Forcing a meaninful future: Brain-spine-muscle interfaces

By Jackson Powell
For future Jackson

Once you know the way broadly, you can see it in all things.
– The Book of Five Rings by Miyamoto Musashi

Contents

0.1 Overview

0.1.1 Purpose

The purpose of this book is to aggregate all of the content and tools I will need in forging the future I desire: building better brain-spine-muscle interfaces.

0.1.2 Content

It is true that total recovery from spinal cord injury, in sever cases, will require both therapeutic and electronic intervention. One must therefore learn of physiology and electronics to be able to solve these problems. I will cover, in some depth, both the electronics, math, and physiology I think I'll need. While programming will likely be an essential tool in this endeavor, I will not discuss programming beyond Verilog (which is an aspect of electronics). This is because programming is something you learn by doing, and writing out C++ algorithms in a PDF would be a waste of time!

Part I

Electronics

Like many of the vagabonds who live in the fields, stray horses seemed to him to be good-natured things. When you're through with them, they ask for nothing; they just go off quietly somewhere by themselves.
– Musashi by Eiji Yoshikawa

Chapter 1

Background Information

1.0.1 Overview

The bulk of this information comes from the Lab Electronics course at the University of Pennsylvania, taught to me by Professor Ashmanskas[1,2]. As a comment on notation: V represents a variable value, while V represents a unit (or at least, that's what it should be; sometimes I forget proper notation). Graphics like this were made using CircuitTikz, whose manual can be found here.

1.1 How Electricity Works

1.1.1 The misconception

One often represents electricity with the canonical metaphor of water flowing in a tube, equating the flow of water, which powers a wheel of some sort, as being the equivalent of electrons pushing through a wire. Incidentally, this simplifying schema illustrates a key misconception in the nature of electricity. This is easiest exemplified in the mode with which electricity reaches one's house from a power-plant, which before arriving will be subject to breaks in the circuit (transformers). In the traditional viewing of electricity, that is taught in early education, this is disconcerting as if electrons can not physically go from a power plant to the lights in your home, then how can their kinetic energy be transferred to them and turn them on?

Regarding the flow of *energy*: when a battery sits without wires attached, around it is an electric field. This field does not dissipate because no electrons flow from it. When wires are attached, charge accumulates on the surface of the wires. This causes a small electric field within the wires, but the drift velocity of electrons within the wire is quite slow—nowhere near the speed of light that you might expect electricity to flow. However, the flow of electrons within the wire is sufficient to drive an electric field which exists all around the wires. From this, we can determine the direction that energy will flow by taking the cross product of the electric and magnetic fields. In fact, if an lightbulb is attached, this means that energy flows from the battery to the bulb in all directions, not through the wires itself. This energy flow induces the vibration of electrons within the bulb's filament, thereby causing light. This means that a net flow of electrons is not required to power a bulb—but rather only their vibrations. Thus is illustrated next:

[1] https://www.hep.upenn.edu/Classes/Phys364_spring23/
[2] http://www.hep.upenn.edu/~ashmansk/

Let us consider now what will happen with an AC circuit (120V AC outlets around your home), where the electromotive force flips with each cycle. In this case, both the electric and magnetic fields switch directions, meaning that their cross product will remain the same and, again, energy flows in all directions to power the lightbulb. Notably, the electrons do not move much (if at all) in this setup—but this is not a surprise, as it is not the electrons that carry the energy anyway. Now, it is still essential to recognize that it is the movement of electrons within the filament of a lightbulb that creates light. This is, indeed, from kinetic energy transferred from electrons bouncing against the metals lattice, dissipating energy in the form of light. The necessary distinction is that it is not electrons that flow all the way from the battery, but rather it is vibrations of those that were, and always will be, within the bulb itself. When you consider it like this, it is straightforward—as the electric field derived from the battery is what provides the electrons with enough kinetic energy to power the bulb.

Interestingly, comparisons to the "water flow" model fail dramatically in the traditional sense, but the Venturi Effect used to describe fluid flow actually succeeds. In adding a bulb you add a resistor, which is comparable to adding a part of a pipe with a smaller diameter. As water will flow faster in this section of the tube, so too will electrons. In order to maintain the same current as is through the rest of the tube, the drift velocity, V_d, must be higher. V_d is proportional to the electromotive force, E, meaning the force is highest within the bulb. Things like V_d are simplified into Ohm's Law ($V = IR$) and not often discussed.

Of course you may say: "well, then why do we use wires at all?" The answer is that wires are helpful in channeling the fields, thereby making them more efficient. But, we do not *need* wires, per se. Think of wireless charging, for example. Knowing all of this, in this work I will almost invariably describe the flow of current as electrons moving through a wire. This is because it is much easier to think of electricity in this way, hence the ubiquitous misconception.

1.1.2 Nuances in the fields

It is worth explicitly highlighting that the electric field that causes the actual flow is from charges along the wires, rather than the battery. This is notable because if this were not the case, the proximity of the bulb to the battery would dictate its brightness. One may wonder how this type of charge distribution can be established so rapidly, and the reason is that the distance an electron needs to travel in order to create such a distribution is subatomic in size—meaning that with movement at the speed of light, the time it takes to establish a surface charge is effectively zero.

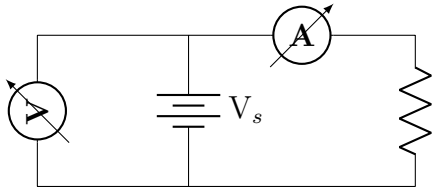
The whole idea is quite unintuitive, so it is appropriate to keep it smushed in the back of your mind, and to only draw it out when encountering things that are otherwise strange—like the aforementioned wireless charging, which should now be much more comprehensible.

1.2 Laws and Devices

The currents flowing in and out of a node will always be equal. The sum of the voltages over an entire circuit will always be equal. Kirchoff's Voltage Law (KVL) can be used to show that wires connected in parallel will have the same voltage across them, and the current flowing in and out of a node will always be the same, defined in Kirchoff's Current Law (KCL). In this way, we can predict the current flowing through a circuit to be $V = IR$. The formula for power is given as: $VI = P$, which means

that when one solves for voltage, and knowing that current is in units of charge/time and power in work/time, voltage is work per unit charge.

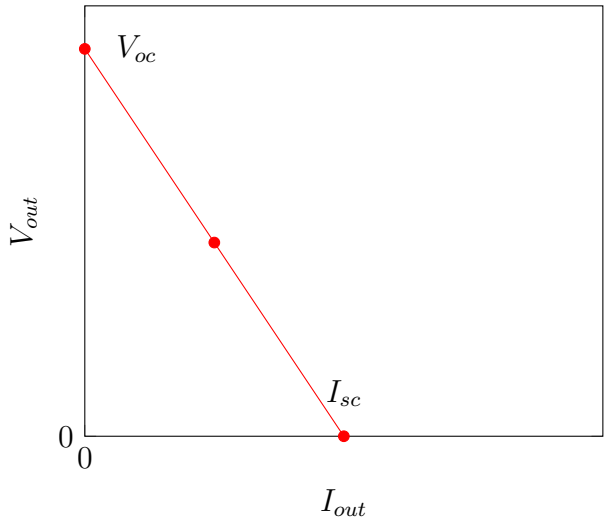
Voltage is measured using a voltmeter, a device in parallel with the load you are interested in measuring. An ammeter is used to measure current, which will be in series with the current you are trying to measure. This means that the voltmeter should have an extremely high resistance, so as to not draw any current, and an ammeter to have a low resistance, so as to not have any voltage drop. These are important considerations, as if the resistance of the load you are measuring is large (say, 1MΩ), it is possible that the voltmeter will have some non-negligible current flow through it. The same goes for if your circuit has very low resistance and you use an ammeter. To illustrate, one would measure the voltage and current coming from a battery as seen below:



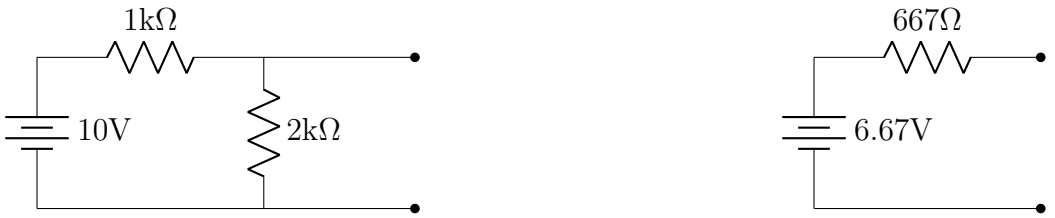
1.2.1 Thevenin and Ideality

A Thevenin Circuit simplifies the circuit to have a single resistance, R_{th}, and a single voltage, V_{th}. R_{th} can be calculated via replacing all of the voltage sources with a wire, and disconnecting all of the current sources. This "short circuits" your circuit and leaves you with only resistors, which can be used to calculate R_{th} using the familiar resistor rules. One can also short circuit the terminals, and determine the current flow, giving us $R_{th} = V_{th}/I_{sc}$.

R_{th} can be measured in a circuit by varying the R_{load} added to a circuit. In this case, you will see the voltage supplied (and corresponding current) change. The slope of this change ($\Delta V/\Delta I$) will equal R_{th}. If you vary the load through the two terminals and measure the voltage across it, you will get a graph that looks something like this:

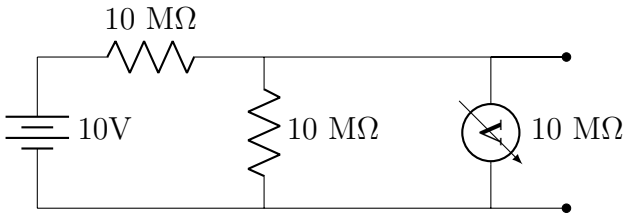


Again, the slope is what gives you R_{th}. I_{sc} is an important value which allows you to calculate V_{th}. The current that flows when you short circuit the load (I_{sc}), multiplied by R_{th}, gives you V_{th}.



The ideality of sources.

