

Simplifying and reducing complex models

Bard Ermentrout *

Department of Mathematics

University of Pittsburgh

Pittsburgh, PA 15260

March 20, 2002

Abstract

Several strategies for reducing the complexity of biologically based models are presented. These methods are primarily based on averaging either over instances or time. In either case, the resulting equations can be directly connected to the original model but often times lead to a much simpler system of equations.

1 Introduction

The skill and innovation of experimental biologists has enabled them to get more and more information about their preparations. This presents a challenge to anyone who wishes to create a mathematical model or simulation of the given system. At what point does the model cease to have explanatory value having become too complex to do anything more than simulate it at a

*Supported in part by NIMH and NSF

variety of parameter values and initial conditions? Often, the models that are proposed have dozens of parameters many of which may not be known for the particular system studied. Furthermore, the complexity of the models makes it difficult to study sensitivity to parameters and initial conditions even on fast computers. This difficulty is magnified when the systems that are simulated are inherently stochastic, for then, one can ask how many sample paths is enough? In addition to the computational difficulties and the incomplete knowledge of parameters, there is also the issue of the interpretation of the output of the model. Large simulations produce a tremendous amount of output and much of it is likely to be useless for the particulars of a given experiment. Finally, for many biological systems, one can only guess at the mechanism. A simulation does not tell you how dependent the behavior is on the particular instance of the mechanism that you have chosen. Only a detailed analysis can tell you that and for complex models and simulations this is difficult at the very least and usually impossible.

These concerns lead many modelers to propose simplified models. The multiple time and space scales in biological systems force one to generally focus on some particular level using often heuristic approximations for the finer details which are neglected. The hope is that the knowledge of the finer levels suggests the correct heuristics for modeling and understanding the higher levels of the system. This principle of reductionism has served the sciences well. In physics, where there are well-defined laws, it is often possible to use a microscopic description to derive a macroscopic model. One does not usually treat every water molecule separately when studying flow through a tube; one instead uses the Navier-Stokes equations which describe the macroscopic behavior of fluids. There are several reasons for this. One is the obvious computational complexity. The second and equally important reason is that the behavior of the individual molecules cannot be studied nor

is it of interest. Modeling a biological system should be viewed in the same manner. If one wants to study the behavior of large scale electrical activity in a region of cortex, then is it necessary to include dozens of channels into each nerve cell? Since the properties of individual neurons are not known, the choice of parameters for each of these channels is at best an average of similar systems. Thus, one approach has been to use simplified models. The difficulty with simplifying is how to choose the simple model, how to connect it to the measured phenomena, and what is its relationship to the details that lurk beneath? Is it possible to construct simplified models which are *quantitative* rather than just qualitative.

In this chapter, we will describe some methods which allow one to derive quantitatively correct models from more complex systems. We will attempt to show that the assumptions of certain dynamical properties and extreme differences in time and space scales can be exploited to produce simple often analyzable models.

1.1 Averaging.

In many models, there are diverse time scales and space scales. Extreme time differences often allow one to assume that during the changes in a fast quantity, a slow quantity can be considered constant. On the other hand, if the slow quantity is of interest, then the fluctuations or changes in the fast quantity occur so fast, the the slow quantity only “sees” the mean or average of them. This intuitively appealing idea can be made rigorous by a procedure called averaging. Related to this is the idea of “spatial averaging” in which one assumes individual influences of one system on another are small but manifold. Thus, one averages over these; such an average or approximation is often called the “mean field approximation.”

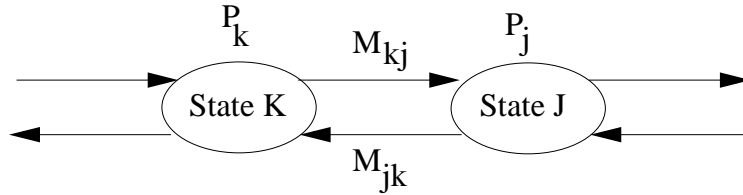


Figure 1: Diagram showing the probability of jumping from state J to K in a Markov chain.

Both averaging and mean field approaches provide a method for making complex models much simpler without losing quantitative details. I will describe a number of these ideas through a variety of examples.

2 Master Equations.

Many biological problems can be cast as continuous time jump processes in which a system switches from one state to the next. The simplest example would be the random opening and closing of a channel. Another example is the growth of an actin polymer in the presence of some cell signal where the states are the length of the polymer. Any system involving rates such as chemical reactions can be viewed as a continuous time process with jumps. If the probability of jumping from one state to the next depends only on the current state (that is, it has no history dependence) then the process is called a Markov process.

