What you should have learned from my lectures

Peter Latham, March 28, 2022

1 Biophysics

1.1 The basic equations

In biophysics, there are only two equations. The first one relates the charge, Q, that accumulates across a bounded surface (such as the membranes surrounding neurons, dendrites and axons) to the voltage difference between the inside and outside,

$$Q = CV \tag{1}$$

where C is the capacitance. This equation assumes that the voltage is the same everywhere. The "same everywhere" assumption doesn't hold *exactly*. Imagine, for instance, injecting a small amount of charge into a neuron. For a brief period the voltage near the point of injection will be different than it is everywhere else. But if the neuron isn't very big, that voltage difference won't last very long. So Eq. (1), is valid for small objects such as neurons, but it breaks down in dendrites and axons, at least in the long direction.

Differentiating both sides with respect to time, and noting that the rate of change of charge is, by definition, current ($dQ/dt = I$ where I is current), we have

$$C\frac{dV}{dt} = I\,. \tag{2}$$

Here signs matter. We'll adopt the convention that voltage, V, is relative to the inside of the cell. Consequently, if current flows into the cell the voltage should go up, and if current flows out of the cell the voltage should go down.

Now we just need to know how the current depends on the voltage, and we'll have an equation for voltage in terms of voltage. If this were regular old physics, where the charge carriers are electrons, we would have $V = IR$, or $I = V/R$. However, in neurons charge carriers are ions, and the concentration of ions is different on the inside and the outside of the cell. (The concentration imbalance is maintained by ion pumps, which accounts for a big chunk of the energy your brain uses. But that's a detail we'll ignore.) Because of the concentration imbalance, current would flow even if the voltage were zero. The voltage required to keep current from flowing is known as the reversal potential, and it's denoted $\mathcal{E}$. Consequently, to good approximation, the current flow is proportional $V - \mathcal{E}$.

But it doesn't quite end there: different ions have different relative concentrations, and so different reversal potentials. They also have different conductances. The three main ionic species are Na^+, K^+ and Cl^-, which stand for sodium, potassium and chloride, respectively; if we included only these, the current would be

$$I = -g_{Na}(V - \mathcal{E}_{Na}) - g_K(V - \mathcal{E}_K) - g_{Cl}(V - \mathcal{E}_{Cl})\,. \tag{3}$$

The leading minus sign arises because we're taking I to be the outward current. There's a lot of sodium and chloride on the outside (think "salt outside the cell") and a lot of potassium on the inside. Consequently, the reversal potential for sodium, $\mathcal{E}_{Na}$, is positive, and the reversal potentials for potassium and chloride, $\mathcal{E}_K$ and $\mathcal{E}_{Cl}$, are negative. You should convince yourself that these last statements about the reversal potentials are correct, based on the relative concentrations inside and outside the cell.

In the general case, we write

$$I = -\sum_x g_x (V - \mathcal{E}_x) \tag{4}$$

where g_x is the conductance of ion x and $\mathcal{E}_x$ is its reversal potential. Inserting this into Eq. (2), and also adding an external current $I_0(t)$ (taken to be inward, by convention), we have

$$C\,\frac{dV}{dt} = -\sum_x g_x (V - \mathcal{E}_x) + I_0(t)\,. \tag{5}$$

This is a fundamental equation, based almost solely on arguments from physics. The interesting part comes from the conductances, which can be either constant, voltage-dependent or neurotransmitter-dependent. Those dependencies just have to be memorized, since they come from experimental observations.

