that may be solved locally at each cross-section across the limit cycle, and an axial eigenvalue problem that follows from assembling the solutions to the transverse problem. (This is the same methodology that is natural in addressing wave propagation in a waveguide of variable cross-section.) The resulting axial dynamical operator inherits both advective and diffusive terms, and is of the form given in equation 1.4, although in this case $f(x, s)$ includes the effects of transverse diffusion.

A.2 Small Stochastic Effects. If the diffusion term in equation 1.4 is small, we may approach the eigenvalue problem,

$$Q\phi_n = \lambda_n \phi_n, \tag{A.5}$$

by straightforward perturbation theory. In equation 1.4, let

$$Q = Q_0 + Q_1, \tag{A.6}$$

where

$$Q_0 = -\frac{\partial}{\partial x} f(x), \; Q_1 = \frac{\partial}{\partial x} D(x) \frac{\partial}{\partial x} \tag{A.7}$$

(we drop the s dependence from the notation, as input is held fixed in the eigenvalue problem), and we assume that inclusion of the small Q_1 induces changes away from both the eigenvalues and eigenvectors of Q_0 :

$$Q_0 \phi_n^{(0)} = \lambda_n^{(0)} \phi_n^{(0)} \tag{A.8}$$

$$\phi_n = \phi_n^{(0)} + \phi_n^{(1)} \tag{A.9}$$

$$\lambda_n = \lambda_n^{(0)} + \lambda_n^{(1)}. \tag{A.10}$$

Substitution of equations A.6, A.9, and A.10 into A.5 (and use of equation A.8) gives the first-order part of equation A.5:

$$Q_0 \phi_n^{(1)} + Q_1 \phi_n^{(0)} = \lambda_n^{(0)} \phi_n^{(1)} + \lambda_n^{(1)} \phi_n^{(0)}, \tag{A.11}$$

where we regard terms with index zero as known, and those with index one as to be found. The perturbed eigenfunction can be expanded in terms of the unperturbed eigenfunctions:

$$\phi_n^{(1)} = \sum_{m \neq n} a_m \phi_m^{(0)}. \tag{A.12}$$

It is notable that the $m = n$ term, which would include an admixture of $\phi_n^{(0)}$ in $\phi_n^{(1)}$, is absent from the sum. This term is absent because its inclusion in equation A.9 would serve simply to rescale the unperturbed eigenvector, and the eigenvector property is not affected by a scale change. Essentially equation A.11 asks, How much must the old eigenvector be *re*directed to be an eigenvector of the modified operator? and a simple scale change is irrelevant to answering this question. So every term of equation A.12 is biorthonormal to the adjoint eigenvector $\widehat{\phi}_n^{(0)}$. If all we want from equation A.11 is the first-order correction to the eigenvalue, $\lambda_n^{(1)}$, we may inner-product that equation with $\widehat{\phi}_n^{(0)}$, and by use of equations 2.22 and 2.27, we at once find

$$\lambda_n^{(1)} = \left(\widehat{\phi}_n^{(0)}, Q_1 \phi_n^{(0)}\right). \tag{A.13}$$

The inner product for the state-space of equation 1.4 is

$$(v, u) = \int_0^1 dx v^* (x)\, u (x). \tag{A.14}$$

We observe that

$$\begin{aligned}(v, Q_0 u) &= -\int_0^1 dx v^* (x) \frac{\partial}{\partial x}\left(f (x)\, u (x)\right) \\ &= \int_0^1 dx \left(f (x) \frac{\partial}{\partial x} v (x)\right)^* u (x) \\ &= \left(Q_0^* v, u\right), \end{aligned} \tag{A.15}$$

where for the second line we integrated by parts, and used the periodic boundary condition discussed in section A.1. By equation A.15 the zeroth-order adjoint operator is

$$Q_0^* = f (x) \frac{\partial}{\partial x}, \tag{A.16}$$

and the adjoint eigenvalue equation,

$$Q_0^* \widehat{\phi}_n^{(0)} = \left(\lambda_n^{(0)}\right)^* \widehat{\phi}_n^{(0)}, \tag{A.17}$$

is

$$f (x) \frac{d}{dx} \widehat{\phi}_n^{(0)} = \left(\lambda_n^{(0)}\right)^* \widehat{\phi}_n^{(0)}, \tag{A.18}$$

which (by direct substitution back into equation A.18) has the solution

$$\widehat{\phi}_n^{(0)}(x) = \exp\left\{\left(\lambda_n^{(0)}\right)^* \int_0^x dx' \left(1/f\left(x'\right)\right)\right\}. \tag{A.19}$$