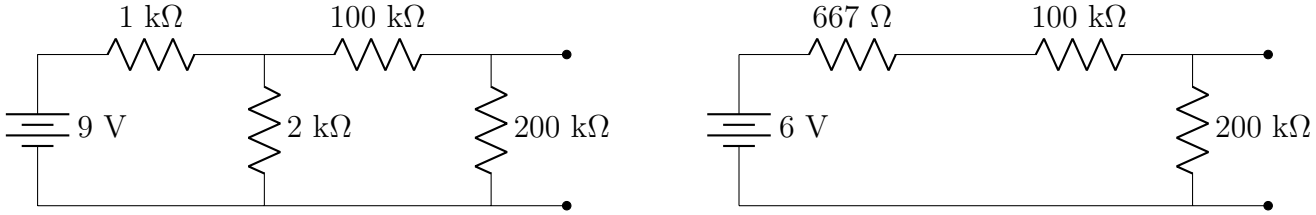
This coaxes us lightly into the topic of source ideality. Imagine all voltage or current sources as having a resistor in parallel with it, but inside of the component itself. An ideal battery, or voltage source, would be able to drive the same voltage, irrespective of the resistor/current. This would be like having a battery whose internal resistance is 0, causing the entirety of the voltage drop to occur on the circuit fragments outside of the battery. In the real world, batteries are not ideal. The canonical illustration of a batteries ideality is in trying to use a 9V battery to start your car. Naturally, the voltage dwindles as the current supplied increases. You can calculate the internal resistance of a battery by adding increasingly large loads to it, thereby giving you the batteries IV curve. It is worth considering this, as if your R_{load} is only $\approx 10 \times R_{th}$, then you may see drooping in the voltage supplied. Another way to state this is to make sure that the *input resistance* of your voltage source is much smaller than the *output resistance* of the upcoming circuit fragment you are attempting to drive. An example of when this fails is as follows:



Because the circuit has a non-negligible resistance relative to the voltmeter, you should expect to read something inconsistent with using an ideal voltmeter. As we are adding a 10 MΩ voltmeter in parallel to our 10 MΩ resistor, we expect that the "R_{load}" in this case will now be 5 MΩ, so the voltage divider at V_{out} will now be $1/3 \times 20$V, or 6.67V.

This contrasts to a current source, whose desired internal resistance is ∞, as you will want no current to flow through it, and to flow entirely through the circuit fragments outside of the component. Once again, as the real world is not ideal, your goal in this case will be to have downstream circuit components whose input resistance is much smaller than the components output resistance.

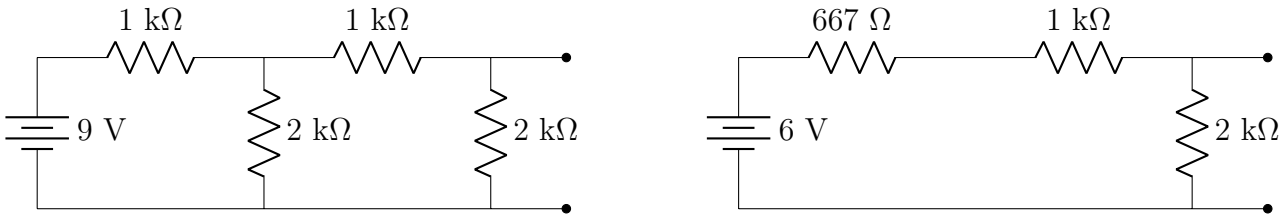
Let us consider another example of input resistance vs. output resistance:



If we draw a black box around the first voltage divider, we can convert it to the circuit schematic on the right because 1 kΩ||2 kΩ = 667 kΩ just as before. We can find the V_{th} by first finding I_{sc}, when

we short the black box's load. So, I_{sc} = 9V/1kΩ = 0.9mA. We multiply I_{sc} (9 mA) by R_{th} (6.67 kΩ) to get a V_{th} of 6 V. If you were to compare the output resistance of the black box, 6.67 kΩ, to the input resistance of the upcoming voltage divider, 300 kΩ, you would find that it is much smaller. This naturally means that the entirety of the 6 V drop will occur over this part of the circuit.

Another way to think about this is as two successive voltage dividers, and qualitatively noting that the second's total resistance is much higher allows us to simplify things greatly. As a reminder, a voltage divider can be solved as $R_A I = R_A \times V_{total}/R_{total} = V_{total} \times R_A/(R_A + R_B)$. If you redraw as the Thevenin equivalent, the first voltage divider can effectively be ignored, because the voltage drop across this component will be minimal. Not by coincidence, since the first voltage divider is a 2/3 divider, and the second is a 2/3 divider, the voltage measured between our two terminals is 2/3 × 2/3 × 9V, or alternatively, 2/3 × 6V (V_{th}). This equivalency does not work when the input and output resistances of each fragment are comparable, here is an example:

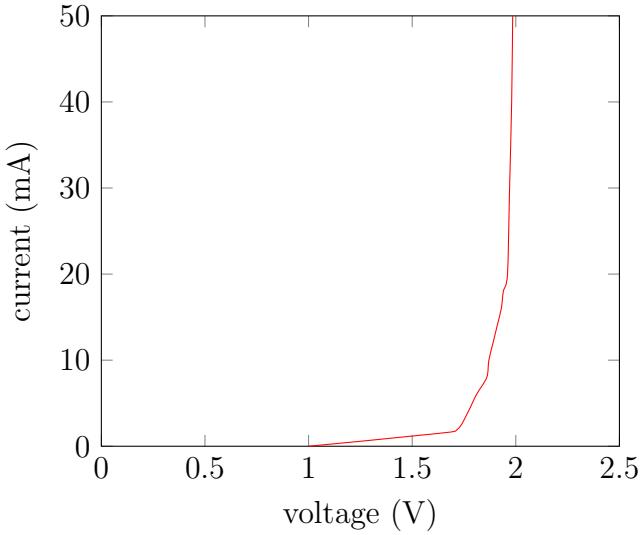


In this case, R_{th} will be the same, and R_{load} will be 3 kΩ. $V_{th} \times R_A/(R_A + R_B) = 3,000/3,667 \approx 5.45$V for the output of the first divider (i.e., between R_{th} and the 1 kΩ resistor). And then naturally, if you were to measure the voltage between the 1 kΩ and 2 kΩ resistors, it would be 2/3 × 5.45V ≈ 3.63V for the voltage at the second divider.

This kind of Thevenin analysis only works when you have a linear IV curve. When might you have a non-linear IV curve?

1.2.2 LED Circuits and PNP

The IV curve across a light emitting diode (LED) should look something like this:

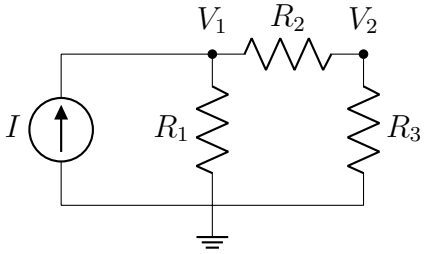


The IV curve for a diode, like an LED, is exponential in that the current slowly increases after the voltage across a diode hits some "threshold," after which the current rises exponentially with voltage. Why is this the case? A diode is a P-N junction bridged by some depletion zone. The P side of the diode contains positively charged elements that act as "holes" (a silly way to say there is an absent electron position). The N side contains elements whose outer layers are loosely filled with electrons (i.e., low ionization energy). Effectively, the P side is devoid of electrons, while the N side has many free to give. What does this mean with regard to current and voltage? It means that the "depletion zone" between the two requires electrons to be able to bridge the gap. This really can't happen unless they have a certain amount of energy, so increasing the voltage helps reach the "threshold" energy requires to pass the depletion zone (think of $P = IV$). Thus, as the electrons somewhat saturate the diode, you can theoretically pass an infinite current through it, as it will be effectively a short circuit.

1.3 Resistor Lattice Digression

One issue with modern BSIs is the usage of ECoGs, or EEGs, or other large measuring devices[3]. Too, they are almost universally hard electronics that require intense surgeries to implant. Therefore, if we could replace with soft electronics, we can cover considerable ground. For example, if one could drill a small hole into a patient's skull and spread over the cortex a fabric that contained electrodes, one could achieve a similar amount of readings with a minimally invasive surgery. It is probable that there will be electrodes small enough to accomplish this, but let's say there aren't. Another way that this could be solved is using a lattice of resistors, with probes at either corner. These corners can exit the brain and be the points at which a computer interfaces with them.

Don't be annoying, just go with the process.



Please, don't be annoying—the solution is trivial but we will be talking about methods you can generalize. This example is from the SPICE method[4]. Considering KCL at the nodes gives us these two equations:

$$\begin{aligned} I &= \frac{V_1}{R_1} + \frac{V_1 - V_2}{R_2} \\ \frac{V_1 - V_2}{R_2} &= \frac{V_2}{R_3} \end{aligned} \tag{1.1}$$

As it goes, we can reformat this so as to easily turn it into a matrix in the following way:

[3]This is expanded on in the later parts.

[4]Thank you Prof. Ashmanskas, http://www.ecircuitcenter.com/SpiceTopics/Overview/Overview.htm

$$
\begin{aligned}
I &= \frac{1}{R_1}V_1 + \frac{1}{R_2}\left(V_1 - V_2\right) \\
I &= \frac{1}{R_1}V_1 + \frac{1}{R_2}V_1 - \frac{1}{R_2}V_2 \\
I &= \left(\frac{1}{R_1} + \frac{1}{R_2}\right)V_1 - \frac{1}{R_2}V_2
\end{aligned}
\tag{1.2}
$$

$$
\begin{aligned}
\frac{V_1 - V_2}{R_2} &= \frac{V_2}{R_3} \\
0 &= -\frac{1}{R_2}V_1 + \frac{1}{R_2}V_2 + \frac{1}{R_3}V_2 \\
0 &= -\frac{1}{R_2}V_1 + \left(\frac{1}{R_2} + \frac{1}{R_3}\right)V_2
\end{aligned}
\tag{1.3}
$$

These sorts of equations will usually be simplified using conductance as below:

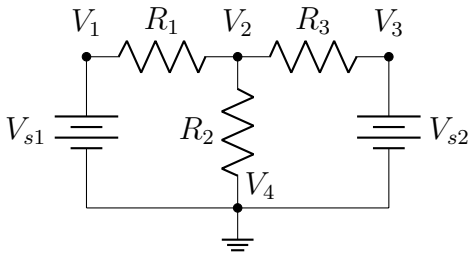
$$
\begin{aligned}
\left(G_1 + G_2\right)V_1 - G_2V_2 &= I \\
-G_2V_1 + \left(G_2 + G_3\right)V_2 &= 0
\end{aligned}
\tag{1.4}
$$

$$
\begin{bmatrix} G_1 + G_2 & -G_2 \\ -G_2 & G_2 + G_3 \end{bmatrix}
\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} =
\begin{bmatrix} I \\ 0 \end{bmatrix}
\tag{1.5}
$$

So if $R_{1,2,3} = 100\Omega$, and $I = 1\text{mA}$, then:

$$
x = \text{A}^{-1}\text{B} = \begin{bmatrix} 67\text{mV} \\ 33\text{mV} \end{bmatrix}
\tag{1.6}
$$