1.2 Constant conductances

In the simplest case all conductances are constant; in that case the neuron is considered to be *passive*. For constant conductances, Eq. (5) can be written

$$\tau\,\frac{dV}{dt} = -(V - \mathcal{E}_L) + g_L^{-1} I_0(t)\,, \tag{6}$$

with τ, $\mathcal{E}_L$ and g_L suitably defined in terms of the g_x and $\mathcal{E}_x$ (you should derive explicit expressions). The solution to this equation is

$$V(t) = \mathcal{E}_L + V(t=0)e^{-t/\tau} + \int_0^t \frac{dt'}{\tau} e^{-(t-t')/\tau} g_L^{-1} I_0(t')\,, \tag{7}$$

which can be written in several other forms,

$$\begin{aligned} V(t) &= \mathcal{E}_L + \int_{-\infty}^t \frac{dt'}{\tau} e^{-(t-t')/\tau} g_L^{-1} I_0(t') \\ &= \mathcal{E}_L + \int_0^\infty \frac{ds}{\tau} e^{-s/\tau} g_L^{-1} I_0(t-s)\,. \end{aligned} \tag{8}$$

You should verify that all three of these are solutions to Eq. (6).

Passive conductances are also important for the propagation of current and voltage in dendrites and axons. For that, see http://www.gatsby.ucl.ac.uk/~pel/tn/notes/biocables.pdf .

1.3 Voltage-dependent conductances

Things get more interesting when conductance depend on voltage. The most common model for that (at least in neuroscience) is what I call a charged ball-on-a-stick model, where the stick/ball can be either open (allowing charged ions to pass) or closed (not allowing them to pass). Because the ball is charged, the transition probability between the open and closed states is a function of voltage,

$$\text{probability of going from closed to open in time } dt = \alpha(V) \tag{9a}$$

$$\text{probability of going from open to closed in time } dt = \beta(V)\,. \tag{9b}$$

If the ball has positive charge, $\alpha(V)$ is an increasing function of V and $\beta(V)$ is a decreasing function of V; if the ball has negative charge it's the other way around. From the transition probabilities, you should be able to show that the open probability, denoted x, evolves according to

$$\tau_x(V)\frac{dx}{dt} = -\big(x - x_\infty(V)\big) \tag{10}$$

where

$$\tau_x(V) = \frac{1}{\alpha(V)+\beta(V)} \tag{11a}$$

$$x_\infty(V) = \frac{\alpha(V)}{\alpha(V)+\beta(V)}\,. \tag{11b}$$

You should also be able to show that if both $\alpha(V)$ and $\beta(V)$ are exponential functions of V, then $x_\infty(V)$ is sigmoidal.

The full story is slightly more complicated: a channel is made up of more than one ball-on-a-stick, the balls-on-a-stick open and close independently, and all have to be open for current to flow. Taking all that into account leads to the Hodgkin-Huxley equation,

$$\tau\frac{dV}{dt} = -(V-\mathcal{E}_L) - \rho_{Na}m^3h(V-\mathcal{E}_{Na}) - \rho_K n^4(V-\mathcal{E}_K) + g_L^{-1}I_0(t) \tag{12}$$

where the m, h and n channels obey equations with the form of Eq. (10). You should know where this equation came from, the approximate time constants, the shapes of the curves $m_\infty(V)$, $h_\infty(V)$ and $n_\infty(V)$, and the approximate values of τ, ρ_{Na}, ρ_K, and the reversal potentials I'll tell you the last three: 10 ms, 400 and 120, respectively.

1.4 Concentration-dependent conductances

Conductances can also depend on the concentration of neurotransmitters and neuromodulators (don't ask me what the difference is); this is especially important for postsynaptic terminals, whose main job is to respond to neurotransmitters. Here the equation is pretty much what one should expect, given the above discussion,

$$\tau\frac{dx}{dt} = -c(t)(1-x) - \beta x \tag{13}$$

where $c(t)$ goes up rapidly when neurotransmitter or neuromodulator is nearby and drops rapidly when it isn't ("rapidly" generally means sub-millisecond). In class we talked about the release of neurotransmitter into the synaptic cleft, with the main neurotransmitters being AMPA, NMDA, GABA$_A$ and GABA$_B$. The first two are excitatory (they activate receptors with a reversal potential near 0 mV) and the second two are inhibitory (they activate receptors with a reversal potential near -80 mV). Neuromodulators can be released more broadly, and have all sorts of complicated effects (including on plasticity) that we don't understand very well.