In the simplest case, a quantity of interest takes on n values and the rate of jumping from k to j is given by $M_{kj}dt$ where dt is the time interval (see figure 1). This is easily simulated and from the simulation, it is possible to obtain the probabilities of being in any given state. The key to analyzing this is to assume that one does the simulation over a long period of time or that we are interested in the average properties of the process. Let $P_j(t)$ be

the probability of being in state j at time t . From the figure, the change in probability P_j is the net influx from P_k minus the outflux from P_j i.e.

$$\Delta P_j = M_{kj}P_k\Delta t - M_{jk}P_j\Delta t.$$

Taking the limit leads to the differential equation:

$$\frac{dP_j}{dt} = \sum_{k=1}^n (M_{kj}P_k - M_{jk}P_j) \equiv \sum_{k=1}^n A_{kj}P_k \quad (2.1)$$

This equation is called the Master Equation. (It is also called the Chapman-Kolmogorov equation.) The initial probabilities must sum to 1 and then it is a simple matter of solving this equation. The steady state probabilities are the desired result. The matrix A has a zero eigenvalue since rows sum to zero. The eigenvector (normalized of course) corresponding to this eigenvalue gives the steady state probabilities. Thus, by looking at this mean or averaged system, the steady state behavior is found.

A more interesting situation arises when one of the variables involved itself obeys a differential equation. Consider, for example, the simple model:

$$\frac{dx}{dt} = z - x \quad (2.2)$$

where z randomly switches between two states, 0, 1 at some constant rate, r . Since x cannot instantly follow the variable, z , we expect the probability, $P(x = X, t)$ to be a continuous function of X . In figure 2, I show histograms for the distribution of x for this system when z randomly switches back and forth between the two states 0 and 1 at a rate r . For fast changes, we expect x to hover around the mean value of z which is 1/2. This is seen in Fig 2a. However, for slow rates, $r = 0.2$ for example, x has enough time to move to $z = 0$ or $z = 1$ so that the distribution is strongly bimodal as seen in the figure. There is a transition from a unimodal to a bimodal distribution as

the rate is decreased. Once again, we can appeal to the Master Equation to understand this transition.

Suppose that z has n states and that

$$\frac{dx}{dt} = f(x, z).$$

Furthermore, suppose that the rates of transition from state j to state k may be x -dependent, $M_{jk}(x)$. Then, we need to define the probabilities of being in state j with $x = X$, $P_j(X, t)$. Averaging once again over many sample paths, one finds that

$$\begin{aligned} \frac{\partial P_j(X, t)}{\partial t} &= -\frac{\partial}{\partial X} [f(X, j)P_j(X, t)] \\ &+ \sum_{k=1}^n (M_{kj}(X)P_k(X, t) - M_{jk}(X)P_j(X, t)). \end{aligned}$$

Applying this to our system we obtain:

$$\begin{aligned} \frac{\partial P_0}{\partial t} &= -\frac{\partial}{\partial X} (-XP_0) - rP_0 + rP_1 \\ \frac{\partial P_1}{\partial t} &= -\frac{\partial}{\partial X} ((1-X)P_1) - rP_1 + rP_0 \end{aligned}$$

Let $P(x, t) = P_0(x, t) + P_1(x, t)$ be the probability of $X = x$ at time t and in either state 0 or state 1. The steady state distribution for this turns out to be:

$$P(x = X) = CX^{r-1}(1-X)^{r-1}$$

where C is a normalization constant. From this it is clear that if $r < 1$, that is, the rates are slow, then the distribution is singular with asymptotes at 0 and 1. On the other hand, for fast rates, $r > 1$ the distribution is continuous, vanishes at the ends, and is peaked at $X = 1/2$. As the transitions become infinitely fast, the distribution becomes infinitely narrow and centered at $X = 1/2$, the limit obtained by solving (2.2) using the average $z = 1/2$. Averaging over sample paths has provided a quantitative method for completely characterizing this simple two-state model coupled to a differential equation.