This method is called *nodal analysis*. You'll find more trouble trying to use a voltage source rather than a current source, and in fixing this, a method of *modified nodal analysis* was invented. Let us take the below example[5]:



In this system, we have 4 nodes, but node 4 largely functions as a reference for the other 3 nodes. A small note to make this a tad clearer is that in thinking about the current flowing into a node, such as into V_1, you will view this as current going from V_2 to V_1, but because everything in electronics is

[5] https://cheever.domains.swarthmore.edu/Ref/mna/MNA2.html

backwards, this is calculated as $(V_1 - V_2)/R_1$, and the opposite for current flowing into V_2. Annoyance aside, using KCL, and then our definitions, we can gather these equations:

$$\begin{aligned} I_1 + \frac{V_1 - V_2}{R_1} &= 0 \\ \frac{V_2 - V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_2 - V_3}{R_3} &= 0 \\ I_2 + \frac{V_3 - V_2}{R_3} &= 0 \\ V_1 &= V_{s1} \\ V_3 &= V_{s2} \end{aligned} \tag{1.7}$$

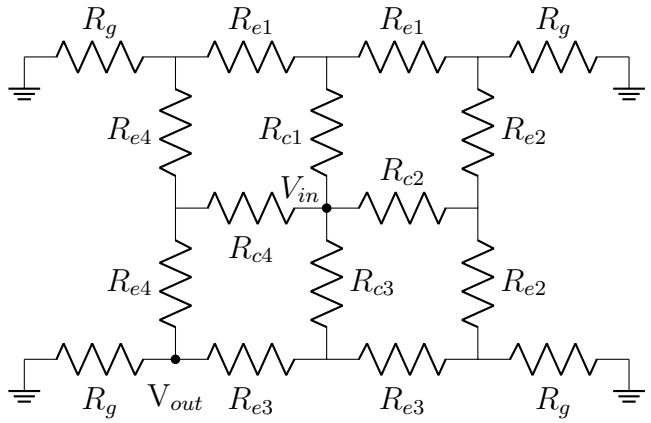
Which we convert to:

$$\begin{aligned} I_1 + G_1 V_1 - G_1 V_2 &= 0 \\ G_1 V_1 + (-G_1 + G_2 + G_3)\, V_2 - G_3 V_3 &= 0 \\ I_3 - G_3 V_2 + G_3 V_3 &= 0 \\ V_1 &= V_{s1} \\ V_3 &= V_{s2} \end{aligned} \tag{1.8}$$

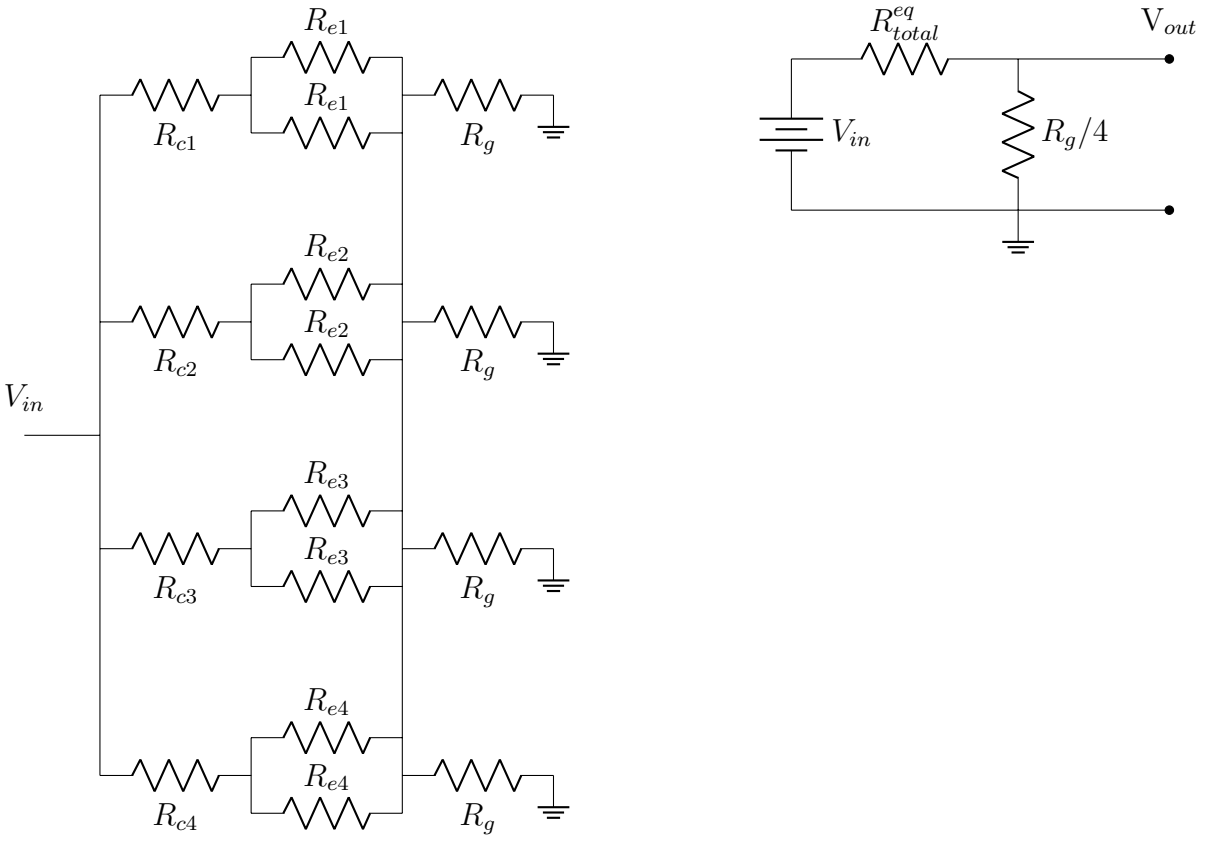
And then:

$$\begin{bmatrix} G_1 & -G_1 & 0 & 1 & 0 \\ -G_1 & G_1 + G_2 + G_3 & -G_3 & 0 & 0 \\ 0 & -G_3 & G_3 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ I_1 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ V_{s1} \\ V_{s2} \end{bmatrix} \tag{1.9}$$

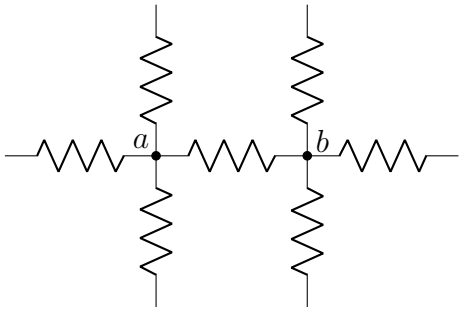
Now, I realize that one doesn't need to build a system of equations for most problems of small size, and they can be generally solved at a glance. Too, in many cases building a system is more hassle than it's worth. But, this will be important to keep in mind for the upcoming chapter **Modeling Circuits**. Let's look at some examples that are easier dealt with in the traditional way:



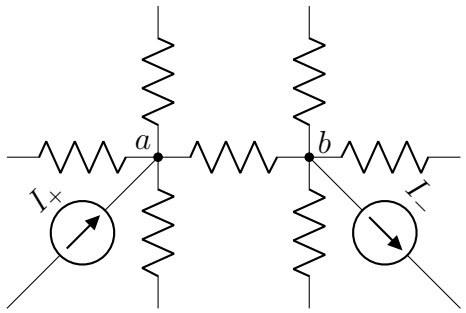
Firstly, apologies for the mess in labeling. The most difficult thing about this problem, in actuality, is how you look at it. Redrawing the circuit helps a great deal. I'll guide you through how one might do this. In truth, the four grounded resistors are in parallel and can be simplified to one wire ($R_g||R_g||R_g||R_g$, or $R_g/4$). You can probably realize that the edge resistors, for example the pair of R_{e1}, are also in parallel and are in series with R_{c1}. Therefore, we have $(R_{e1}||R_{e1}) + R_{c1}$. I'll call this R_1^{eq}. Then you'll realize simply that all of the equivalent R are also in parallel. This gives us $R_1^{eq}||R_2^{eq}||R_3^{eq}||R_4^{eq}$, which I will call R_{total}^{eq}. Now, our circuit simplifies to:



You may have had trouble at first glance because of all the "looped" sections of the circuit. It appears that current can flow any which way. Keep KVL in mind, and recall that current will only flow from high to low voltage. Therefore, you would never have current flowing from, say, R_{e2} to R_{e3}. A popular problem of the same flavor is the *infinite resistor grid*, which looks something like this:

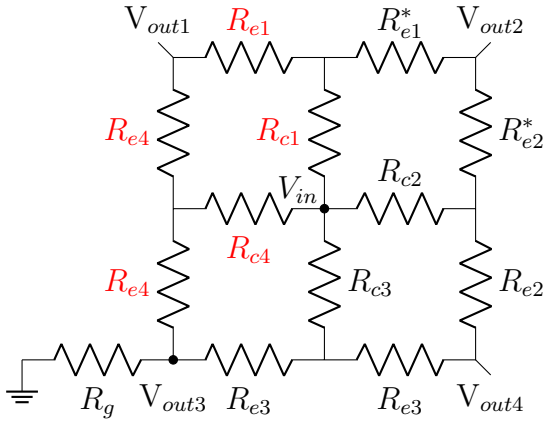


Imagine that the lattice of resistors continues out in every direction infinitely, in a grid-like fashion. The problem asks you to determine the equivalent resistance between points a and b. Considering the infinite nature, the more immediate method to solve would be to take some limit of parallel and series resistors. I actually think this is an important solution to keep in mind, as it can be generalized to other systems better. However, in this case the easiest solution is to use *superposition*. It goes as follows:

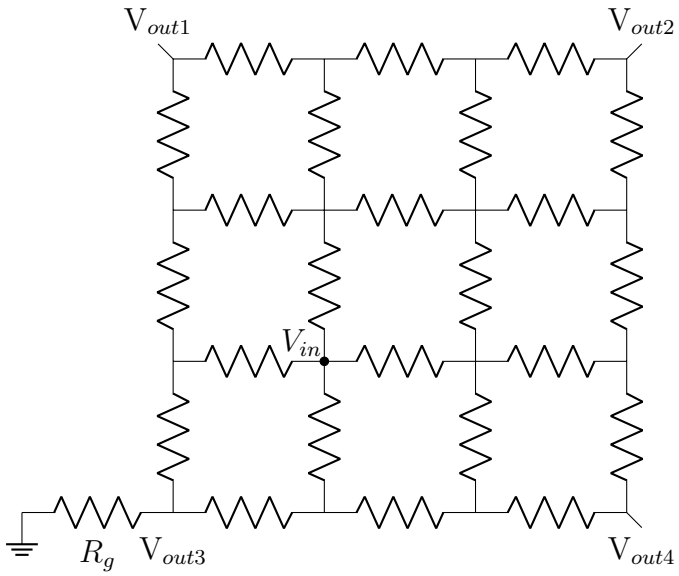


If you were to wire a positive current source to node a, you can immediately know that the current will be split equally in 4 ways, as the lattice is infinite. Then, if you apply a negative current source to node b, you can immediately know that it will draw $1/4I$ from each resistor as well. The total current flowing between nodes a and b will then be $1/2I$. Given that $V = IR$, and by superposition, we can know that $V_{ab} = (I/2)R$, so $R_{ab} = V_{ab}/I = (I/2)R/I = R/2$. The in-line math is kind of ugly, but I think you get the gist.