The variable x, which is dimensionless, lies between 0 and 1; to get the current, we need to multiply by the maximum conductance, modulated by the reversal potential. In the general case, for neurotransmitter of type z,

$$I_z = -g_z x_z (V - \mathcal{E}_z)\,. \tag{14}$$

Note that the reversal potential depends on neurotransmitter type. This didn't have to be the case, but for reasons we don't understand that's what evolution chose to do.

And, as usual, there's a twist: NMDA receptors are different. These are blocked by magnesium, which is a doubly charged positive ion (denoted Mg^{++}), so they're not active unless the voltage at the postsynaptic receptor is relatively high. More quantitatively, the NMDA current is

$$I_{NMDA} = -\frac{g_{NMDA} x_{NMDA} (V - \mathcal{E}_{NMDA})}{1 + ([Mg^{++}]/3.57 \text{ mM}) \exp(-V/16.1 \text{ mV})} \tag{15}$$

where $[\cdot]$ is shorthand for concentration and mM is milli-molar. One way to get high voltage at the postsynaptic receptor is via a backpropagating action potential: if there's a spike at the soma, the voltage propagates through the dendritic tree, and raises the voltage everywhere. This makes NMDA receptors coincidence detectors (they're active when pre and postsynaptic spikes happen at about the same time). This is perfect for learning, and will be discussed below.

1.5 Neurotransmitter release

You should know the sequence that leads to neurotransmitter release from presynaptic terminals (it's what causes $c(t)$ in Eq. (13) to increase). This is kind of complicated,

1. A presynaptic spike travels along an axon and arrives at the presynaptic terminal.
2. That causes, via voltage-gated calcium channels, a local increase in calcium, denoted Ca^{++}, inside the presynaptic terminal.
3. That causes (sometimes; more on that below) one or more vesicles (in the cortex usually one; in other places, more) to fuse to the cell membrane and release neurotransmitter.
4. That in turn causes $c(t)$ to increase (see Eq. (13)), which results in current flowing into, or out of, the postsynaptic terminal.

5. Voltage diffuses along the dendrites, eventually leading to a change in voltage at the soma.

6. Which will eventually lead to a spike, and the whole thing starts over.

It's a miracle this works!

For reasons that are not clear (probably due to biophysical constraints), sometimes no neurotransmitter is released. Lack of release is called a failure, and it's captured by the release probability, p, which is surprisingly low – only around 1/2 (with a big range). And to make matters more complicated, p changes: it tends to drop if there's release (depression), because there are fewer vesicles available, and it tends to increase when a presynaptic spike arrives (facilitation), because calcium builds up. Between presynaptic spikes it relaxes back to its base value, which can drift slowly. You should be able to write down simple differential equations describing these processes.

1.6 Synaptic plasticity

Finally, there's synaptic plasticity: both the release probability and the postsynaptic conductance change in an activity-dependent way; it's what we believe is responsible for learning. There are lots and lots of experiments investigating how weights change. Those experiments generally look at changes as a function of pre and postsynaptic activity. The result: it's complicated, and depends on the synapse. The most common form of plasticity at excitatory synapses (or at least the most studied) is NMDA-dependent plasticity: when NMDA channels open (which, as discussed above, happens when pre and postsynaptic spikes occur at about the same), calcium enters the cell, and can cause the insertion or deletion of AMPA channels.

But this is only one type of plasticity, and of course it doesn't apply when the presynaptic neuron is inhibitory (since the receptor is GABA, not AMPA, mediated). There are lots of other types, and we're just beginning to work them out. And what's seriously missing is the dependence on an error signal, which is necessary to determine whether to increase or decrease the synaptic strength. This is a *huge* and complicated field, and we only touched the surface.