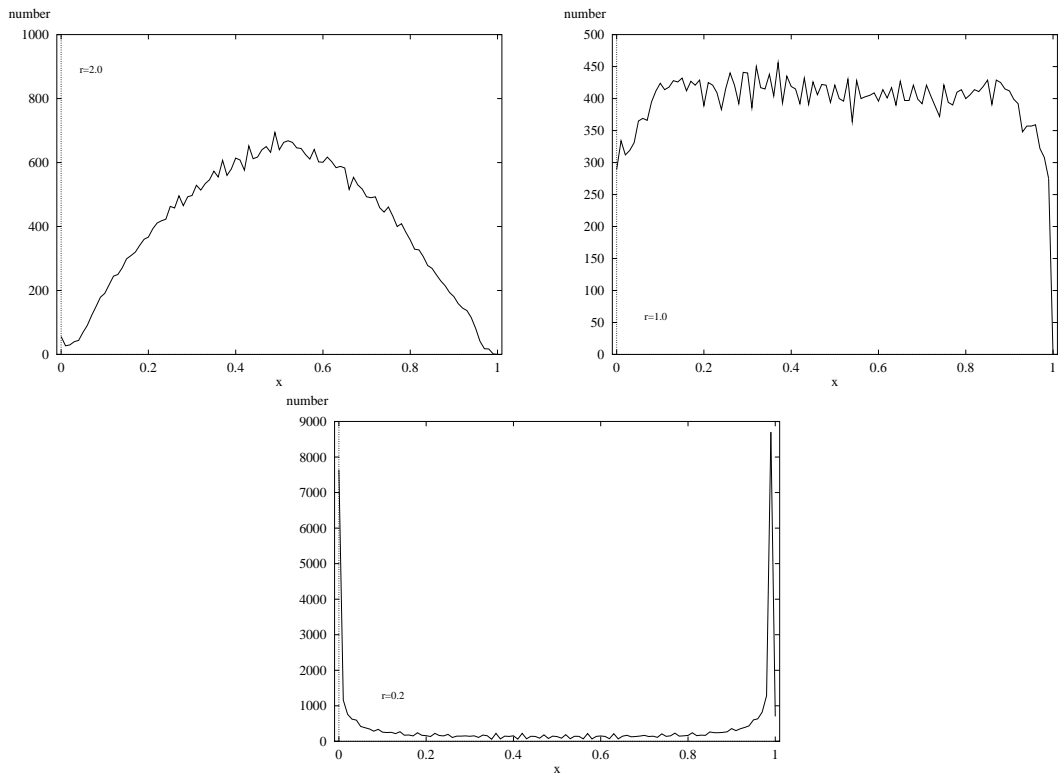


Figure 2: Histograms for the state variable, x in the two-state Markov model as the rate varies from fast $r = 2$ to intermediate, $r = 1$ to slow $r = 0.2$

2.1 Application to a model for fibroblast orientation.

I will next consider an example from cell biology. Here we consider the effect of density on the behavior of fibroblasts in culture. (Edelstein-Keshet and Ermentrout, 1991). It is known that fibroblasts will align with each other with a probability that depends on their relative angles of motion. In dense cultures, many patches of parallel cells can be found. In sparse cultures, there are few aligned patches. Thus, we would like to understand how the density affects the appearance of parallel patches. We can treat the formation of arrays as a stochastic process in which cells will switch from one angle to another angle when they interact with each other. Spatial distribution is obviously quite important but here we will neglect it and simply study the alignment problem. We distinguish cells that are moving (free) from cells that are attached (bound) to the culture dish surface. The following types of interactions are allowed:

1. Random shifts in alignment from an angle θ to a neighboring angle $\theta \pm \phi$ at rate r for free cells
2. Cells that bump into other bound cells of orientation θ' will reorient to that angle and stick at a rate that is linearly related to the fraction of cells with angle θ' and the difference between the angles, $K(\theta - \theta')$.
3. Bound cells will free up at a rate γ and free cells become bound at a rate η .

Since the area of the culture dish does not change over time and only the number of cells does, we will treat this number N as the parameter. For simplicity assume m discrete angles. Let $F_j(t)$ denote the probability of a cell being free and having orientation θ_j and let $B_j(t)$ be the probability of a

cell being bound and having an orientation θ_j . The Master Equation yields:

$$\begin{aligned} \frac{dF_j}{dt} &= r(F_{j-1} - 2F_j + F_{j+1}) - N \sum_i K(i-j) B_i F_j \\ &+ \gamma B_j - \eta F_j \\ \frac{dB_j}{dt} &= \eta F_j - \gamma B_j + N \sum_i K(i-j) B_j F_i. \end{aligned}$$

The number of cells, N , appears in the equations because the larger the number of cells the greater the rate of collision between bound and free cells. One solution to this equation is $F_j = f$ and $B_j = b$ where both f and b are constant. The stability of the uniform state can be analyzed as a function of N and the critical value of N can be found for a spatial instability. This can then be compared to the stochastic simulations and the biological system. Related reductions of discrete probabilistic simulations are given in Ermentrout and Edelstein-Keshet, (1991).

2.2 Mean field reduction of a neural system

. The Master Equation essentially averages over many sample paths in a system and leads to a set of equations for the probability of any given state of the system. Another way to average a system that has intrinsic randomness is the so-called “mean-field” approximation. Here, one uses that idea that there are many interacting subunits that are tightly coupled and so the effect is one of the average of all of them. (See e.g., Cowan, 1968, or van Vreeswijk and Sompolinsky, 1998). In this section, we apply the ideas of averaging over units to reduce a random system of many units to a deterministic system with just two equations. The original model attempts to understand the effect that cortical processing has on thalamic input in the somatosensory whisker barrel area of the rat (Pinto, et al 1996). The model barrel contains N_e excitatory cells and N_i inhibitory cells. Each cell is coupled to the other cells with a

randomized weight (positive for excitatory and negative for inhibitory) and each cell receives excitatory input from N_T thalamic neurons, again with randomized weights. The model is cast as an integral equation. If V is the voltage then the probability of firing an action potential is $P(V)$. A typical example is

$$P(V) = 1/(1 + e^{-\beta(V-V_{thr})}).$$

The parameter β determines the sharpness of the probability and V_{thr} is the voltage at which there is a 50% chance of firing a spike.