Another option of a similar flavor is:



In this case, we will think about what each V_{out} will read. One thing to note immediately is that the net current at V_{out2} will be 0. Therefore, R^*_{e1} and R^*_{e2} don't really matter, and can be neglected in our calculations. This once again is easiest to solve when re-visualized. The gist being that you have two quadrants (the R_{c1}, R_{c4}, which I colored to have red labels, and the R_{c2}, R_{c3} quadrants) where V_{in} is split into. The voltage drop across these quadrants being $3R||R$. We then add another R, for the one that connects to R_g, and voila, we are done. Let's call this R_q, so to find the equivalent resistance from V_{in} to V_{out3} it is $R_{q1}||R_{q2}$. I think you get the idea. The reason for going through these simple models is to prepare ourselves for some of the upcoming, more complicated ones. Let us level up now:


V$_{out1}$
V$_{out2}$
V_{in}
R_g
V$_{out3}$
V$_{out4}$

Chapter 2

Capacitors

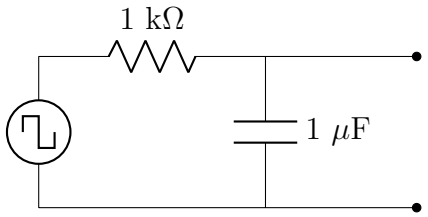
2.1 Overview

Getting a good intuition of capacitor's behavior is an essential part of understanding electronics. A capacitor is like a storage of charge, which can maintain a voltage for a period after the voltage supplied to the circuit is cut off. The basic structure of a capacitor is an anode and a cathode separated by a dialectric plate. Charge will be stored across it, and is released when the voltage supply dwindles.

Capacitor discharge will be exponential decay in the form:

$$\begin{aligned} V_C &= V_S \times e^{-t/RC} \\ \tau &= RC \\ V_C &= \mathrm{V}_S \times e^{-t/\tau} \end{aligned} \tag{2.1}$$

This is quite similar to neurons, isn't it! So some examples, to build up a bit of intuition. Let's say you have the following circuit:



RC will be $1\mathrm{k}\Omega \times 1\mu\mathrm{F} = 1 \times 10^{-3}$, or $\tau = 1\mathrm{ms}$. What does this actually mean, though? It means that if you charge this capacitor up to 10V, then $V_C = V_S \times e^{-t/RC} \rightarrow 10V \times e^{-t/1\mathrm{ms}}$. So 1 ms after voltage is removed, and or, one time constant τ after voltage is removed, the voltage at the capacitor will be $10 \times e^{-1} \approx 10 \times 0.367 = 3.67\mathrm{V}$. If you were using a square wave which charged the capacitor to 10V, with a very high frequency (100 kHz, for example), then the period would be $1/100\mathrm{kHz} = 0.01$ ms. Therefore, you would not expect the capacitor to ever fully discharge, and it would maintain a constant, high voltage. This circuit happens to be what is called a *low-pass filter*, which we will discuss in-depth later. But, you can see how it might get this moniker, as this very high frequency is not allowed to pass due to the capacitor's time constant.

This was meant to serve as a basic intro to what a capacitor is. Now we can begin discussing the nuances.

2.2 Filtration Conceptually

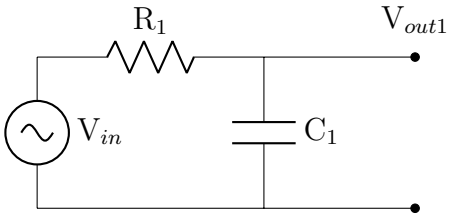
Firstly, what is filtration? Filtration refers to our filtering out of background signal, or any undesigned signal, in order to isolate our signal of interest. For example, perhaps you are interested in a sine wave whose period is 1 ms, but this is superimposed by a bunch of other sine waves of variable periods. You can extract the 1ms sine wave with some clever circuits, and build your following circuit fragments based on it.

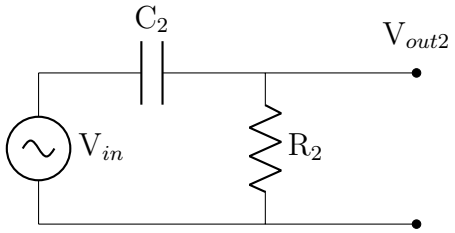
Why filter?

You may ask yourself, why would we need a circuit that can filter out signals / frequencies? An example given by Professor Ashmanskas is: imagine you are building a sensor that tests the water level in a pool. You don't want the pool to overflow when it rains, or get too low when the Summer comes and evaporates water—so you devise an automatic system to add or take out water as needed. One could simply add a sensor to the side of the pool, but it will be subject to the constant waves formed by people swimming, and its readings will be horribly off-base, and uninterpretable. Each time someone cannon-balls in and splashes it, the sensor will think that the pool is greatly overflowing. Thus, one can devise a circuit that filters out all of these little fluctuations (this would be called a *low-pass filter*, for passing things that occur on a long time-scale—low frequency).

2.2.1 Integration and Differentiation

Capacitors have this incredible ability to perform complex math, including taking derivatives or integrals of your wave form. Let's think about how this may occur using the circuit mentioned in the previous part. Firstly, recall that the current flowing through the resistor must equal the current flowing through the capacitor, and that $Q = CV$:





We can then solve for the voltage drop across R_1 as:

$$\begin{aligned} Q &= C_1 V_{out1} \\ \frac{d}{dt}Q &= \frac{d}{dt}C_1 V_{out1} \\ I &= C_1 \frac{dV_{out1}}{dt} \end{aligned} \tag{2.2}$$

$$\begin{aligned} IR_1 &= V_{in} - V_{out1} \\ \frac{dV_{out1}}{dt} &= \frac{1}{R_1 C_1}\left(V_{in} - V_{out1}\right) \end{aligned} \tag{2.3}$$

Therefore, if V_{out} is very small compared to V_{in}, you get:

$$\begin{aligned} \frac{dV_{out1}}{dt} &= \frac{1}{R_1 C_1} V_{in} \\ \int \frac{dV_{out1}}{dt} &= \int \frac{1}{R_1 C_1} V_{in} \\ V_{out1} &= \frac{1}{R_1 C_1} \int V_{in} \end{aligned} \tag{2.4}$$

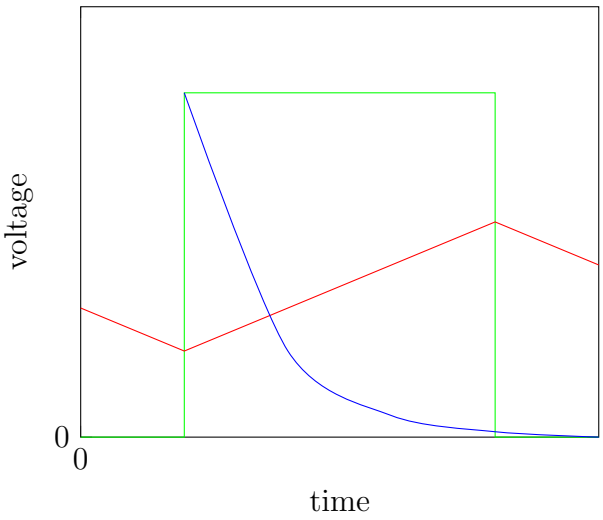
And or, that V_{out} integrates V_{in}. How will this change if we swap the positions of the resistor and capacitor in circuit 2? Once again, consider when V_{out} is much smaller than the input.

$$\begin{aligned} I &= C_2 \frac{d}{dt} (V_{in} - V_{out2}) \\ \frac{V_{out2}}{R_2} &= C_2 \frac{d}{dt} (V_{in} - V_{out2}) \\ \frac{V_{out2}}{R_2} &= C_2 \frac{dV_{in}}{dt} \\ V_{out2} &= R_2 C_2 \frac{dV_{in}}{dt} \end{aligned} \tag{2.5}$$

Thus, in this case V_{out} approximates the derivative of V_{in}.

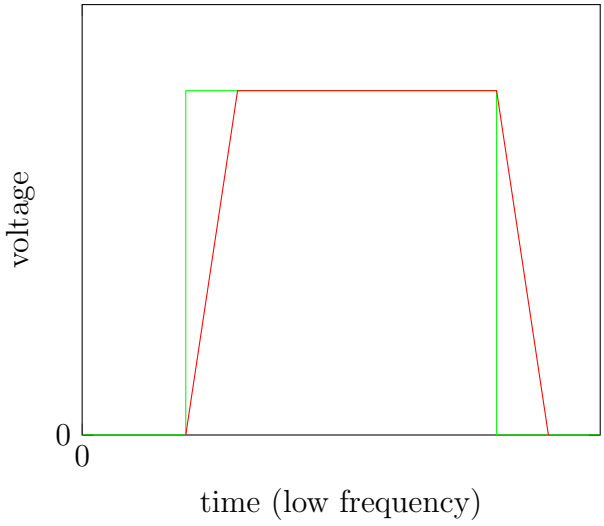
Qualitative thinking.

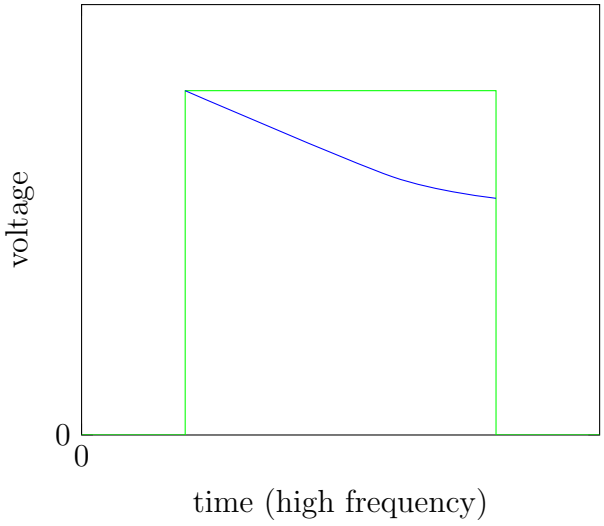
The best lead in to thinking about filtration, to me, is thinking simply about how the graphs look like when something integrates or differentiates a square wave. Let us not do any math, and think only qualitatively.



