Companion calculations for tutorial on reduced models

Alex Roxin

April 7, 2011

Contents

1 What this is

This is a rather bare-bones presentation of the calculations that were discussed in the tutorial. I obviously skipped lots of algebraic steps along the way but I tried not to leave out any conceptual steps.

2 A rate equation with input current

Using heuristic arguments one arrives at a two-component model for determining the firing rate of neurons self-consistently in a recurrent network. The equations are

$$\tau\frac{dI}{dt} = -I + r, \tag{S.1}$$

$$\frac{dr}{dt} = -r + \phi(wI) \tag{S.2}$$

where I is the input current and r is the mean firing rate in the network. The sign of the parameter w in the second equation tells us whether the recurrent connectivity is excitatory or inhibitory. That is, if the recurrent connections are excitatory, then an increase in the input to a cell will increase the firing rate, while in the inhibitory case, increasing the input decreases the firing rate. The steady-state values of the input and firing rate are $(I, r) = (\bar{I}, R)$, where $\bar{I} = R$ and so $R = \phi(wR)$. This second equation is the self-consistent constraint on the firing rate. Depending on the shape of the input-output function ϕ there may be only one or several solutions.

The stability of the steady-state solution (fixed point) can be determined by perturbing the solution with the ansatz $(I, r) = (\bar{I}, R) + (\delta I, \delta r)e^{\lambda t}$, where λ is the linear growth rate. Clearly, if $\lambda < 0$ then the solution is (linearly) stable, while if $\lambda > 0$ the perturbation will grow. This would indicate that the solution is linearly *unstable* and one therefore expects a new, different solution to emerge. When doing linear stability calculations, the only thing to remember is that nonlinear functions get Taylor-expanded. In this simple model there is only one nonlinearity. Therefore we expand $\phi(wI) = \phi(w\bar{I} + w\delta I) = \phi(w\bar{I}) + \phi^{'}(w\bar{I})\delta I + \ldots$ Note that the perturbative term has a prefactor which is the slope of the function ϕ evaluated at the fixed point. This slope is known as the *gain* which gives the effect of a small change in input on the output. In the rest of this write-up I will leave it implicit that all functions are evaluated at the fixed point when calculating stability, i.e. I will just write $\phi^{'}$. Higher order, nonlinear terms in δI are not kept because this is a *linear* stability analysis. Later on I will discuss a calculation where higher order terms are kept. It is then a much more involved calculation.

After expanding ϕ we find that the leading order terms cancel since they are just the fixed point equations. We are left with the linear system

$$\begin{pmatrix} \tau\lambda + 1 & -1 \\ -w\phi^{'} & \lambda + 1 \end{pmatrix} \begin{pmatrix} \delta I \\ \delta r \end{pmatrix} = 0. \tag{S.3}$$

Usually in linear algebra we are faced with equations such as $\mathcal{A}\mathbf{x} = \mathbf{b}$, where $\mathcal{A}$ is a matrix and $\mathbf{x}$ and $\mathbf{b}$ are vectors. Then, we just find the solution by inverting the matrix to get $\mathbf{x} = \mathcal{A}^{-}\mathbf{b}$. Here that is clearly not possible. In this case it must be that both of the algebraic equations are equal to zero, otherwise there would be no solution. It's easy to show that this is equivalent to setting the determinant of the matrix equal to zero. This gives

$$\tau\lambda^2 + \lambda(1+\tau) + 1 - w\phi^{'} = 0. \tag{S.4}$$

Solving for λ, one finds

$$\lambda = -\frac{(1+\tau)}{2\tau} \pm \frac{1}{2\tau}\sqrt{(1-\tau)^2 + 4\tau w\phi'}. \tag{S.5}$$

From Eq.S.5 it is clear that only steady instabilities are possible in this simple model, not oscillatory ones. How does one see this? A steady instability means that as some parameter is varied, one of the growth rates or *eigenvalues* λ passes through zero along the real axis (it has no imaginary component). The resulting instability causes the firing rate to 'blow up'. This happens here for $w\phi^{'} > 1$ which can only happen in an excitatory network $w > 0$. On the other hand, an oscillatory instability occurs if the real part of λ crosses zero with a non-zero imaginary part. Then there will be a pair of complex eigenvalues. The resulting instability oscillates as it grows. Here this could only happen if the argument of the square-root function became negative. This happens for $|w|\phi^{'} > \frac{(1-\tau)^2}{4\tau}$ as long as the network is inhibitory ($w < 0$). However, when this occurs, the real part of the eigenvalue is just $-\frac{(1+\tau)}{2\tau}$ which is always negative. Therefore, this predicts that when the recurrent connectivity in an inhibitory network is sufficiently strong, the response of the system will be oscillatory, but damped.