The potential of the k^{th} excitatory neuron is :

$$\begin{aligned} V_k^e(t) &= \sum_j w_{ee}^{kj} \int_0^t P_e(V_j^e(s)) e^{-(t-s)/\tau_e} ds \\ &+ \sum_j w_{te}^{kj} \int_0^t P_e(V_j^T(s)) e^{-(t-s)/\tau_e} ds \\ &- \sum_j w_{ie}^{kj} \int_0^t P_i(V_j^i(s)) e^{-(t-s)/\tau_i} ds \end{aligned}$$

where the probabilities of thalamic cell spikes are given as inputs. A similar equation holds for the inhibitory neurons. Since the cells fire asynchronously and the weights are randomly distributed about some mean, an obvious approximation is to sum up all the excitatory (inhibitory) cells and then divide by their number to arrive at a mean potential. The question is what equations does this mean satisfy. A preliminary transformation to a differential equation eases the analysis. Let

$$S_j^e(t) = \int_0^t P_e(V_j^e(s)) e^{-(t-s)/\tau_e} ds.$$

This is the weighted average of the firing probabilities taking into account the past history. Then

$$\begin{aligned} \tau_e \frac{dS_k^e}{dt} + S_k^e &= P_e(V_k^e) \\ V_k^e &= \sum_j (w_{ee}^{kj} S_j^e + w_{te}^{kj} S_j^T - w_{ie}^{kj} S_j^i) \end{aligned}$$

We define the mean field

$$S_e = \frac{1}{N_e} \sum_k S_k^e,$$

and find that we have to deal with sums of the form

$$\frac{1}{N_e} \sum_k P_e \left(\sum_j w_{ee}^{kj} S_j^e + \dots \right).$$

The only approximation that is made is to interchange the nonlinearity and the sum. That is, we approximate

$$\frac{1}{N_e} \sum_k P_e(V_k^e) \approx P_e \left(\frac{1}{N_e} \sum_k V_k^e \right).$$

Since P is a sigmoid curve and the most sensitive behavior occurs near the linear portion, this is not a bad approximation. In fact, one can even compensate for this approximation by modifying the averaged sigmoid. The final result is a set of equations for the mean of the activity in a whisker barrel:

$$\tau_e S_e' = P_e(w_{ee} S_e^e + w_{te} S^T - w_{ie} S^i)$$

with analogous equations for the inhibition. The thalamic drive, S^T is given as input to the model and

$$w_{ee} = \frac{1}{N_e} \sum_{jk} w_{ee}^{jk}.$$

The other weights have similar definitions.

The reduced mean field model turns out to mimic the behavior of the full system extremely well and in fact was used to make experimental predictions to some novel forms of stimulus in later work (Pinto et al, 1996).

3 Deterministic systems

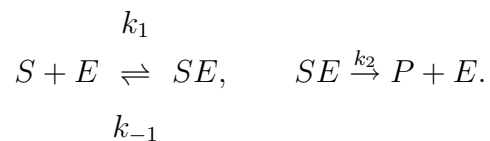
There are many ways to reduce the dimension and complexity of deterministic systems. Again, all of these exploit the differences in time scales or space

scales. Here we will concentrate on techniques for reduction exploiting time scales. The idea is very simple – if some process occurs over a much faster time scale than the one of interest, look at the mean of the fast process and eliminate it from the system or conversely, hold the slow processes as constant parameters and study the fast process. This approach has many advantages:

- A smaller systems arises
- There is a direct quantitative connection to the detailed system
- Computations are hard when there are drastically different time scales; eliminating the fast or slow variables makes the computations easier.
- Separation of time scales enables one to understand the mechanisms underlying the behavior.

3.1 Reduction of dimension by using the pseudo-steady state.

Most biochemists are aware of this method as it is the idea that is used to derive the Michaelis-Menten equations (Murray, 1989). Recall that one wants to model the following reaction:



Letting $s = [S], e = [E], c = [SE], p = [P]$, we get

$$\begin{aligned} s' &= -k_1es + k_{-1}c, & e' &= -k_1es + (k_{-1} + k_2)c \\ c' &= k_1es - (k_{-1} + k_2)c, & p' &= k_2c. \end{aligned}$$

Clearly, p is obtained just by integrating c . Summing the equations for c and e implies that $c' + e' = 0$ so that $c + e = e_0$. We have immediately eliminated two of the four equations and must only look at a two-dimensional system. By introducing dimensionless variables and parameters, we can eliminate one more equation. As an aside, I want to point out the incredible usefulness of rendering a model dimensionless; this allows one to compare parameters which prior to scaling had different units and so could not be compared. This type of comparison enables one to see where small and large parameters lie and thus direct the reduction of dimension. Murray (and many others) introduce the following scaled parameters and variables:

$$\begin{aligned} \tau &= k_1 e_0 t, & u(\tau) &= s(t)/s_0, & v(\tau) &= c(t)/e_0 \\ \lambda &= \frac{k_2}{k_1 s_0}, & K &= \frac{k_{-1} + k_2}{k_1 s_0}, & \epsilon &= \frac{e_0}{s_0}, \end{aligned}$$

where s_0 is the initial substrate. With this scaling, the equations are:

$$\frac{du}{d\tau} = -u + (u + K - \lambda)v, \quad \epsilon \frac{dv}{d\tau} = u - (u + K)v.$$

Since the amount of enzyme compared to the initial substrate is generally very small, the parameter ϵ is quite small. The approximation is to set this to zero. This implies $v = u/(u + K)$ so that we finally end up with

$$\frac{du}{d\tau} = -u + (u + K - \lambda) \frac{u}{u + K} = -\lambda \frac{u}{u + K}.$$

There is only one differential equation to solve!