If the green waveform is our V_{in} from the previous example, then our red curve is similar to how V_{out1} may look, and the blue curve is similar to how V_{out2}. That is, the blue curve *kind of* differentiates the green curve, because the relatively instant rise signifies a very positive derivative, marked by this blue spike. Similarly, the red curve *kind of* integrates the green curve, because the area under the green curve slowly accumulates, hence the constant slope of the red curve.

So if we say that the frequency is extremely slow, and the red curve is the voltage being measured at V_{out2}, how might the derivative curve look?





Similarly, if the frequency is very high, how might the integral look? Take a second to ponder the two graphs above and gather a bit of intuition on it. This should allow you to qualitatively state that an integrating circuit will be a *low-pass filter*, and a differentiating circuit will be a *high-pass filter*. As in the left graph, as the frequency gets smaller and smaller, the red plot will more and more closely match the green plot. This is the basis of capacitor filtration, an essential tool in electronics!

2.3 Frequency Dependence

This section will primarily be quantitative. The important bits, though, are the qualitative understandings, and the final results of this quantitative section. What you find in the interim is likely not worth understanding fully. Let's begin by supposing we apply a cos wave to a resistor. The relationship between current and voltage is as follows:

$$\begin{aligned} f &= \frac{\omega}{2\pi} \\ V(t) &= V_{pp}\cos(\omega t) \\ V(t) &= I(t)\mathrm{R} \\ I(t) &= (V_{pp}/R)\cos(\omega t) \end{aligned} \tag{2.6}$$

Therefore, if you were to solve something like V_{out}/V_{in} your pesky cos terms will cancel. Unfortunately, this is not so for capacitor equations. We can see this below:

$$\begin{aligned} Q &= CV \\ I(t) &= C\frac{d\left(V_{pp}\cos(\omega t)\right)}{dt} \\ I(t) &= -\omega C V_{pp}\sin(\omega t) \end{aligned} \tag{2.7}$$

This means that the current and the voltage will be 90° out of phase. Another way would be to write this as $A\cos(\omega t + \phi)$. Let us think about what will happen with our standard low-pass filter. Using KVL, we can state that:

$$\begin{aligned} V_{pp}\cos(\omega t) &= RC\frac{dv_C}{dt} + v_c \\ v_c &= A\cos(\omega t + \phi) \\ V_{pp}\cos(\omega t) &= RC\frac{d}{dt}A\cos(\omega t + \phi) + A\cos(\omega t + \phi) \end{aligned} \tag{2.8}$$

2.3.1 Imaginary Numbers Digression.

You probably looked at the previous equation and it is solvable (probably?) but that it would be not worth your while to do so, and that there is likely a better way to go about it. Recall Euler's Relation[1]:

$$e^{j\omega t} = \cos(\omega t) + j\sin(\omega t) \tag{2.9}$$

One of the requirements of a linear circuit is superposition, and that is *kind of* the argument that allows us to use imaginary numbers. You can think that if you are using math that includes both a real and imaginary component, as long as you keep track of the real, your output will still be correct—but, don't think too hard about this. In the upcoming sections, I will use $\mathbf{v}_c$ to denote the complex number. Let us examine:

$$\begin{aligned} V_{pp}e^{j\omega t} &= RC\frac{d\mathbf{v}_c}{dt} + \mathbf{v}_c \\ \mathbf{v}_c &= Ae^{j\omega t} \\ V_{pp}e^{j\omega t} &= RC\frac{d}{dt}Ae^{j\omega t} + Ae^{j\omega t} \end{aligned} \tag{2.10}$$

We can differentiate, and then obtain an expression for A, and solve for the voltage at the capacitor as:

$$\begin{aligned} V_{pp}e^{j\omega t} &= j\omega RCAe^{j\omega t} + Ae^{j\omega t} \\ V_{pp} &= j\omega RCA + A \\ V_{pp} &= A(1 + j\omega RC) \\ \frac{V_{pp}}{(1 + j\omega RC)} &= A \\ \mathbf{v}_c &= \frac{V_{pp}}{1 + j\omega RC}e^{j\omega t} \end{aligned} \tag{2.11}$$

We now want to find the real component, which begins by rewriting the expression in its polar form:

[1] Electrical Engineers use j instead of i for imaginary numbers. The claim is that it is easier to keep track of in the math.

$$\begin{aligned}
\mathbf{v}_c &= \left(\frac{1}{\sqrt{1+\omega^2R^2C^2}}e^{j\phi}\right)V_{pp}e^{j\omega t} \\
\phi &= \tan^{-1}(-\omega RC) \\
\mathbf{v}_c &= \frac{1}{\sqrt{1+\omega^2R^2C^2}}V_{pp}e^{j(\omega t+\phi)}
\end{aligned} \tag{2.12}$$

From here we can simply take the real part and be on our way:

$$v_c = \frac{V_{pp}}{\sqrt{1+\omega^2R^2C^2}}\cos(\omega t+\phi) \tag{2.13}$$

This is one way to go about this problem. Another way, which will be discussed in the next section, is using impedance.

2.4 Impedance

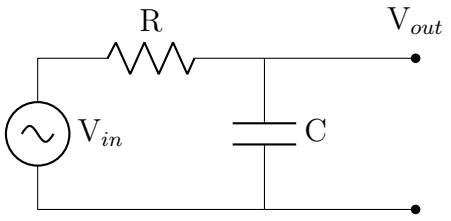
We mentioned earlier in equation (2.11) that the relationship between an input voltage and what is measured at a capacitor for a low-pass filter is:

$$V_{in}\frac{1}{1+j\omega RC} = v_c \tag{2.14}$$

If we were to divide this fraction by $j\omega C$, and simplify using the representation Z_C, we would get:

$$\begin{aligned}
V_{in}\frac{1/j\omega C}{1/j\omega C+R} &= v_c \\
V_{in}\frac{Z_C}{Z_C+R} &= v_c
\end{aligned} \tag{2.15}$$

This looks just like a voltage divider equation! The concept of *impedance* is used to summarize other circuit fragments, like capacitors or inductors, using some resistance equivalent Z. The impedance of a resistor, Z_R, is simply R. Z_C is $1/j\omega C$. Let's think of how this pertains to our low-pass filter.



If we measure at V_{out}, we can think of it like a voltage divider, giving us:

$$V_{out} = \frac{Z_C}{Z_C+R}V_{in} \tag{2.16}$$

You may wonder if $Z_C = 1/j\omega C$ has any real basis, or if it is simply an extraction from the above math. In reality, you can find it quite simply through:

$$\begin{aligned} I_c &= C\frac{dv_c}{dt} \\ I_c e^{j\omega t} &= CA\frac{d}{dt}e^{j\omega t} \\ I_c e^{j\omega t} &= j\omega CAe^{j\omega t} \\ I_c &= j\omega CA \\ I_c\frac{1}{j\omega C} &= A \end{aligned} \tag{2.17}$$

This is the equivalent of Ohm's law, where $1/j\omega C$ is the resistance. Understanding this, we can find our high-pass filter's equation to be:

$$V_{out} = \frac{R}{Z_C + R}V_{in} \tag{2.18}$$

2.5 Filtration Quantitatively

We are finally prepared to talk about filtration quantitatively! As was mentioned in the impedance discussion, resistor-capacitor (RC) circuits can be formulated as voltage dividers. We know that the impedance of a capacitor is $1/j\omega C$, which has some frequency dependence from ω (or, $2\pi f$). Therefore, we can get the sense that V_{out} may change depending on this frequency. With a bit of re-writing, we find:

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= \frac{Z_C}{Z_C + R} \\ \frac{V_{out}}{V_{in}} &= \frac{1}{1 + j\omega RC} \\ \frac{V_{out}}{V_{in}} &= \frac{1}{\sqrt{1 + (2\pi fRC)^2}} \end{aligned} \tag{2.19}$$

We can see that as the frequency goes up, this converges to $1/\infty$, or $1/1$ as frequency goes down. Whereas, for a high pass filter:

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= \frac{R}{Z_C + R} \\ \frac{V_{out}}{V_{in}} &= \frac{j\omega RC}{1 + j\omega RC} \\ \frac{V_{out}}{V_{in}} &= \frac{2\pi fRC}{\sqrt{1 + (2\pi fRC)^2}} \end{aligned} \tag{2.20}$$

Which converges to ∞/∞ as frequency goes up, and $0/1$ as frequency goes down.

2.5.1 Corner Frequency and Phase Shift

The way one quantifies this is with the "corner frequency," which you will more often heart as f_{3dB}. You'll notice that for both the high-pass and low-pass filter equations described above, the ratio of V_{out} to V_{in} is $1/\sqrt{2}$ when the frequency inputted is $f = 1/2\pi RC$. $1/\sqrt{2}$ corresponds to about 0.7, meaning that at this f_{3dB}, the output is about 70% the amplitude of the input. Thus, it is a good marker of how well your filter will work. If you are trying to pick a low pass filter that filters out anything at 1000 kHz and higher, you'll want to pick an RC circuit combination that is well below it.

Notably, filtration causes a phase shift. This shouldn't be surprising when you recall that a low-pass and high-pass filter also integrate and differentiate respectively. Let us take, for example, a sin input. If it is passed through a low-pass filter, would expect it to be integrated to $-\cos$. If passed to a high-pass, it would be differentiated to cos. Therefore, as $-\cos$ is $-90°$ ($\sin(x - \frac{\pi}{2}) = -\cos(x)$) relative to sin, and cos is $+90°$ ($\sin(x + \frac{\pi}{2}) = \cos(x)$), we would expect a low-pass filter to generate an output that lags V_{in}, while a high-pass will generate one which precedes V_{in}.

2.6 Summarizing Thoughts

Let us look at these two circuits again, and see what we can tell just from a glance.

For the left circuit, since the current through the resistor and capacitor must be the same, we can say $Q = CV_{out}$, and thus $I = CV'_{out}$, that $V_{in} = IR = RCV'_{out}$, so V_{out} integrates V_{in}. Knowing that it integrates, we can intuit that it must be a low-pass filter, based on the graphs presented earlier. We can also know it is a low-pass filter if we recall that $Z_C = 1/j\omega C$, and that a voltage divider must look like $Z_C/R + Z_C$, which would give us $1/j\omega C + 1$. And, since it integrates, we know that $\int \sin = -\cos$, so there will be a $-90°$ phase shift at very high frequencies.