Periodicity demands that

$$\widehat{\phi}_n^{(0)}(1) = \widehat{\phi}_n^{(0)}(0), \tag{A.20}$$

and this requires us to choose the eigenvalues

$$\lambda_n^{(0)} = i \cdot 2\pi n \cdot \left\{\int_0^1 dx \left(1/f(x)\right)\right\}^{-1} = in\omega_0, \tag{A.21}$$

which is of a form in agreement with equation 2.41. If we now go back to equation 1.1, in the absence of noise it gives

$$dt = dx/f(x), \tag{A.22}$$

which we may immediately integrate from $x = 0$ to $x = 1$ to obtain the transit time, or reciprocal firing frequency:

$$T = \int_0^1 dx/f(x). \tag{A.23}$$

As firing period and radian frequency are related by

$$\omega_0 = 2\pi/T, \tag{A.24}$$

this is in exact agreement with the fundamental radian frequency that appears in the eigenvalue evaluation, A.21.

In the same manner as what we have just done, from equation A.7, the unperturbed dynamical operator's eigenvalue equation is

$$-\frac{d}{dx}\left(f(x)\,\phi_n^{(0)}\right) = \lambda_n^{(0)}\phi_n^{(0)}, \tag{A.25}$$

which, as back-substitution again confirms, has the solution

$$\phi_n^{(0)}(x) = C\left(1/f(x)\right)\exp\left\{-\lambda_n \int_0^x dx' \left(1/f\left(x'\right)\right)\right\}. \tag{A.26}$$

If in equation A.26 we choose the constant

$$C = T^{-1}, \tag{A.27}$$

then the eigenfunctions and their adjoints are indeed biorthonormal, as equation 2.27 becomes

$$\left(\widehat{\phi}_m^{(0)}, \phi_n^{(0)}\right) = C \int_0^1 \frac{dx}{f(x)} \exp\left\{\frac{2\pi i}{T}(n-m)\int_0^x \frac{dx'}{f(x')}\right\}. \tag{A.28}$$

In fact equation A.22 serves as a crafty substitution that reduces equation A.28 to an easy integral and confirms equation 2.27 for this case.

We may now evaluate the integral, equation A.13, for the eigenvalue's perturbation due to diffusion. We first make the observation that

$$\begin{aligned}(v, Q_1 u) &= \int_0^1 dx v^*(x)\frac{d}{dx}D(x)\frac{d}{dx}u(x) \\ &= -\int_0^1 dx \left(\frac{d}{dx}v(x)\right)^* D(x)\frac{d}{dx}u(x),\end{aligned} \tag{A.29}$$

again by parts integration and periodicity. Thus equation A.13 becomes

$$\lambda_n^{(1)} = -\int_0^1 dx \left(\frac{d}{dx}\widehat{\phi}_n^{(0)}(x)\right)^* D(x)\frac{d}{dx}\phi_n^{(0)}(x). \tag{A.30}$$

The derivative that involves $\widehat{\phi}_n^{(0)}$ we can simplify directly from the adjoint eigenvalue equation, A.18, to obtain a term with a factor $\lambda_n^{(0)}$. To evaluate the right-most derivative in equation A.30, we may bring equation A.25 to the form

$$\frac{d\phi_n^{(0)}}{dx} = -\frac{\left(\lambda_n^{(0)} + f'(x)\right)}{f(x)}\phi_n^{(0)}. \tag{A.31}$$

From equations A.19 and A.26, we have further that

$$\widehat{\phi}_n^{(0)*}(x)\,\phi_n^{(0)}(x) = \frac{1}{T}\frac{1}{f(x)}. \tag{A.32}$$

When these ingredients are substituted into equation A.30, we obtain the evaluated form,

$$\lambda_n^{(1)} = \left(\lambda_n^{(0)}\right)^2 \int_0^1 \frac{dx}{T} \frac{D(x)}{\left(f(x)\right)^3} + \lambda_n^{(0)} \int_0^1 \frac{dx}{T} \frac{f'(x) D(x)}{\left(f(x)\right)^3}$$

$$= \left(\lambda_n^{(0)}\right)^2 A + \lambda_n^{(0)} B, \tag{A.33}$$

where A is clearly positive and the two integrals A and B are independent of the eigenvalue number n. (However, they do depend on the input s, which controls the fundamental radian frequency ω_0.) The eigenvalue without diffusion, $\lambda_n^{(0)}$, is given by equation A.21; substitute that into equations A.33 and A.10, which gives the eigenvalue through zeroth and first orders, and we have

$$\lambda_n = -n^2\omega_0^2 A + in\omega_0 (1 + B). \tag{A.34}$$

Here (unlike in equation 2.42) ω_0 is the fundamental radian firing frequency in the absence of diffusion, and A and B, assumed small, are the two integrals that appear in equation A.33, which are linear in the strength of diffusion. The presence of diffusion modifies the fundamental firing frequency, through the term B in equation A.34; and equation 2.42 is confirmed under the often reasonable assumption of small diffusion and in the context of the fairly accurate model presented here.