3 A rate equation for an inhibitory network with delay

Seeing as the rate equation with current for an inhibitory network never exhibits any instabilities, we might as well just use a one-component model for the rate. However, numerical simulations of recurrently connected networks of inhibitory neurons robustly show oscillatory dynamics. In the tutorial I showed this was due to an *effective* delay in the response of postsynaptic cells. This can be captured by including a fixed time-delay in the rate equation. Doing so gives

$$\dot{r} = -r + \Phi(-wr(t-D) + I), \tag{S.6}$$

where the delay is D. Note that in this equation, as before, I have rescaled time by the time constant of the rate dynamics. This should be roughly the time-constant of the membrane potential which is $5-20$ms for cortical cells. Therefore the parameter D is actually the *ratio* of the effective delay to the membrane time constant and is small.

The steady-state $r = R$ is determined self-consistently by $R = \Phi(-wR + I)$. The linear stability of the steady-state is found with the ansatz $r = R + \delta r e^{\lambda t}$. Plugging this into Eq.S.6 one has

$$\lambda = -1 - w\Phi' e^{-\lambda D}, \tag{S.7}$$

which is a transcendental equation for the eigenvalue λ. Since we are interested in oscillatory instabilities we take $\lambda = i\omega$, which places us right on the stability boundary. Separating the real and imaginary parts in Eq.S.7 yields

$$\begin{aligned} \omega &= w\Phi' \sin \omega D, && \text{(S.8)} \\ 1 &= -w\Phi' \cos \omega D. && \text{(S.9)} \end{aligned}$$

Dividing the first equation through by the second reveals that the frequency of oscillation depends on the delay as $-\tan \omega D = \omega$. In the limit $D \to 0$ it is easy to see (draw $-\tan \omega D$ and ω and see where they intersect), that $\omega \to \frac{\pi}{2D}$. Actually there are infinitely many roots to this equation, but the instability with the lowest frequency is the first to go unstable. This is because from Eq.S.8, the critical values of the coupling and gain at the instability boundary are equal to the frequency.

4 Rate equations for two-choice decision making

The simplest dynamical model for two-choice decision making with some biological relevance would take the form of three coupled equations

$$\begin{aligned} \dot{r}_1 &= -r_1 + \Phi(w_{11}r_1 + w_{12}r_2 - w_{13}r_I + \nu_1), && \text{(S.10)} \\ \dot{r}_2 &= -r_2 + \Phi(w_{22}r_2 + w_{21}r_1 - w_{23}r_I + \nu_2), && \text{(S.11)} \\ \tau\dot{r}_I &= -r_I + \Phi_I(w_{31}r_1 + w_{32}r_2 - w_{33}r_I + \nu_I), && \text{(S.12)} \end{aligned}$$

where r_1 and r_2 are the firing rates of neurons in populations which encode the developing decision for the two possible choices and r_I is the activity of a population of inhibitory neurons. Again, I have normalize time by the time constant of the two excitatory populations, and I'm using ν instead of I for the current so as not to confuse it with the inhibition. For the sake of simplicity we can make some assumptions. First, take $\Phi_I(x) = [x]_+$ which is a linear function if $x > 0$ and otherwise zero. Inhibitory neurons tend to fire at higher rates than excitatory neurons in the cortex, and this puts them in a regime where the single-cell fI curve can be approximated as linear, so this is not a bad assumption. Also, take $\tau \rightarrow 0$ which means that the inhibition is much faster than excitation. Then, we can solve for r_I and plug it into the other two equations. Additionally, we assume that the connectivity between r_1 and r_2 is symmetric, i.e. there is no hard-wired bias for r_1 or r_2 'winning'. Then the equations can be written

$$\begin{aligned} \dot{r}_1 &= -r_1 + \Phi(sr_1 - cr_2 + \nu_1), \quad &\text{(S.13)} \\ \dot{r}_2 &= -r_2 + \Phi(sr_2 - cr_1 + \nu_2), \quad &\text{(S.14)} \end{aligned}$$

where s is the effective self-coupling, c is the effective inhibitory cross-coupling and ν is the effective external drive. A simple exercise would be to relate these parameters to those in the original, three-component system.