I'll leave you with the model that theorists like most. In this model, the synaptic strength from presynaptic neuron j to postsynaptic neuron i, denoted w_{ij}, changes according to

$$\Delta w_{ij} = \eta \int dt\, dt'\, K(t-t') S_i(t) S_j(t') \tag{16}$$

where η is the learning rate and $S_i(t)$ refers to the spike train of neuron i; using t_i^μ for the time of the μ^{th} spike on neuron i,

$$S_i(t) \equiv \sum_\mu \delta(t - t_i^\mu)\,. \tag{17}$$

Note that the above learning rule can be written

$$\Delta w_{ij} = \eta \sum_{\mu,\nu} K(t_i^\mu - t_i^\nu)\,. \tag{18}$$

Different kernels have been observed. Symmetric ones tend to be called Hebbian, while antisymmetric ones are called STDP (for spike-timing-dependent plasticity). While theorist like the learning rule given in Eq. (16) (mainly because it's relatively simple), the actual learning rule (which varies from one synapse to the next) is a lot more complicated, and typically involve nonlinearities in spike times.

2 Network dynamics

2.1 Randomly connected networks

The writeup on randomly connected networks, which can be found at http://www.gatsby.ucl.ac.uk/~pel/tn/notes/networks.pdf , is pretty complete. But to understand what we've done in class, you just need to keep in mind a few things.

First, you need to know how sums of large numbers of random variables scale. In particular, if

$$\mathbb{E}[x_i] = \mu \tag{19a}$$

$$\text{Covar}[x_i, x_j] = \delta_{ij}\sigma^2\,, \tag{19b}$$

then, in the large n limit,

$$\frac{1}{n}\sum_{i=1}^{n} x_i \sim \mu + \frac{\sigma}{\sqrt{n}}\,\xi_i \tag{20}$$

where ξ_i is a zero mean, unit variance Gaussian random variable,

$$\xi_i \sim \mathcal{N}(0,1)\,, \tag{21}$$

and the "$\sim$" in Eq. (20) means the left and right hand sides have the same distribution.

This means sums of large numbers of random variables self-average: their empirical sum is equal to the actual sum. Or, more accurately, it's equal if you only care about leading order in n. That's extremely useful, because it simplifies lots of calculations. And the next order in n, which is smaller by a factor of $1/\sqrt{n}$, is also pretty simple, because it's Gaussian.

Note that it's critical that the random variables are uncorrelated. If instead they were correlated,

$$\text{Covar}[x_i, x_j] = \Sigma_{ij}\,, \tag{22}$$

things change. The average is still the same

$$\mathbb{E}\left[\frac{1}{n}\sum_i x_i\right] = \mu\,, \tag{23}$$

but this quantity does not, in general, self-average. To see that, compute the variance,

$$\text{Var}\left[\frac{1}{n}\sum_i x_i\right] = \frac{1}{n^2}\sum_{ij}\Sigma_{ij}\,. \tag{24}$$

Unless the off-diagonal terms of the covariance matrix are zero on average (which is not so common), the right hand side is $\mathcal{O}(1)$. Thus, the standard deviation of the average is on the same order as the mean. And it gets worse: we can no longer invoke the central limit theorem, so the distribution of the average isn't even Gaussian (at least not in general). The point is: assuming a random variable is uncorrelated is a very strong assumption, and should be made with care.

To see what this has to do with networks of neurons, I'll consider two examples. The first is in the writeup on networks mentioned above (although this version is somewhat simplified); the second isn't.

In the first example, we'll consider an equation describing the equilibrium firing rates in a network of n neurons,

$$\nu_i = \phi\left(\frac{1}{n^{1/2}}\sum_{j=1}^{n} w_{ij}\nu_j\right) \tag{25}$$

where the weights, w_{ij} are pulled *iid* from some distribution. All we know about that distribution is that the mean is μ and the variance is σ^2. Our goal is to find the distribution of the firing rates. We could use Eq. (20) directly, but we'll instead use a slightly different method, mainly because it's somewhat easier. But it amounts to the same thing.