In general, if the equations have the form:

$$\epsilon \frac{dX}{dt} = F(X, Y), \quad \frac{dY}{dt} = G(X, Y), \quad X \in R^m, \quad Y \in R^n,$$

then, one sets $\epsilon = 0$, solves for $X(Y)$ and plugs this back into the equation for Y . We call X the fast variable(s) and Y the slow variable(s). This approach has been used by Rinzel (1985) to reduce the 4-dimensional Hodgkin-Huxley

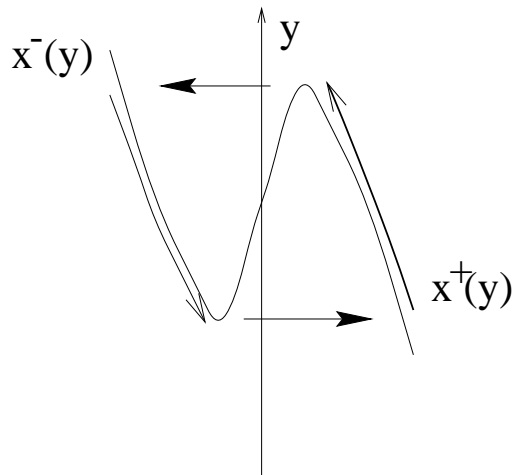


Figure 3: Illustration of a relaxation oscillator. As the slow variable creeps up the right branch, it encounters the knee and must jump to the left branch. Then the slow variable moves down until the lower knee is reached. A jump to the right branch occurs.

equations to a two-dimensional model. The advantages of a two-dimensional model are that the powerful methods of phase-plane analysis can be used to understand the equations. Abbott and Kepler (1990) generalize Rinzel's ideas and describe a way to simplify arbitrary neural models. In fact, they reduce the six-dimensional Connor-Stevens model to just two dimensions using the differences in time scales.

In many interesting cases, the equation $F(X, Y) = 0$ may have several branches of solutions for a given value of Y . This can lead to oscillations and other phenomena. For example, the classic van der Pol oscillator can be written as:

$$\epsilon \frac{dx}{dt} = x(1 - x^2) - y, \quad \frac{dy}{dt} = x.$$

It is easiest to visualize what is going on by looking in the (x, y) plane. In figure 3, I plot the curve, $y = x(1 - x^2)$ and $x = 0$ which are respectively

where $dx/dt = 0$ and $dy/dt = 0$. There is one equilibrium point at the origin and it is unstable. Setting $\epsilon = 0$, we must solve $x - x^3 = y$. For $|y| < \frac{2}{3\sqrt{3}}$, this equation has three roots. Suppose, we take the most positive root, $x^+(y) > 0$. (See figure 3). Since $x^+(y) > 0$ this means that y will increase. As long as y is below the maximum of the cubic, we can continue to define $x^+(y)$. However, once y exceeds this maximum, the only possible root is $x^-(y)$ and the system “jumps” to the left branch. However, $x^-(y)$ is negative, so y begins to decrease until it reaches the minimum of the cubic and x must jump to the right branch. Thus, an oscillation is formed. This type of model is called a “relaxation oscillator” and is the basis for many models of excitable and oscillatory media. (See Murray, 1989, for many examples applied to biology.) The approximate period of the oscillation is given by the amount of time it takes to traverse the two outer branches of the cubic. If we rescale time, introducing $T = t/\epsilon$ then the van der Pol model becomes:

$$\frac{dx}{dT} = x(1 - x^2) - y \quad \frac{dy}{dT} = \epsilon x.$$

Setting $\epsilon = 0$, we see that y is essentially constant, so that x will tend to one of up to three fixed points of the equation $x(1 - x^2) = y$. As we saw above, for values of y near zero, there are up to three roots. The middle fixed point is unstable so only the outer ones are possible steady states. This multiplicity of stable states coupled with the slow movement of another variable (in this case y) has led to an oscillation which jumps between two fixed points. What if the “ x ” system is two-dimensional or higher? What if the two stable states are not just fixed points? In fact, in many neural models, there are slow conductances. Holding these constant often leads to subsystems which have as stable behavior both a fixed point and a rhythmic solution. Thus, jumping between states produces behavior which is alternately nearly constant and rapidly oscillating. Such behavior is called bursting and is important in

many physiological and biochemical systems. Wang and Rinzel (1995) use this decomposition of slow and fast systems to characterize and dissect a variety of bursting neurons. While the details are more complicated, the basic ideas are exactly as described in the example illustrated in Figure 3.