4.1 Linear Stability

The structure of the steady-state solutions in Eqs.S.13-S.14 is a bit more interesting than those seen in the previous sections, and this is because of a symmetry of the system. First, we will assume $\nu_1 = \nu_2 = \nu$, so both populations receive the same input. Then there is a fixed point where both populations have the same activity $(r_1, r_2) = (R, R)$, where $R = \Phi((s-c)R + \nu)$. Both there also may be a fixed point $(r_1, r_2) = (R_{\text{low}}, R_{\text{high}})$. If so, then you can show there also must be a fixed point $(r_1, r_2) = (R_{\text{high}}, R_{\text{low}})$, and this is due to the fact that the system is reflection symmetric, i.e. it is invariant under the transformation $(r_1, r_2) \rightarrow (r_2, r_1)$. This is precisely the property one would want in a two-choice decision-making circuit. The idea is that the system sits initially at $(r_1, r_2) = (R, R)$ (the 'spontaneous state') and then the sensory input drives it towards one of the two reflection symmetric fixed points. Whichever population ends up in the R_{high} fixed point 'wins'. How does this happen? It depends on the way in which the two new fixed points appear and how stable the 'spontaneous state' is. Now we will look at the stability of the spontaneous state.

The stability is found by taking $(r_1, r_2) = (R, R) + (\delta r_1, \delta r_2)e^{\lambda t}$. Plugging this into Eqs.S.13-S.14 leads to the linear system

$$\begin{pmatrix} \lambda + 1 - s\Phi^{'} & c\Phi^{'} \\ c\Phi^{'} & \lambda + 1 - s\Phi^{'} \end{pmatrix} \begin{pmatrix} \delta r_1 \\ \delta r_2 \end{pmatrix} = 0. \tag{S.15}$$

Again, this system only has a solution if the determinant of the matrix is equal to zero. This leads to a quadratic equation in λ, the roots of which are $\lambda = -(1 - s\Phi^{'}) \pm c\Phi^{'}$. A steady instability (which presumably leads to one of the other two stable fixed points) occurs when $c\Phi^{'} = 1 - s\Phi^{'}$. This could happen if the external drive to both populations is increased, say in response to some sensory input. Why? Because the spontaneous state is a low-rate state. For low rates, any reasonable Φ should be concave up (like x^2 as opposed to $\sqrt{x}$). This is the case for cortical pyramidal cells, for example. And this means that the gain increases with increasing input. Therefore $c\Phi^{'}$ gets bigger and $1 - s\Phi^{'}$ gets smaller until, for some large enough input, they are equal and there is an instability. At the bifurcation then, the eigenvalues are

$$\lambda_c = 0, \tag{S.16}$$

$$\lambda_s = -2c\Phi^{'}. \tag{S.17}$$

The eigenvector associated with each eigenvalue can be found by plugging the corresponding eigenvalue into Eq.S.15. In the case of λ_c (c for critical) this yields

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} r_1^c \\ r_2^c \end{pmatrix} = 0, \tag{S.18}$$

the solution of which is $\mathbf{r}_c = (r_1^c, r_2^c) = (1, -1)$, while for λ_s (s for stable) one has

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} r_1^s \\ r_2^s \end{pmatrix} = 0, \tag{S.19}$$

the solution of which is $\mathbf{r}_s = (r_1^s, r_2^s) = (1, 1)$. This means that at the bifurcation, perturbations which lead to an increase in the activity of one population and a decrease in that of the other will not decay away, while simultaneous increases and decreases in their activity are strongly damped. This is precisely the type of dynamics which lead to a ‘winner-take-all’ behavior.

4.2 Center manifold reduction

We have now determined that as the external input to the two populations of neurons is increased, at some point a bifurcation occurs for which the activity in the two populations will be strongly negatively correlated. Specifically, fluctuations which cause the activity in one population two increase at the expense of the other will not decay away, while those which increase or decrease both simultaneously are highly damped. This type of behavior, whereby fluctuations are strongly damped except along a particular direction (one speaks of a particular *manifold*), is, in fact, generic at a bifurcation. One says that the other directions are *slaved* to the dynamics along the critical manifold. Now we will go through a calculation which will yield the approximate functional form of the critical manifold, as well as an equation which gives the dynamics along this manifold. This is the dynamics which describes the developing decision. The method we will use is called *Center Manifold Reduction* (CMR). For practical applications I personally prefer another method called *Multi-scale Analysis* which yields the same results but which is more widely applicable and more flexible. Nonetheless, CMR is much more intuitive, being entirely geometric in nature. Basically, we will do a change of coordinates so that the dynamics along the center manifold, the dynamics we care about, is in the X direction. Then, the dynamics orthogonal to the X-axis, which in the case of this simple 2D system is just the Y direction (although there will be a slight twist), just gets projected (slaved) onto the X direction. That's it. Very simple conceptually. We will need to do some linear algebra, and then some straightforward but rather involved algebra.