First, let

$$w_{ij} = \mu + \delta w_{ij}\,, \tag{26}$$

so that the sum over j becomes

$$\frac{1}{n^{1/2}}\sum_{j=1}^{n} w_{ij}\nu_j = n^{1/2}\mu\overline{\nu} + \frac{1}{n^{1/2}}\sum_{j=1}^{n}\delta w_{ij}\nu_j \tag{27}$$

where the k^{th} moment of the firing rate is given by

$$\overline{\nu^k} = \frac{1}{n}\sum_{i=1}^{n}\nu_i^k\,. \tag{28}$$

Second, treat the second term as a Gaussian random variable. It's mean is zero (by design), and so all we need is its variance, with respect to index, i. That's given by

$$\text{Var}\left[\frac{1}{n^{1/2}}\sum_{j=1}^{n}\delta w_{ij}\nu_j\right] = \frac{1}{n}\sum_i\left[\frac{1}{n^{1/2}}\sum_{j=1}^{n}\delta w_{ij}\nu_j\right]^2 = \frac{1}{n}\sum_{jj'}\nu_j\nu_{j'}\frac{1}{n}\sum_i\delta w_{ij}\delta w_{ij'}\,. \tag{29}$$

Because the δw_{ij} are uncorrelated random variable, terms with $j \neq j'$ are zero on average (you should verify this – remember, there are a lot more of those terms than terms with $j = j'$), and so, in the large n limit,

$$\text{Var}\left[\frac{1}{n^{1/2}}\sum_{j=1}^{n}\delta w_{ij}\nu_j\right] = \frac{1}{n}\sum_j\nu_j^2\frac{1}{n}\sum_i\delta w_{ij}^2 = \sigma^2\overline{\nu^2}\,. \tag{30}$$

Our big leap is to treat the second term in Eq. (27) as a zero mean Gaussian random variable with variance $\sigma^2\overline{\nu^2}$. When we do that, Eq. (25) can be written

$$\nu_i = \phi\left(n^{1/2}\mu\overline{\nu} + \left(\sigma^2\overline{\nu^2}\right)^{1/2}\xi_i\right) \tag{31}$$

where ξ_i is a zero mean, unit variance Gaussian random variable. The problem is that we don't know the first and second moments, which are given by

$$\overline{\nu^k} = \frac{1}{n}\sum_{i=1}^{n}\phi^k\left(n^{1/2}\mu\overline{\nu} + \left(\sigma^2\overline{\nu^2}\right)^{1/2}\xi_i\right) \tag{32}$$

for $k = 1, 2$. So it seems like all this analysis doesn't help much; we still have to do a sum. However, the sum depends on the firing rates only through the first two moments, and all the i-dependence is in the Gaussian random variable ξ_i. The latter observation means we can turn the sum into an integral, giving us the mean field equations

$$\overline{\nu^k} = \int d\xi\,\frac{e^{-\xi^2/2}}{(2\pi)^{1/2}}\,\phi^k\left(n^{1/2}\mu\overline{\nu} + \left(\sigma^2\overline{\nu^2}\right)^{1/2}\xi\right)\,. \tag{33}$$

Solving this for for $k = 1, 2$ gives us the first two moments. Once we know those, the distribution of firing rates is found from

$$P(\nu) = \int d\xi\,\frac{e^{-\xi^2/2}}{(2\pi)^{1/2}}\,\delta\left(\nu - \phi\left(n^{1/2}\mu\overline{\nu} + \left(\sigma^2\overline{\nu^2}\right)^{1/2}\xi\right)\right) \tag{34}$$

where $\delta(\cdot)$ is the Dirac delta function.