3.2 Averaged equations

. A related way to reduce complexity is to again exploit time scales and average, keeping only the slow variables. This approach is valid if the fast variables have a unique solution when holding the slow variables fixed. (In the previous section, the nonuniqueness of the roots of the fast equation were exploited in order to produce an oscillation.) That is, suppose the as the slow variables move around, at any given value of them, there is only a single stable behavior of the fast system. Then we can apply the methods of this section.

The basic idea can be gleaned by using the scaled time equations:

$$\frac{dX}{dT} = F(X, Y), \quad \frac{dY}{dT} = \epsilon G(X, Y).$$

For ϵ small the slow variables Y do not change too much so we can treat them as constant in the fast system. Then the fast system will evolve in time until it reaches a (possibly time-dependent) steady state, $X_{ss}(t, Y)$ which depends on the “parameters” Y . We plug this into the Y equation and obtain the slow system

$$\frac{dY}{dT} = \epsilon G(X_{ss}(t, Y), Y). \quad (3.1)$$

There are two typical situations. Either the fast system tends to a fixed point so that X_{ss} is independent of time; or the fast system is periodic in which case X_{ss} is periodic in time. In the former, the slow equation is reduced to:

$$\frac{dY}{dT} = \epsilon G(X_{ss}(Y), Y)$$

which is just an equation for the variables Y . We have eliminated X altogether. If the fast system tends to a time-dependent solution, such as a periodic solution, then we can use the ‘‘Averaging Theorem’’ which states that the behavior of

$$\frac{dY}{dT} = \epsilon G(Y, T) \text{ with } G(Y, T + P) = G(Y, T)$$

is close to the behavior of the averaged system

$$\frac{d\bar{Y}}{dT} = \epsilon \frac{1}{P} \int_0^P G(\bar{Y}, T) dT.$$

Formally, we can perform the same averaging even if the stimulus is not periodic but varies rapidly compared to ϵ . In any case, the resulting averaged equation depends only on the slow variable(s), Y .

There are many useful applications of this method; we present three of them.

3.2.1 Example 1: Hebbian learning.

The main mechanism for unsupervised learning is called Hebb’s rule. This states that the strength of connection between two neurons depends on the coincidence of activity between them. We will use the concepts of this subsection to derive standard model for the growth of weights to a single neuron. Let $V(t)$ be the potential of the neuron and let there be n inputs, $I_1(t), \dots, I_n(t)$ with corresponding weights, w_1, \dots, w_n . Thus at any given time, the potential satisfies:

$$\tau \frac{dV}{dt} = -V + \sum_j w_j I_j(t).$$

The inputs can be changing with time but we assume that the change is not real fast, so that

$$V(t) \approx \sum_j w_j I_j(t).$$

Hebb's rule, in this case, simply says that the weight, w_j will grow depending of the correlation between the input and the output, $V(t)I_j(t)$. Thus,

$$\frac{dw_j}{dt} = \epsilon F(w_j, V(t)I_j(t))$$

would represent a change in the weights. The simplest case is just linear growth:

$$\frac{dw_j}{dt} = \epsilon V(t)I_j(t) = \epsilon \sum_k I_k(t)I_j(t)w_k.$$

Since the growth rate is small, $\epsilon \ll 1$, we average over the inputs and obtain:

$$\frac{dw_j}{dt} = \epsilon \sum_k C_{kj}w_k$$

where C is the zero time correlation matrix of the inputs. The solution to a linear constant coefficient differential equation has the following form:

$$\vec{w}(t) = \sum_k \phi_k e^{\epsilon \lambda_k t},$$

where λ_k are the eigenvalues of C and ϕ_k the corresponding eigenvectors. If C is a positive matrix, the Perron-Frobenius theorem implies that the maximal eigenvalue of C has nonnegative components. Thus, the eigenvector corresponding to the largest eigenvalue will grow fastest and the weights will tend to be proportional to it. This is the basis of "Principal Component Analysis" used in statistics. By using averaging, we have reduced a complex problem to a simple exercise in linear algebra.

3.2.2 Example 2. Neural networks from biophysics.

The typical biophysical model of a network of neurons has the form

$$C \frac{dV_j}{dt} = -I_{ion,j} + I_{appl,j} - \sum_i g_{ij}^{ex} s_{ij}^{ex}(t)(V_j - E_{ex}) - \sum_i g_{ij}^{in} s_{ij}^{in}(t)(V_j - E_{in})$$

where the synapses satisfy

$$\frac{ds_{ij}^l}{dt} = \alpha_{ij}^l(V_i)(1 - s_{ij}^l) - \beta_{ij}^l s_{ij}^l.$$

Here, we have divided the synapses into excitatory and inhibitory although there can be more populations. Furthermore, the model cells are only single compartments but could easily be extended to more. How can we connect this to the simple notion of a neural network or so-called firing rate model in a quantitative manner?