The first step is to expand the nonlinearities in a Taylor series. As promised, we are extending the linear analysis to higher order. Specifically, assuming the ansatz $(r_1, r_2) = (R_{cr}, R_{cr}) + (\bar{r}_1, \bar{r}_2)$ and $\nu = \nu_{cr} + \bar{\nu}$, where R_{cr} and ν_{cr} are the firing rate and the input right at the bifurcation point, and $\bar{r}_1, \bar{r}_2, \bar{\nu} \ll 1$ we can write

$$\Phi(sr_1 - cr_2 + \nu) = \Phi_0 + \Phi_0^{'}(s\bar{r}_1 - c\bar{r}_2 + \bar{\nu}) + \frac{\Phi_0^{''}}{2}(s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^2 + \frac{\Phi_0^{'''}}{6}(s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^3 + \ldots, \tag{S.20}$$

and we won't need any higher order terms. Also, the subscript 'naught' means that the function is evaluated at the bifurcation point. Then, we find that all of the leading order terms cancel since they are just the fixed point equations. This leaves the following system of equations

$$\dot{\mathbf{r}} = \mathcal{L}\mathbf{r} + \mathbf{N}(\mathbf{r}), \tag{S.21}$$

where $\mathbf{r} = (\bar{\nu}, \bar{r}_1, \bar{r}_2)$ and

$$\mathcal{L} = c\Phi_0^{'} \begin{pmatrix} 0 & 0 & 0 \\ 1/c & -1 & -1 \\ 1/c & -1 & -1 \end{pmatrix}, \tag{S.22}$$

$$\begin{aligned} \mathbf{N}(\mathbf{r}) &= \frac{\Phi_0^{''}}{2} \begin{pmatrix} 0 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^2 \\ (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^2 \end{pmatrix} \\ &\quad + \frac{\Phi_0^{'''}}{6} \begin{pmatrix} 0 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^3 \\ (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^3 \end{pmatrix}. \end{aligned} \tag{S.23}$$

The system has now three components as opposed to two. Why? This is a trick. We are now treating the input as a state variable. Basically this overcomes what Crawford refers to as a an 'awkward discrepancy' [1]. Namely, in our analysis we consider the dynamics in some interval around the bifurcation point, but the center manifold only exists *exactly* at the bifurcation point, i.e. for $\bar{\nu} = 0$. So instead of varying $\bar{\nu}$ as a parameter, which would cause the center manifold to disappear, we treat it as a variable. It turns out this is just fine to do. If this point is hard to grasp, it doesn't matter. Formally we haven't changed *anything*, we've just rewritten the equations.

Now we rotate our coordinate system. You may remember from linear algebra that if a matrix is diagonalizable, then it can be expressed as $\mathcal{L} = \mathcal{Q}\mathbf{\Lambda}\mathcal{Q}^{-1}$, where $\mathbf{\Lambda}$ is the matrix of eigenvalues, and $\mathcal{Q}$ is the column matrix of eigenvectors. Once we have found them we can apply the matrix $\mathcal{Q}^{-1}$ from the left in Eq.S.21 which gives

$$\mathcal{Q}^{-1}\dot{\mathbf{r}} = \mathcal{Q}^{-1}\mathcal{Q}\mathbf{\Lambda}\mathcal{Q}^{-1}\mathbf{r} + \mathcal{Q}^{-1}\mathbf{N}(\mathbf{r}), \tag{S.24}$$

$$\dot{\mathbf{X}} = \mathbf{\Lambda}\mathbf{X} + \mathcal{Q}^{-1}\mathbf{N}(\mathbf{X}), \tag{S.25}$$

where $\mathbf{X} = (\nu, X, Y) = \mathcal{Q}^{-1}\mathbf{r}$. It just takes a little work to find the eigenvalues and eigenvectors (hint: two of them are just the same as before) and show that

$$\mathbf{\Lambda} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2c\Phi_0' \end{pmatrix}, \tag{S.26}$$

$$\mathcal{Q} = \begin{pmatrix} 2c & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \tag{S.27}$$

$$\mathcal{Q}^{-1} = \begin{pmatrix} 1/2c & 0 & 0 \\ 0 & 1/2 & -1/2 \\ -1/2c & 1/2 & 1/2 \end{pmatrix}, \tag{S.28}$$

and so $\eta = \bar{\nu}/2c$, $X = (r_1 - r_2)/2$ and $Y = (r_1 + r_2)/2 - \bar{\nu}/2c$.