For the right circuit, since $V_{in} - V_{out} = V_C$, we know $CV'_C = I$, so $V_{out} = RCV'_{in}$. Therefor, it differentiates, which means it mus be a high-pass filter. Too, it must be of the form $R/Z_C + R$. And if it differentiates, then $\sin' = \cos$, so there must be a $+90°$ phase shift at very low frequencies.

In other words, you can tell a lot just at a glance! This is without any math!

Chapter 3

Writing Hardware

3.1 Introduction

Verilog[1] is a language used to describe electronics, and allows you to avoid the physical action of wiring. This is the reason for the designation **writing hardware**. In this way, you can pick your poison: debugging code, or debugging breadboards. Importantly, though, Verilog is capable of computation and writing data files that go beyond circuit descriptions. Thus, it is not a "markdown" language and is Turing Complete[2].

Tools like Field Programmable Gate Arrays (FPGAs) allow for this, as their internal composition is something of an array of transistors, which can be rewired through code in order to meet the demands of the programmer.

3.1.1 Creating Modules

In Verilog, a circuit is called a `module`. Each module is defined between a `module` and `endmodule`, which can be named as shown in the example below. Different ports connect the module to things outside of the module.

```
module example1(o, i1, i2);
// example1 is the name of our module, and o, i1, and i2 are our ports
// it is convention to list outputs first

output o; // this defines o as an output
input i1, i2; // this defines i1 and i2 as inputs

endmodule
```

Gates are also initialized like modules. The way to do this is with the built in primitives for AND and OR gates (`and` and `or` respectively). For example:

```
module example2(o, i1, i2);

```

[1] A large part of the background information and general syntax comes from the YouTube channel: CompArchIllinois.
[2] Or at least, I think it is. I can never remember the exact definition of Turing Complete :)

```
output o;
input i1, i2;
wire wire1, wire2; // this initializes two wires called wire1 and wire2

or or1(wire1, i1, i2);
// this makes an OR gate named or1 with inputs i1 and i2, and output called
    wire1
and and1(o, wire1, wire2);
// this makes a NOT gate named not1 with input i2, and output called wire2
// one of the outputs of the OR gate feeds into the AND gate (via wire1) in
    this example

endmodule
```

The order in which things are initialized do not matter. It is very important to not reuse wire or other variable names, as Verilog will read these as being connected irrespective of where they are intended to be. As mentioned, the code above uses modules built into Verilog, but you could make your own module in the following way:

```
module andgate(output o1, input i1, input i2);
    assign o1 = i1 & i2;
    // for OR you would use |, and for XOR you would use ^
endmodule;
```

3.1.2 Bus Notation

Bus notation is used to simplify the pins used (in Verilog, this is called a vector). For example, a multiplexer or an adder will have many inputs, which would be inconvenient to initialize individually. Instead, we can use something like this:

```
module adder(c, a, b); // a, b, c are 3 bus inputs we will use
    output[3:0] c; // initializes 4 wires within our c bus
    input[3:0] a, b; // initializes 4 wires within our a and b buses
endmodule
```

Firstly, note that Verilog is 0 indexed, so $[3:0]$ includes 4 wires. In a circuit schematic, busses are drawn as thicker wires with a slash through them and a number denoting the amount of wires in the bus. If we wanted to call individual wires from our busses into the `andgate` module we declared earlier, we could do it as:

```
andgate(c[0], a[0], b[0]);
```

And we can connect busses together, or wires together, using the assign command like before. For example:

```
wire wire3;
assign wire3 = c[2];
assign c[2:0] = a[2:0]; // wire3 will now be connected to bus a[2] through
   bus c[2]
```

3.2 Constants and Variables

Verilog allows us to define constants using 3 parameters, defined as their size, method of encoding, and value. For example, 8'hd7 corresponds to a size of 8 bits, hexadecimal encoding, and the value d7 (equivalently, 11010111). You can use this in `boolean` comparisons, as below:

```
wire wire4;
'define CONST1 3'b011; // CONST1 is the name of the constant
wire4 = (a[2:0] == CONST1);
// This is a tad complicated. The gist is: all of the bit values in a[2:0]
   will be compared to CONST1 in a NXOR style statement. That will then be
   compared to all of the other bits in an AND style statement. So wire4
   will be on only if all wires in correspond to CONST1.
// The C++ equivalent would be something like:
//    (a[2] == 1'b0) && (a[1] == 1'b1) && (a[0] == 1'b1)
// Also, this could be completely wrong. I can't check any of this without
   an actual FPGA in front of me! So who knows!
```

This does bring up a worthwhile point, which is that everything you write in Verilog has a direct circuit component underlying it (I suppose the same is true for any program, but it's more... explicit with Verilog).

Part II

Math and Models

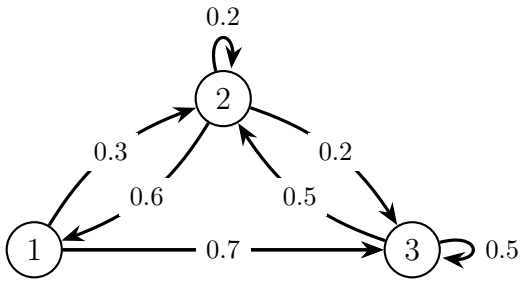
All truly strong people are kind.
– Vagabond by Takehiko Inoue

Chapter 4

Moving Through States

4.1 Markov-chains

Markov chains are useful in predicting the next state desired.



Transition matrix:

$$\mathrm{A} = \begin{bmatrix} 0 & 0.3 & 0.7 \\ 0.6 & 0.2 & 0.2 \\ 0 & 0.5 & 0.5 \end{bmatrix} \tag{4.1}$$

Chapter 5

Neuron Modeling

5.1 Hodgkin-Huxley

5.1.1 The Main Form

The pair won the nobel prize for this model, which formed the basis of our understanding of action potentials. Beyond neurons, it was used in modeling pacemakers of the heart, and muscle cell depolarizations before better models existed. The basis is simply KCL:

$$C\dot{V} = I - I_{Na} - I_K - I_{Leak} \tag{5.1}$$

Because the equations can be found in nearly any textbook[1] or Wikipedia page, I will focus on some of the conceptual understanding I had issues with at first. The complete equation Hodgkin and Huxley arrived at is as follows:

$$C\dot{V} = I - \bar{g}_{Na}m^3h(V - E_{Na}) - \bar{g}_K n^4(V - E_K) - \bar{g}_L(V - E_L) \tag{5.2}$$

There are a few main points to make here. Firstly, this model considers only 3 currents. Na and K are self explanatory, but *Leak* represents the small amount of current that will always occur in cells due to the many routes of charged particles passing through the membrane. It is restorative, in that it pushes the membrane potential back to the resting voltage.

5.1.2 Gating and Conductance

The $\bar{g}$ represent the maximal conductance of these ions. But, shouldn't conductance be variable, depending on how many channels are open? Yes, that is what m, h, and n are for. These three variables are effectively kinetic fits of the opening and closing dynamics of sodium and potassium channels. Again, I will not mention these equations explicitly as they can be found anywhere. Conceptually,

[1] Izhikevich, *Dynamical Systems Neuroscience*

there are three things to know:

Firstly, m is an activation curve for sodium, and the power to the 3rd represents that there are three activation gates. h is an inactivation gate for sodium. Potassium has 4 activation gates, n, and no inactivation gate. Gating can be any number of things, for example, h could be a conformtional change that occurs in the channel after it has been open for $0.1ms$ that closes it again. Naturally, the gating for every channel will be different. Because the *Leak* current is an ensemble of many channel types, it will not have "gating" per se.

Secondly, m, h, and n all are between 0 and 1 and represent the **proportion of channels open**. For instance, if $n = 1$, then 100% of potassium channels will be open. This is why we multiply by the maximal conductance.

Thirdly, m, h, and n are dependent upon voltage, which affords them a time constant τ. This is the conceptually most difficult part. The experimental explanation may be beneficial in understanding. Hodgkin and Huxley realized that these three gating variables will converge to different values depending on the voltage. This makes sense, because we know potassium channels are voltage gated, we would expect the gating variable n to converge to around 1 as the voltage increases. But, the rate at which channels open and close is different. Therefore, their experiments were done to vary the voltage and determine how long it took the conductance of the channels to converge to some value. Does this make sense? In simplest terms: channels open and close at different rates, and that depends on the voltage.

What is the implication of this? Again, look up the exact equations if you are interested. Otherwise, trust the following: m has a time constant τ_m which is very small compared to τ_h and τ_n. Meaning, sodium channels will open the fastest in response to a voltage increase, causing depolarization of the cell. After some delay, sodium channel inactivation (h) and potassium channel activation (n) will kick in, causing repolarization and then hyperpolarization.

These are all derivatives.

One of the most difficult conceptual understandings I had was that $\dot{V}$, m, h, and n are all rates that depend on different time constants, which take voltage as their input. So, the derivative of voltage depends on the derivative of m, h, and n, which depend on voltage. The cyclic nature of this makes it strange, but still doable. Use the general form of derivative, $x_{i+2} = x_{i+1} + (x_{i+1} - x_i)/t$, follow the math, and you will survive.

5.2 Fitzhugh-Nagumo Reduction

5.2.1 Why would we simplify this system?

Reduction implies we are reducing the amount of variables. But why would we do this? The system is already incredibly generalized. We only consider two ion channels and are looking at a static neuron. How can we be accurate if we simplify this system any further?

Let's start by doing a simple thought experiment regarding the previous model:

$$C\dot{V} = I - \bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_K n^4 (V - E_K) - \bar{g}_L (V - E_L)$$

As mentioned, m, h, and n have their own time constants $\tau_{m,h,n}$. That means you'll need to do at least 6 calculations in order to determine $\dot{V}$, which, because it is a derivative, has its own time constant τ_v. Thus, the whole equation is 4^{th} dimensional with respect to time and requires at least 7 or so calculations per time step. If you'd like to simulate an action potential for around 10ms with a time step of 0.01ms, that means you'll perform around 7,000 calculations. Which is not so bad!

However, let's say you want to attempt a propagating action potential. Many people would model this on an infinitely long neuron/wire, but for the sake of this thought experiment let's say you're just interested in a 1 cm neuron/wire for 10 ms. To account for this spatial consideration, you'll need to add in another term besides I which receives current input from the previous segment of the neuron. So, this brings us up to at least 8,000 calculations.