A.3 Rapidly Oscillating Eigenfunctions. The eigenfunction problem given by equation 1.4

$$\frac{d}{dx}\left(-f(x) + D(x)\frac{d}{dx}\right)\phi_n = \lambda_n \phi_n, \tag{A.35}$$

may be expected to yield solutions $\phi_n(x)$ that oscillate more and more rapidly across the range of x, as $|n|$ and consequently $|\lambda_n|$ become larger. This offers us another approximation procedure, the "WKB" approximation, that does not hinge on small diffusion. We assume that the solution locally looks like a sinusoidal wave, but with a wavelength and amplitude that change gradually with position:

$$\phi_n(x) = A(x) \exp\left\{i \int_0^x dx' k\left(x'\right)\right\}. \tag{A.36}$$

In the very special case where f and D are constant, equation A.36 solves A.35 *exactly* with A and k constant, and where k must satisfy

$$-ifk - Dk^2 = \lambda_n. \tag{A.37}$$

If we now insist on solutions with unit period, we must have $k = 2\pi n$, whereupon equation A.37 agrees exactly with equation 2.42. For variable f and D, substitution of equations A.36 into A.35 yields the result that, so long as $k(x)$ is very large compared to the first derivatives of $f(x)$ and $D(x)$, neglect of the relatively much smaller terms again leads to equation A.37 locally for every value of x, or equivalently

$$k = \frac{-i}{2D}\left\{f \pm \sqrt{f^2 + 4D\lambda_n}\right\}. \tag{A.38}$$

If we let the diffusion D go to zero in equation A.38 (with the choice of the minus sign), then substitution of A.38 into A.36 again gives an exact solution, as we retrieve the zeroth-order exact solution of the previous subsubsection, which in fact was in the form of equation A.36. Substitution of $k(x)$ from equation A.38 into A.36 yields a condition that specifies the eigenvalue once we insist that $\phi_n(x)$ be periodic, as in equation A.20:

$$\int_0^1 dx \frac{-1}{2D(x)}\left\{f(x) \pm \sqrt{\left(f(x)\right)^2 + 4D(x)\lambda_n}\right\} = -2\pi i n. \tag{A.39}$$

This is a transcendental equation to be solved for λ_n. If n is large, then clearly $|\lambda_n|$ must be chosen large to solve the equation. But if $|\lambda_n|$ is large, then we can approximate k in equation A.38 by

$$k \cong \frac{-i}{2D}\left\{f + 2\lambda_n^{1/2} D^{1/2}\left(1 + \frac{1}{8}\frac{f^2}{D\lambda_n}\right)\right\} \tag{A.40}$$

(where $\lambda^{1/2}$ implies both positive and negative signs) and the eigenvalue equation, A.39, becomes

$$\begin{aligned} -in &= \int_0^1 \frac{dx}{2\pi}\frac{f}{2D} + \lambda_n^{1/2}\int_0^1 \frac{dx}{2\pi}\frac{1}{D^{1/2}} + \frac{1}{\lambda_n^{1/2}}\int_0^1 \frac{dx}{2\pi}\frac{1}{8}\frac{f^2}{D^{3/2}} \\ &= \alpha + \lambda_n^{1/2}\left(\beta + \frac{\gamma}{\lambda_n}\right) \end{aligned} \tag{A.41}$$

where α, β, γ are the three specific integrals, which do not depend on λ_n or n. Carry α to the left side by subtracting it from equation A.41, and square:

$$\begin{aligned} -n^2 + 2in\alpha + \alpha^2 &= \lambda_n\left(\beta^2 + \frac{2\beta\gamma}{\lambda_n} + \frac{\gamma^2}{\lambda_n^2}\right) \\ &= \beta^2\lambda_n + 2\beta\gamma + \frac{\gamma^2}{\lambda_n}. \end{aligned} \tag{A.42}$$

If $|\lambda_n|$ is large, the right-hand term with the coefficient $1/\lambda_n$ may be dropped, and the rest may be solved for λ_n :

$$\lambda_n = -\frac{1}{\beta^2}n^2 + i \cdot 2\frac{\alpha}{\beta^2}n + \left(\frac{\alpha^2}{\beta^2} - \frac{2\gamma}{\beta}\right). \tag{A.43}$$

This result, which should hold for large n, is again in agreement with equation 2.42, except for the right-hand asymptotic off-set term, which does not grow with n.