Also

$$\begin{aligned} \mathcal{Q}^{-1}\mathbf{N} &= \frac{\Phi_0''}{4}\begin{pmatrix} 0 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^2 - (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^2 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^2 + (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^2 \end{pmatrix} \\ &+ \frac{\Phi_0'''}{12}\begin{pmatrix} 0 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^3 - (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^3 \\ (s\bar{r}_1 - c\bar{r}_2 + \bar{\nu})^3 + (s\bar{r}_2 - c\bar{r}_1 + \bar{\nu})^3 \end{pmatrix} \end{aligned} \tag{S.29}$$

Now we need to express $\mathbf{r}$ as a function of $\mathbf{X}$. One finds that $\bar{\nu} = 2c\eta$, $\bar{r}_1 = \eta + X + Y$ and $\bar{r}_2 = \eta - X + Y$. Then, after some algebra one can write

$$\begin{aligned} \mathcal{Q}^{-1}\mathbf{N} &= \frac{\Phi_0''}{4}\begin{pmatrix} 0 \\ 4(s+c)^2\eta X + 4(s^2 - c^2)YX \\ 2(s+c)^2(\eta^2 + X^2) + 4(s^2 - c^2)\eta Y + 2(s-c)^2Y^2 \end{pmatrix} \\ &+ \frac{\Phi_0'''}{12}\begin{pmatrix} 0 \\ 2X^3 + \dots \\ \dots \end{pmatrix}. \end{aligned} \tag{S.30}$$

Wherever there are dots it means you won't need these higher order terms. Next, we consider the equation for Y. It is

$$\dot{Y} = -2c\Phi_0'Y + \frac{\Phi_0''}{4}\Big(2(s+c)^2(\eta^2 + X^2) + 4(s^2 - c^2)\eta Y + 2(s-c)^2Y^2\Big) + \dots \tag{S.31}$$

Note that the equation for η is just $\dot{\eta} = 0$ so there is nothing to be done there. Compared to X, which has a linear growth rate of 0, perturbations in Y decay away quickly to a manifold $Y_{CM} = h(\eta, X)$. This is just the center manifold. We expand Y in powers of η and X and solve order by order. It turns out that to leading order the only nonzero term is $Y = h(\eta, X) = h_{02}X^2$. You can do a proper expansion for all orders and find they will be set to zero if you like. Then, we plug this into Eq.S.31 keeping in mind that $\dot{Y} = \frac{\partial h}{\partial \eta}\dot{\eta} + \frac{\partial h}{\partial X}\dot{X}$. I will skip the algebra here, which you can fill in. The answer is

$$Y = \frac{\Phi_0^{''}(s+c)^2}{4c\Phi_0^{'}}X^2, \tag{S.32}$$

and so the center manifold is quadratic in X. Now, we plug this result into the equation for X and retain only the leading order terms. After not too much work you will find

$$\dot{X} = \Phi_0^{''}\frac{(s+c)^2}{2c}\bar{\nu}X + \Big(\frac{(\Phi_0^{''})^2(s^2-c^2)(s+c)^2}{4c\Phi_0^{'}} + \frac{\Phi_0^{'''}}{6}(s+c)^3\Big)X^3 + \dots \tag{S.33}$$

Recall that $X = (r_1 - r_2)/2$. So this equation gives you the nonlinear dynamics of the difference in the activities of the two populations. This is precisely what one would like. Also note that while the parameter dependence of the coefficients depends on the model we choose, the form of the equation *does not*. Therefore, if we choose another model for two-choice decision making, which is different in the details we will still end up with the equation $\dot{X} = \alpha\bar{\nu}X + \beta X^3$. Only α and β will have changed.

You may have noticed through all of this that the input to both of the populations is identical which admittedly is not the most typical situation for decision making tasks. Usually, there is a slight bias in the sensory input. This is easily incorporated into the CRM framework. Consider a small bias in inputs such that $\nu_1 = \nu_{cr} + \bar{\nu} + \Delta\nu_1$ and $\nu_2 = \nu_{cr} + \bar{\nu} + \Delta\nu_2$. We will treat the bias term as a nonlinearity, i.e. we will include it in $\mathbf{N}$. After rotating the coordinates, it leads to term

$$\mathbf{N}_{\text{bias}} = \frac{\Phi_0^{'}}{2}\begin{pmatrix} 0 \\ \Delta\nu_1 - \Delta\nu_2 \\ \Delta\nu_1 + \Delta\nu_2 \end{pmatrix} \tag{S.34}$$

which in the equation for X adds a term $\frac{\Phi_0^{'}}{2}(\Delta\nu_1 - \Delta\nu_2)$. The same can be done for any noise term present in the rate equations. The assumption is that the noise amplitude is small.

References

[1] J. D. Crawford. Introduction to bifurcation theory. *Rev. Mod. Phys.*, 63:991–1037, 1993.