Suppose that the synapses are *slow* relative to the dynamics of the membrane. (This is not in general a good assumption but for NMDA and GABA-B synapse, it may not be unreasonable.) Then the synapses can be held “frozen” as parameters and the fast dynamics of the neurons will reach a steady state of either a fixed point or periodic behavior. Let G_j^{ex} (G_j^{in}) be the total excitatory (inhibitory) conductance into cell j . Then the behavior of the cell is determined by just these two parameters, so that $V(t) = V_{ss}(t; G^{ex}, G^{in})$. Suppose that the time course of a synapse depends only on the presynaptic neuron. That is $s_{ij}(t) = s_i(t)$. Then,

$$\frac{ds}{dt} = \alpha(V_{ss}(t; G^{ex}, G^{in}))(1 - s) - \beta s.$$

We now average this and get equations for s

$$\frac{ds}{dt} = \bar{\alpha}(G^{ex}, G^{in})(1 - s) - \beta s.$$

To close the system, we note that G^{ex}, G^{in} are just linear sums of the synaptic gating functions, s . In Ermentrout (1994) we apply this technique to a specific model and obtains a neural-net-like set of equations. More recently, Chen et al derive a greatly simplified model for wave propagation in a thalamic network which produces nearly identical behavior to a large scale biophysically based computational model and *which can be solved in closed form*.

3.2.3 Example 3. Weakly coupled oscillators.

As a final example, I describe a technique that has been used in a variety of biological systems that comprise many coupled rhythmic elements. The idea is that if the elements are rhythmic, then one of the most important quantities is the relative timing of the rhythms. For example, in quadruped locomotion, the relative phases of the four limbs are precisely what define the gait of the animal. Similarly, in the swimming behavior of the lamprey, the key question is how is the metachronal traveling wave formed. This wave controls the timing of the muscular contractions required for locomotion. Since phase is the relevant experimentally measured quantity, models that consider only the phases as variables make sense. The problem is how to connect the abstract phase model with a concrete mechanistic model for the oscillators. In this section, we show how this is done by using averaging.

The model equations take the form

$$\frac{dX_j}{dt} = F_j(X_j) + \epsilon G_j(X_1, \dots, X_N)$$

We assume that each subsystem (which can represent many variables) has a periodic solution with all the periods identical. (They need not be identical, but we will absorb the differences in the G_j .) Furthermore, in order to make this mathematically rigorous, the strength of interaction between the oscillators must be “weak.” A natural question to ask is what does “weak” mean? There is no simple answer to this; heuristically it means that the coupling is not so strong as to distort the oscillation other than to shift its phase. That is, the waveform of the oscillating components should not be changed much by coupling, however they can be shifted.

Let $X_j^0(t)$ be the oscillation in absence of coupling. Then it can be proven (Kuramoto, 1989) that if ϵ is sufficiently small, the solution to the coupled system is $X_j^0(t + \theta_j)$ where θ_j is a phase-shift. The phase-shift satisfies a set

of equations of the form:

$$\frac{d\theta_j}{dt} = \omega_0 + \epsilon H_j(\theta_1 - \theta_j, \dots, \theta_N - \theta_j).$$

The parameter ω_0 is the uncoupled frequency and the functions, H_j are $2\pi/\omega$ periodic in each of their arguments. The beauty of this reduction is that even if the basic oscillators lie in a high-dimensional space, there is only one equation for each oscillator. Furthermore, the experimentally important variable, the phase or timing, is the only state variable in the reduced system. Computing H_j from a given model or experimental preparation is difficult but there is a nice heuristic way to do it. Suppose for simplicity that each oscillator is coupled only to a single component of the other oscillators. For example, if these represent neurons, then the coupling appears only in the somatic voltage equation. An easily computable (both experimentally and numerically) quantity for an oscillator is the phase response curve (PRC). Choose a visible event in your oscillator eg the appearance of a spike. Since the system oscillates, the spike occurs at $t = 0, P, 2P, \dots$ where P is the period of the oscillator. At $t = t_0 < P$ give the oscillator a brief stimulus. This will change the time of the next spike to say, P_0 . The PRC, χ is defined as

$$\chi(t_0) = \frac{P - P_0}{P}.$$

Thus the PRC measures the fraction of the period lost or gained as a function of the timing of the perturbation. The function H_j is now easily computed:

$$H_j(\theta_1 - \theta_j, \dots) = \frac{1}{P} \int_0^P \chi(t) [G_j(X_1^0(t + \theta_1 - \theta_j, \dots))]_1 dt$$

where $[G]_1$ is the first component (and by assumption, the only nonzero one) of G . That is, we just *average* the effects of the coupling against the effect of a perturbation.

