9.49 / 9.490 Neural Circuits for Cognition
Homework 2
Due: Wednesday October 16, by midnight.

In this assignment, we will explore bistable switches, linear integration, and low-pass filtering. You will use the tools of numerical simulation and graphical and linear stability analysis to analyze the fixed points of a system and their stability. You will also explore state-space plots,.

1) **Bistable switch, neural integrator.** Consider the case of a single neuron with an autapse (a synapse that connects back to the same neuron), with a sigmoidal firing rate function given by $f(g) = e^g/(1 + e^g)$. The rate-based equation for the summed synaptic input g can be written as

$$\tau \frac{ds}{dt} = -s + f(Ws + b) \tag{1}$$

where W is strength of the autapse and b is the sum of the cell's threshold/baseline activation with any feedforward drive to the cell.

a. Let $b = b_0 = -W/2$. Draw (by hand) the curves s and the curve $f(Ws + b)$ as a function of s, taking care to show the intercepts and asymptotic values of curves where appropriate. Derive the threshold value W^* for values of W above which the system exhibits bistability and below which it does not. Explain which fixed points are stable and why. [Hint: the sigmoid $f(x)$ and the line x intersect at three points only when the sigmoid is steeper than the line at $x = 1/2$. Thus, the critical point will occur when the curves have the same slope.] Thus, only when positive feedback is sufficiently strong does the system develop interesting bistable behavior.

b. Bifurcation: What happens for $W < W^*$? Graphically show all the fixed points (draw by hand), and explain their stability by graphical stability analysis. Make a plot (by hand, so qualitative) of locations of the fixed points as a function of W, as W sweeps from below to above W^*. Use one kind of marker for the stable fixed points, and another for the unstable ones. This plot illustrates why the dynamical system is said to undergo a "pitchfork bifurcation" at the critical parameter value W^*.

c. Simulate the autapse, and numerically verify your analytical results about W from [a.] and [b.]: that is, show that the value W^* is indeed the critical value and show what happens for W above and below the critical value.

d. Hysteresis: If the bistable switch is initialized at one of its stable fixed points, it takes a fairly strong drive to push it out of its *basin of attraction* into that of the other stable state. Try this in your simulation from [c.]: initialize the system in the lower state, and give a series of well-spaced input pulses $\Delta b(t)$, of duration τ each and amplitudes ranging from $0.1w$ to w. Generate three plots from your simulation of the following variables as a function of time: s, the firing rate ($r = f(Ws + b)$), and the pulses Δb.

e. Assuming some fixed value of $W > W^*$, derive what increment δb must be added to b_0 to destabilize the lower stable state. [Hint: b slides the sigmoid left/right. First graphically determine which direction of sliding eliminates the lower state as a fixed point, and visualize what the condition must be for the elimination to occur – similar to the condition above on the slopes of the two curves. Then implement this condition algebraically to find analytically when the lower fixed point ceases to exist.] Verify this analytically derived value by comparison with your numerical results from [d.].

2) **Linear stability analysis.**

a. Perform a linear stability analysis around each fixed point of problem 1a, above and below W^*. Verify through linear stability analysis your conclusions from graphical stability analysis, including as a function of W and b. Compare your results with those from Problem 1.

3) **Linear neural integrator, leaky integration, low-pass filtering.**

a. Integrator: Adapt your code from Problem 1 above to convert the bistable switch into a perfect integrator. An integrator must have a continuum of fixed points, rather than just two. For this the neural transfer function must be linear. (Recall from earlier in the course – our linear algebra section – that linear systems can have zero, one, or infinitely many solutions. By contrast, nonlinear systems can have more than one but not infinitely many solutions.) You will also need to adjust the weight W to cancel the exponential decay of the neuron: What is this critical value W^*? (Show by writing down the autapse equation with linear transfer function.) Experiment with and show the response (output) of the integrator in your code to a series of short pulses (positive and negative), and to a sine wave of period equal to τ. Interpret. Plot s and the firing rate r.

b. Mutual inhibition and state-space analysis: Change your code from a linear network with a single neuron to one with two neurons inhibiting each other

symmetrically with weights

$$\mathbf{W} = \begin{pmatrix} 0 & -w \\ -w & 0 \end{pmatrix} \tag{2}$$

Derive the condition on w to have one integrating mode and one leaky (stable) one; (refer back to class notes if needed; show your derivation), and set w to this value. Use the series of short input pulses (positive and negative) as before, to demonstrate that the system is an integrator. Note that inputs are now vectors in two dimensions, $b(t) = (b_1(t), b_2(t))^T$. The input must have some projection onto the integrating mode (eigenvector): you can choose them to be parallel to the integrating mode. Make a state-space plot: plot $s_1(t)$ against $s_2(t)$ across the sequence of pulses and responses; use discrete markers for each time-point, not a continuous line across time. Do the same for $r_1(t)$ against $r_2(t)$. Indicate on this plot the integrating mode (eigenvector) and the leaky mode (eigenvector); describe what you see. Now set the input kicks b to be parallel to some random direction in the 2-dimensional space and interpret what you see in the state space: how is the (s_1, s_2) plot different from when b is aligned to only the integrating mode?

c. Leaky integrator as a low-pass filter: Returning to the linear autapse, if $W < W^*$, the integrator is leaky. The equation for the dynamics, driven by an input $b(t)$, can be written as:

$$\tau' \frac{dx}{dt} = -x + b'(t) \tag{3}$$

What are the network time-constant τ' and input gain b' in terms of the original quantities? Now, set $W^* = 0.5$ and numerically integrate this equation for sinusioidal input of period T and input amplitude 1. Find the amplitude of the response s of the network, repeat for different periods of the sine (be sure to use values much smaller than and much larger than τ'), and make a plot of amplitude versus frequency. Equation 3 is the equation of a linear low-pass filter of the inputs b: A low-pass filter lets through inputs of low frequency, and attenuates inputs of high frequencies. Explain briefly why your numerical result is consistent with this name.

d. The solution to Equation 3 is given by $s(t) = \mathrm{conv}\,(b, K)\,(t)$, where $K(t)$ is an exponentially decaying kernel (filter) with time-constant τ': $K(t) = \frac{1}{\tau'} e^{-t/\tau'}$, and conv refers to the convolution operation (recall that $\mathrm{conv}(u, v)(t) = \int_{-\infty}^{\infty} u(\tau) v(t-\tau) d\tau = \int_{-\infty}^{\infty} u(t-\tau) v(\tau) d\tau$)). Derive $\mathrm{conv}\,(b, K)\,(t)$, where $b(t)$ is the cosine function $\cos(\omega t)$ with frequency ω. [*Hint:* Write $\cos(x) = \frac{e^{ix}+e^{-ix}}{2}$, where i is the

imaginary number defined by $i = \sqrt{-1}$; this rewriting makes it easy to evaluate the integral in the convolution.] In the end, express your result with real denominators only (note that $(a + ib)(a - ib) = a^2 + (ib)^2 = a^2 - b^2$). What does convolution with the exponential filter do to cosine waves of different frequencies? Plot this analytical solution together with the numerical result from [c.] – they should be an excellent match!