Acknowledgments

I owe a particular debt to Lawrence Sirovich. Others who contributed valuable ideas are Daniel Amit, Dmitrii Manin, and Ahmet Omurtag. This work has been supported by NIH/NEI EY11276, NIH/NIMH MH50166, ONR N00014-96-1-0492.

Publications from the Laboratory of Applied Mathematics may be found at http://camelot.mssm.edu/.

References

Abbott, L. F., & van Vreeswijk, C. (1993). Asynchronous states in networks of pulse-coupled oscillators. *Phys. Rev. E., 48*, 1483–1490.

Bailey, N. T. J. (1964). *The elements of stochastic processes with applications to the natural sciences*. New York: Wiley.

Everson, R. M., Prashanth, A. K., Gabbay, M., Knight, B. W., Sirovich, L., & Kaplan, E. (1998). Representation of spatial frequency and orientation in the visual cortex. *Proceedings of the National Academy of Science*, *95*, 8334–8338.

Gerstner, W. (1995). Time structure of the activity in neural network models. *Phys. Rev. E., 51*, 738–758.

Gerstner, W. Population dynamics of spiking neurons: Fast transients, asynchronous states and locking. In press. *Neural Computation*.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol., 117*, 500–544.

Johannesma, P. I. M. (1969). *Stochastic neural activity: A theoretical investigation*. Unpublished Ph.D. dissertation, University of Nijmegen.

Knight, B. W. (1972a). Dynamics of encoding in a population of neurons. *J. Gen. Physiol., 59*, 734–766.

Knight, B. W. (1972b). The relationship between the firing rate of a single neuron and the level of activity in a population of neurons: Experimental evidence for resonant enhancement in the population response. *J. Gen. Physiol., 59*, 767–778.

Knight, B. W., Manin, D., & Sirovich, L. (1996). Dynamical models of interacting neuron populations. In *Symposium on Robotics and Cybernetics: Computational Engineering in Systems Applications (Gerf, E.C. ed) Cite Scientifique, Lille, France*.

Knight, B. W., Omurtag, A., & Sirovich, L. (2000). The approach of a neuron population firing rate to a new equilibrium: An exact theoretical result. In press. *Neural Computation*.

Kuramoto, Y. (1991). Collective synchronization of pulse-coupled oscillators and excitable units. *Physica D, 50*, 15–30.

McCormick, D. A., & Huguenard, J. R. (1992). A model of the electrophysiological properties of thalamocortical relay neurons. *J. Neurophysiol., 68*, 1384–1400.

Nykamp, D., & Tranchina, D. A population density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning. In press. *J. Comp. Neurosci*.

Omurtag, A., Knight, B. W., & Sirovich, L. (1998). On the simulation of large populations of neurons. In press. *J. Comp. Neurosci*.

Risken, H. (1996). *The Fokker–Planck equation. Methods of solution and applications*. Berlin: Springer-Verlag.

Sirovich, L., & Everson, R. (1992). Management and analysis of large scientific datasets. *International Journal of Supercomputer Applications, 6*, 50–68.

Sirovich, L., Everson, R., Kaplan, E., Knight, B. W., O'Brien, E., & Orbach, D. (1996). Modeling the functional organization of the visual cortex. *Physica D, 96*, 355–366.

Sirovich, L., Knight, B. W., & Omurtag, A. (2000). Dynamics of neuronal populations: The equilibrium solution. In press. *SIAM*.

Sirovich, L., Knight, B. W., & Omurtag, A. (1999a). Dynamics of neuronal populations: Eigentheory. In preparation.

Sirovich, L., Knight, B. W., & Omurtag, A. (1999b). Dynamics of neuronal populations: Stability theory. In preparation.

Sirovich, L., Knight, B. W., & Omurtag, A. (1999c). Dynamics of neuronal populations: Low dimensional approximations. In preparation.

Stein, R. B. (1965). A theoretical analysis of neuronal variability. *Biophys. J., 5*, 173–194.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology* (Vol. 2). Cambridge: Cambridge University Press.

Wilbur, W. J., & Rinzel, J. (1982). An analysis of Stein's model for stochastic neuronal excitation. *Biol. Cybern., 45*, 107–114.

Received January 11, 1999; accepted April 12, 1999.