One of the easiest applications is to look at a pair of mutually coupled identical oscillators and ask whether or not they synchronize:

$$\begin{aligned}\frac{d\theta_1}{dt} &= \omega + \epsilon H(\theta_2 - \theta_1) \\ \frac{d\theta_2}{dt} &= \omega + \epsilon H(\theta_1 - \theta_2)\end{aligned}$$

Let $\phi = \theta_2 - \theta_1$ and we get

$$\frac{d\phi}{dt} = \epsilon(H(-\phi) - H(\phi))$$

The zeros of the right-hand side are the allowable phase-shifts between the two oscillators. In particular $\phi = 0$ is always a root. It is stable if $H'(0) > 0$. Thus, one can ask what kinds of interactions lead to stable synchrony. This is a hot research topic and there have been many papers on the subject.

This approach has been used to study the swim generator of the lamprey where the circuitry is not known in detail. In spite of this lack of detailed knowledge, with very few assumptions, it is possible to suggest experiments and determine mechanisms by simply analyzing the general structure of chains of phase models. A review of this approach can be found in Williams and Sigvardt (1995).

4 Discussion and caveats

Averaging is an intuitively appealing method for reducing the complexity of biological systems which operate on many different time and space scales. There are mathematical techniques that can be brought to bear on complex models that enable us to reduce the dimensionality of the system and to connect one level of detail with another. There are additional mathematical methods that can be used to reduce models to lower dimensions. For example, normal form analysis (Hoppensteadt and Izhikevich, 1997) has been used

to reduce complex neural networks to simple “canonical models”, low dimensional equations that have all of the properties of their higher-dimensional relatives.

We have presented a number of case studies illustrating these techniques for simplifying as applied to specific biological systems. The advantages of the simplified models are quite obvious. At the very least, simulations involve far fewer equations and parameters; at best, a complete analysis of the behavior becomes possible. In many cases the information that they give is quantitative and thus the simplified models can and have been used to suggest specific experiments.

However, there are several questions that arise in this approach to modeling. When are important details being neglected? This is a very difficult question to answer since there will always be aspects that a simplified model cannot tell you but the more detailed model will. This is an obvious consequence of simplifying. The more dangerous problem is that sometimes these neglected details have consequences for the reduced model and then the behavior of the reduced model will be misleading. How can you know that important details are missing? The only way to know is to do simulations. A negative answer can never be definitive since there may always be a different instance that you haven't tried which produces a qualitatively different answer to your simplification. Finally, in the extreme case, the models can become so simplified and far from the original experimental system as to be phenomenological or metaphorical. This is a trap into which many modelers fall, particularly those who neglect to study any of the underlying biology. An entire decade of “Catastrophe Theory” models illustrates this phenomena. The best way to avoid the pitfalls inherent in simplification is to continue to maintain contact with the experimental results and at every step of the procedure attempt to justify and if possible quantify the assumptions made

in going from the details to the simplification.

References

- [1] L.F. Abbott and T.B. Kepler, 1990, Model neurons: from Hodgkin-Huxley to Hopfield, in *Statistical Mechanics of Neural Networks*, (L. Garrida, ed.) Springer-Verlag, Berlin, pp 5-18.
- [2] Z. Chen, B. Ermentrout, and XJ Wang, 1998, Wave propagation mediated by GABA-B synapse and rebound excitation in an inhibitory network: a reduced model approach, *J.Computat. Neuro.* 5:53-60.
- [3] J.D. Cowan, 1968, Statistical mechanics of neural nets, in ER Caimanello, ed, *Neural Networks*, 181-188, Springer-Verlag, Berlin.
- [4] L Edelstein-Keshet and GB Ermentrout, 1991, Models for contact-mediated pattern formation: Cells that form parallel array, *Differentiation* 29:33-58.
- [5] Ermentrout GB 1994 Reduction of conductance-based models with slow synapses to neural nets, *Neural Comp.* 6:679-695.
- [6] GB Ermentrout and L Edelstein-Keshet, 1993, Cellular automata approaches to biological modeling, *J. Theoretical Biology*,160:97-133.
- [7] Hoppensteadt F and Izhikevich E 1997 *Weakly Connected Neural Nets*, Springer-Verlag Berlin.
- [8] Kuramoto Y 1984 *Chemical Oscillations, Waves, and Turbulence*, Springer-Verlag New York.
- [9] Murray JD 1989 *Mathematical Biology*. Springer-Verlag New York.

- [10] D.J. Pinto, J.C. Brumberg, D.J. Simons and G.B. Ermentrout, 1996, A quantitative population model of whisker barrels: re-examining the Wilson-Cowan equations, *J. Comput. Neurosci.* 3:247-264.
- [11] J. Rinzel, 1985, Excitation dynamics: insights from simplified membrane models, *Fed. Proc.* 44, 2944-2946.
- [12] C. van Vreeswijk and H. Sompolinsky, 1998, Chaotic balanced states in a model of cortical circuits, *Neural Computation* 10:1321-1373.
- [13] XJ Wang and J. Rinzel, 1995, Oscillatory and bursting properties of neurons, in *The Handbook of Brain Theory and Neural Networks*, (MA Arbib, ed) MIT Press, pp 686-691
- [14] TL Williams and KA Sigvardt, Spinal cord of lamprey: generation of locomotor patterns, in the *Handbook of Brain Theory and Neural Networks* (op cit) pp 918-921.