The second example is the classical Hopfield network, which has the discrete update equation

$$x_i(t+1) = \tanh\left(\frac{1}{n}\sum_j J_{ij}x_j(t)\right)\,. \tag{35}$$

The weights, J_{ij} are given by

$$J_{ij} = \sum_{\mu=1}^{p}\eta_i^\mu\eta_j^\mu \tag{36}$$

where

$$\eta_i^\mu = \begin{cases} +1 & \text{probability } 1/2 \\ -1 & \text{probability } 1/2\,. \end{cases} \tag{37}$$

To solve this we'll define the overlaps, m_μ, via

$$m_\mu(t) \equiv \frac{1}{n}\sum_i \eta_i x_i(t)\,. \tag{38}$$

The update rule for m_μ is given by

$$m_\mu(t+1) = \frac{1}{n}\sum_i \eta_i^\mu \tanh\left(\frac{1}{n}\sum_j \sum_\nu \eta_i^\nu \eta_j^\nu x_j(t)\right) = \frac{1}{n}\sum_i \eta_i^\mu \tanh\left(\sum_\nu \eta_i^\nu m_\nu(t)\right) . \quad (39)$$

If $p < n$ we have fewer equations, but besides that it's not immediately clear how this helps. However, we'll now make a wild guess: m_μ is large and the rest of the overlaps, $m_{\nu\neq\mu}$, are small. This suggest that we treat m_μ and m_ν differently, so we write

$$m_\mu(t+1) = \frac{1}{n}\sum_i \eta_i^\mu \tanh\left(\eta_i^\mu m_\mu(t) + \sum_{\nu\neq\mu} \eta_i^\nu m_\nu(t)\right) . \quad (40)$$

In the previous examples sums were turned into Gaussian integrals. Here the sum over i can also be turned into a statistical average, but now over the distribution of η_i^μ, which is binary. However, to do that we first need to deal with the second term inside the tanh. That term we can treat as a Gaussian random variable. Its mean is zero (because the η_i^ν are zero mean), so we just need its variance, which is given by

$$\mathrm{Var}\left[\eta_i^\mu \sum_{\nu\neq\mu} \eta_i^\nu m_\nu(t)\right] = \frac{1}{n}\sum_i \left(\sum_{\nu\neq\mu} \eta_i^\nu m_\nu(t)\right)^2 . \quad (41)$$

As usual, only the diagonal terms (in ν) survive, giving us

$$\mathrm{Var}\left[\eta_i^\mu \sum_{\nu\neq\mu} \eta_i^\nu m_\nu(t)\right] = \sum_{\nu\neq\mu} m_\nu^2(t) \equiv \sigma_m^2 . \quad (42)$$

Note that σ_m^2 depends on time, a dependence we'll suppress. It's possible to find σ_m^2 self-consistently, but we won't do that. Instead, we'll pretend like it's known, and treat the second term inside Eq. (40) as a random variable,

$$m_\mu(t+1) = \frac{1}{n}\sum_i \eta_i^\mu \tanh\left(\eta_i^\mu m_\mu(t) + \sigma_m \xi_i\right) \quad (43)$$

where ξ_i is a zero mean, unit variance Gaussian random variable. Importantly, ξ_i and η_i^μ are uncorrelated. We can, therefore, write

$$m_\mu(t+1) = \int d\xi \, \frac{e^{-\xi^2/2}}{(2\pi)^{1/2}} \left\langle \eta^\mu \tanh\left(\eta^\mu m_\mu(t) + \sigma_m \xi\right)\right\rangle_{\eta^\mu} \quad (44)$$

where η^μ is ± 1, each with probability 1/2. Because ξ and η are independent, we can perform the average over η^μ, leaving us with

$$m_\mu(t+1) = \int d\xi \, \frac{e^{-\xi^2/2}}{(2\pi)^{1/2}} \tanh\left(m_\mu(t) + \sigma_m \xi\right) . \quad (45)$$

As usual, the quenched noise just smooths the gain function. However, we have to solve for σ_m self-consistently. That's not totally easy, but it's not especially hard either. We won't do that here; really the main point is that sums over indices sometimes turn into Gaussian integrals, but sometimes they turn into averages over discrete variables. It's important to know when to use which!