You'd probably want to divide up the neuron into segments on the order of 1 μm. This multiplies our 8,000 calculations by an additional 100,000, giving us 800,000,000 to worry about. Still, this is not horrendous. But, this considers a 1D wire. Neurons are 3D dimensional. We are already considering a system that is 4^{th} dimensional with respect to time, and now we desire to consider 3^{rd} dimensional with respect to space. Imagine trying to calculate the flux through a $1000 \times 1000 \times 1000$ resolution box (i.e., perhaps μm^3 with good resolution). The surface area of this box is thus 6×10^6. Now extend this surface area to include the length of the wire and the area of the soma and dendrites, giving you thousands of millions of points to calculate per time iteration. And, we are still only considering two ion channels. Neurons have dozens and dozens of channels all with different gating kinetics. It does not consider things like lateral inhibition, birufcation, dendritic input, etc. I'll not bother telling you how many calculations we need to perform beyond this point—but it would be large.

5.2.2 How to Reduce

What do we know about the time constants mentioned in the previous section? Roughly speaking, some are fast and some are slow. The upswing of an action potential is on a fast time constant, and the repolarization is on a slow time constant. We also know that the upswing portion is roughly a positive feedback loop, so as voltage increases, so should the derivative of voltage.

This helps us arrive at least at the following:

$$\dot{V} = V \times f(x) \tag{5.3}$$

Simply meaning that the derivative should scale with voltage in some way. We also know that there are at least two "equilibrium points" in a neuron. Meaning, when the neuron is at rest, the $\dot{V}$ will be zero. And, when the neuron reaches the peak of the action potential, the same is true. This will allow us to immediately assume something interesting:

$$\dot{V} = V(V - V_{rest})(V - V_{max}) \tag{5.4}$$

We are already almost there. What we have just done is said that when either $V = V_{rest}$ or $V = V_{max}$, the $\dot{V}$ will not change. These are all on the aforementioned "fast" time scale, and as this is representative of the activation of the action potential, it is effectively a simplification of the sodium channel

dynamics. This is also extremely easy to measure experimentally.

On our second time scale, the slow time scale, we have the inactivation/repolarization function. How will this look like? Just as with the first equation, we will want this curve to increase in magnitude with voltage. Because n represented the potassium channel activation in the previous segment, we can use that as our repolarization function here.

$$\dot{n} = V - \gamma n \tag{5.5}$$

What does this say? It says that our repolarization curve $\dot{n}$ will increase with respect to voltage. But, it will also decrease with respect to itself according to some scaling factor γ.

Now we have reduced our function down to two dimensions and can combine terms:

$$\begin{aligned} \dot{V} &= V(V - V_{rest})(V - V_{max}) - n \\ \dot{n} &= V - \gamma n \end{aligned} \tag{5.6}$$

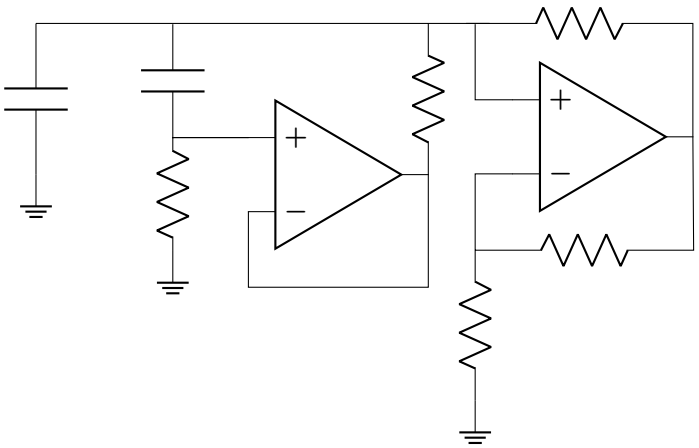
But, we still want voltage to be affected by an injected current, so we can simply add this term back in. And it is also in this equation that we will add our spatial dependence to reach the following:

$$\dot{V} = I_{app} + [V(V - V_{rest})(V - V_{max}) - n] + D\frac{\partial V^2}{\partial x^2} \tag{5.7}$$

D is our spatial dependence, which represents the diffusion of charge around the neuron membrane. And that's it, for now!

Circuits digression.

A lot of the original work done by Fitzhugh and Nagumo used circuit equivalents in order to model neurons. One such example is as follows[2]:



The rightward opamp functions as a Schmitt Trigger, and the entire thing is effectively an opamp oscillator with a second opamp in the middle. The purpose is to simulate an excitable system, like a neuron that is continually firing. Excitable systems systems are those that fire and have some refractory period before firing again (for example forest fires, or even your toilet).

[2]This circuit is adapted from *Mathematical Physiology*, by James Keener & James Sneyd (1998).

5.3 Diffusion

Note that in the coming sections, $[f]$ is used to describe the concentration of some molecule f. But, this principal can be applied to anything, including the spreading of voltage across some surface.

5.3.1 Forward Euler's

Diffusion is accomplished using some diffusion coefficient (D_c) multiplied by some measure of the proportion in one compartment verses another (often $\partial^2[f]/\partial x^2$). D_c can be tuned however desired. The important bit is the second derivative of $[f]$ with respect to space. This can be done using the general form of a second derivative, as written below:

$$f'' = \frac{f_{x+1} - 2f_x + f_{x-1}}{x^2} \tag{5.8}$$

One may wonder how one would solve for an edge case, as the general form of a second derivative requires three data points. There are some nuances, but in general the solution is simply the first derivative of the non-edge side. That is, since the second derivative is the difference in derivatives, that leaves simply the derivative of one side minus zero. This is like applying a closed end to your surface. You can ponder how to solve for an open end, if that ever arises.

This method has some slight issues in which the $[f]$ can occasionally go negative. The way in which this occurs is stated below (note that now $[f]$ is used instead of f to signify concentration at a value x and time t). We can first describe a simplified version of the problem:

$$\begin{aligned}
[f]_i &= f_0 \exp(-mt) \\
\frac{d[f]}{dt} &= \frac{[f]_{i+1} - [f]_i}{\Delta t} = -m[f]_i \\
[f]_{i+1} &= (1 - m\Delta t)[f]_i = [f]_i - m\Delta t[f]_i \\
[f]_i &= (1 - m\Delta t)[f]_{i-1} = (1 - m\Delta t)^2[f]_{i-2} \\
[f]_i &= (1 - m\Delta t)^i[f]_0
\end{aligned} \tag{5.9}$$

You can see easily, from this, that if Δt is too big, you will abandon the characteristic decay you'd expect from $f_0 \exp(-mt)$, and instead get some diverging oscillatory function. How this applies to our interest in diffusion is described below:

$$\begin{aligned}
\frac{[f]_x^{t+1} - [f]_x^t}{\Delta t} &= \frac{[f]_{x+1}^t - 2[f]_x^t + [f]_{x-1}^t}{x^2} \\
[f]_x^{t+1} - [f]_x^t &= \frac{\Delta t}{x^2}\left([f]_{x+1}^t - 2[f]_x^t + [f]_{x-1}^t\right) \\
[f]_x^{t+1} &= \frac{\Delta t}{x^2}[f]_{x+1}^t + \left(1 - 2\frac{\Delta t}{x^2}\right)[f]_x^t + \frac{\Delta t}{x^2}[f]_{x-1}^t
\end{aligned} \tag{5.10}$$

Therefore, if we want to ensure that the concentration is always positive, we are constrained by:

$$\begin{aligned} 1 - 2\frac{\Delta t}{x^2} &= 0 \\ \Delta t &< \frac{x^2}{2} \end{aligned} \tag{5.11}$$

The relevance of this being that if one were interested in modeling on a very small Δx, then one would have to use a Δt that is not physiological, and thus waste a great deal of computing power in doing so. This can be avoided explicitly using some other methods, discussed next.

5.3.2 Backward Euler's

This form serves to solve the time-scale dilemma by swapping $[f]_{x+1}$, and can be used with any Δt. Let us consider the same example from above:

$$\begin{aligned} [f]_i &= f_0 \exp(-mt) \\ \frac{d[f]}{dt} &= \frac{[f]_{i+1} - [f]_i}{\Delta t} = -m[f]_{i+1} \\ [f]_{i+1} - [f]_i &= -m\Delta t[f]_{i+1} \\ [f]_{i+1} &= \frac{1}{1 + m\Delta t}[f]_i \end{aligned} \tag{5.12}$$

Naturally, there is no longer a concern of the size of Δt. Though, one immediate concern is the difficulty of solving your equation for $[f]_{i+1}$. Getting back to the diffusion interest, we now have:

$$\begin{aligned} \frac{[f]_x^{t+1} - [f]_x^t}{\Delta t} &= \frac{[f]_{x+1}^{t+1} - 2[f]_x^{t+1} + [f]_{x-1}^{t+1}}{x^2} \\ [f]_x^{t+1} - [f]_x^t &= \frac{\Delta t}{x^2}\left([f]_{x+1}^{t+1} - 2[f]_x^t + [f]_{x-1}^{t-1}\right) \end{aligned} \tag{5.13}$$

Which simply replaces the previous $[f]_x^t$ with $[f]_x^{t+1}$. This leaves us with three unknowns (those being $[f]^{t+1}$ at $x-1, x, x+1$). We must use linear algebra to solve by first rewriting the left and right side as vectors and a matrix in the following way:

$$[\vec{f}]^{t+1} = \begin{bmatrix} [f]_0^{t+1} \\ [f]_1^{t+1} \\ \vdots \\ [f]_x^{t+1} \end{bmatrix}; [\vec{f}]^t = \begin{bmatrix} [f]_0^t \\ [f]_1^t \\ \vdots \\ [f]_x^t \end{bmatrix} \tag{5.14}$$

and

$$\mathrm{A} = \frac{\Delta t}{x^2}\begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & -2 & 1 & & \vdots \\ \vdots & 1 & -2 & 1 & \vdots \\ \vdots & & 1 & -2 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{5.15}$$

The corners $1, 1$ and x, x were intentionally omitted, as what one desires to do with this is dependent on how they would prefer to treat their edges. As described before in the *Forward Euler's* method, one can use instead of the $(1, -2, 1)$ pattern, simply $(\varnothing, -1, 1)$ pattern, which signifies a closed edge. Therefore, together now we get:

$$\begin{aligned} [\vec{f}]^{t+1} - [\vec{f}]^{t} &= \mathrm{A}[\vec{f}]^{t+1} \\ \mathrm{I}[\vec{f}]^{t+1} - \mathrm{A}[\vec{f}]^{t+1} &= [\vec{f}]^{t} \\ [\vec{f}]^{t+1} &= (\mathrm{I} - \mathrm{A})^{-1}[\vec{f}]^{t} \end{aligned} \tag{5.16}$$

Where I is the identity matrix.

Remarks.

Foward Euler's is explicit, and will be preferred whenever the differential equations are non-stiff[3]. It is the more accurate of the two methods, and can be less computationally intensive if your decay rates are all slow. Backward Euler's is implicit, and can not be used to solve everything, but is doable in most cases. The extra computation required to solve the system of equations more than makes up for potential limitations in your Δt.

[3]On approximately the same time scale.

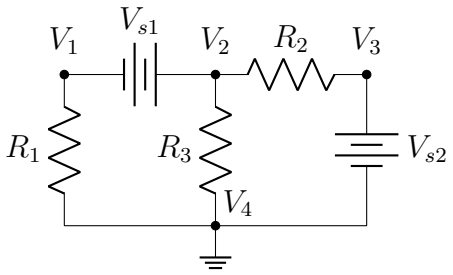
Chapter 6

Modeling Circuits

The idea of this algorithm comes from[1] and is an expansion of what was discussed earlier in the **Electronics** part of this book. Not by coincidence, both of the examples in the earlier section had symmetric matrices. The general way to extract equations from a circuit is in the following pattern:

$$\begin{bmatrix} \mathrm{G} & \mathrm{V} \\ \mathrm{V}^T & \mathrm{Z} \end{bmatrix} \tag{6.1}$$

G is essentially the way nodes are connected by resistors (represented as conductances). The diagonal represents the total number of resistors connected to a diagonal, and the other bits represents how nodes are connected by resistors. I will use the example provided in the footnote directly to illustrate:



There are 4 nodes, but as mentioned before, Node 4 is grounded, a reference, and thus not included in our calculations. This will mean our matrix G must be $n \times n$, where n is $N-1$. Node 1 is connected only to R_1, resolving the $(0,0)$[2] position of the matrix to be G_1. Node 2 is connected to R_2 and R_3, giving us $(1,1) = G_2 + G_3$. Node 3 to only R_3, giving us $(2,2) = G_3$. Node 2 and Node 3 are connected via R_2, meaning the $1,2$ and $2,1$ positions will have a G_2, though notably, it will be $-G_2$:

$$\mathrm{G} = \begin{bmatrix} G_1 & 0 & 0 \\ 0 & G_2 + G_3 & -G_2 \\ 0 & -G_2 & G_2 \end{bmatrix} \tag{6.2}$$

There are two voltage sources ($V = 2$), and 3 nodes, meaning the matrix must be $n \times m$, where $n = N - 1$ and $m = V$. You fill the matrix as if the positive terminal of the j^{th} voltage source is connected to the i^{th} node, point $(i,j) = 1$, or $(i,j) = -1$ for the negative terminal. So for the above, since the negative terminal of V_{s1} is connected to Node 1, $(0,0) = -1$, and as the positive terminal connects to Node 2, $(1,0) = 1$. Lastly, as Node 3 connects to the positive terminal of V_{s1}, $(2,1) = 1$.

[1] https://lpsa.swarthmore.edu/Systems/Electrical/mna/MNA3.html

[2] You are a computer scientist in addition to a mathematician, so of course our matricies must be 0 indexed.

$$\mathrm{V} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.3}$$

$$\mathrm{V}^T = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.4}$$

And lastly, Z for zero is an $n \times n$ matrix of zeros where $n = V$:

$$\mathrm{V}^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{6.5}$$

Giving us an $\mathrm{A}x = \mathrm{B}$ of:

$$\begin{bmatrix} G_1 & 0 & 0 & -1 & 0 \\ 0 & G_2 + G_3 & -G_2 & 1 & 0 \\ 0 & -G_2 & G_2 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ I_1 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ V_{s1} \\ V_{s2} \end{bmatrix} \tag{6.6}$$

You may, rightfully, say "*Uh, who cares?*," since we arleady know that each row corresponds to a different equation. Well, there are a few reasons. Firstly, doing this algorithmically allows us to avoid accidentally underdetermining our matrix using the ol' eye-balling it technique. More importantly, though, the algorithmic approach allows us to solve this via code. Once I can write a bit more neatly, I'll likely upload some snippets here.

Part III

Physiology

Preoccupied with a single leaf, you won't see the tree. Preoccupied with a single tree, you'll miss the entire forest.
– Vagabond by Takehiko Inoue

6.1 Perspective

It is very important to avoid taking a quantized approach to studying physiology. You'll notice neuroscience majors may know some superb nuances regarding the role of HCN channels in overexcitable neural disorders, like epilepsy. However, they may be totally unaware of their canonical role as pacemakers in the heart. So too may they be unaware of the rich history of modeling I_h currents by computational biologists, preferring to look only at a channel as it functions in neurons and is testable in a cell culture. Hence, one must not become preoccupied with a single leaf.

Another bit of wisdom comes from Yoshikawa's Musashi. There is a chapter in which vagabonds are tilling a marsh's untenable land. Each time a storm came, all work was reset as flooding returned the land to mud. One realized that they had been tilling the field in the shape of a square, unconsciously aiming for a uniform, symmetrical crop that is typical of any farm. In doing so, they intended to bend nature from its gruff state into something it was not. Nature, not allowing this, cyclically reclaimed the land. But, if they tilled the land in an oblong fashion, asymmetrical and with the natural bends of the marsh, along the path that water would flow in a storm—perhaps the fields would survive the storm. Rather than trying to oppose the universe, one must first know its way, and be guided by it.

Surgical intervention is most often an opposition to the way of nature. Indeed, the most clear evidence of this is in today's BSIs which attempt to escape endogenous circuits all together. Perhaps one should focus more on enhancing nature's way first, before opposing it. And to enhance nature's way, one must understand it fully.

Chapter 7

Muscles

A large part of this information comes from Dee Silverthorn's *Human Physiology* textbook—arguably the best textbook of all time.

7.1 Skeletal Muscle

7.1.1 A Cellular Level

Structure.

Skeletal muscles are composed of muscle fibers, which are composed of myofibrils. Muscles use a silly nomenclature, in which "sarco" is added to words. For example, the whole structure sits within the sarcoplasm (cytoplasm), fibers are intertwined with the the sarcoplasmic reticulum (equivalent to endoplasmic reticulum) and surrounded by the sarcolemma (equivalent to the cell membrane). An individual myofibril is made up of overlapping actin and myosin segments, held together, to some degree, by titin. The myosin heads are the canonical structure you imagine, which bud off of the end of a chain like leaves on a branch and bind to actin. Undeniably, the most uninteresting part of muscle physiology is as follows: The "centerline" of the myosin networks is called the M line, while the "centerline" of the actin network is called a Z disk. Titin branches off from the Z disks to bind to the ends of myosin chains, providing both some elasticity and support for myosin. Actin chains are centered around a line of nebulin, which too provides structural support and organization. The I band is considered to be the unbound part of the actin structure, centered around the Z disks. The H zone is the unbound part of the myosin structure, centered around the M line. The A band is the entire length of the myosin fibers, thereby encapsulating both the bound part of the actin-myosin complex and the H zone, meaning it too is centered around the M line. Therefore, when contraction occurs, the size of the H zone and I band decreases, while the size of the A band stays the same.

Contraction-relaxation.

Myosin heads desire to bind to actin, but are blocked by tropomyosin. Ca^{2+} can bind to troponin, bound to tropomyosin, to cause conformational changes resulting in the exposure of the actin to the myosin heads. In doing so, binding can occur, followed by myosin's power stroke. The energy for the stroke comes from hydrolyzing ATP, which had already occured by the time myosin attached to the actin. The powerstroke allows the release of the ADP and P_i. ATP then can bind to the empty active site of myosin, which causes the release of the head and prepares it for another cycle. Notably, when ATP does not bind to myosin, the muscle will be stuck in the rigor state.

In skeletal muscles, the source of Ca^{2+} is a combination of the extracellular Ca^{2+} flowing inward, and further release from the sarcoplasmic reticulum. The story goes as follows: a motor neuron releases acetylcholine onto the motor end plate (an area which a high density of sodium channels). This causes a depolarization, which propagates down the muscle fiber. Structures called T-tubulues sink lower into the tissue, allowing for more direct access to the inner processes (visually, these look similar to gyri in the brain or the crypts of the intestinal wall). The T-tubules are lined with dihyropyridine (DHP) channels, an L-type VGCC (specifically $Ca_v1.1$). DHP and ryanodine receptors (RyR) can be mechanically coupled, which influx of Ca^{2+} in through DHP mechanically opens RyR channels of the sarcoplasmic reticulum (a large store of Ca^{2+}). The free calcium is lowered through things like active pumping out of the sarcoplasm.

A steady supply of ATP is needed to maintain pumping, and it is said that at any given time, there is 8 or so twitches worth of ATP within the muscle fiber. Therefore, frequent production and alternate stores are required for continuous movement. One such storage is phosphocreatine, whose phosphate group can be quickly transferred to ADP through creatine kinase. Muscles therefore contain high levels of this enzyme, and **testing for it in the bloodstream can be a good proxy for muscle damage**.

Notably, it is very difficult to fully deplete a muscle of its ATP. Other forms of fatigue begin before this can possibly occur, which include CNS or PNS feedback. An example of this may be that acetylcholine is not synthesized fast enough to continually stimulate muscle fibers. Continual simulation of muscle fibers, beyond what is allowable under normal conditions, **will fully deplete ATP levels and therefore cause damage to muscles**. Another consideration is the continuous use of ATP may result in P_i buildup in the sarcoplasm, making release of ADP + P_i from myosin less likely to occur. Too, this opens the possibility of calcium phosphate forming, which can be quite damaging if it crystallizes further. There are also ion concentration changes to consider, and continued stimulation can result it tetanus.

Disorders Digression.

There are many ways one can lose control of their muscles. Nerve damage, for example, will halt the release of acetylcholine onto the motor plate. In a similar manner, botulism is a result of the botulinum toxin blocking release of acetylcholine, resembling the effect of nerve damage. In muscular dystrophy, the protein dystrophin is absent, which would normally attach actin to the cell membrane. Evidently, this can result in membrane permeability, calcium influx, and thus activation of digestive enzymes that breakdown muscle fibers. In McArdle's disease, muscles simply cannot convert glycogen to glucose-6-phosphase, causing the energy supply to be limited.

7.2 Smooth Muscle

Smooth muscle is much more variable than skeletal, differing by location, contraction pattern, required inputs, and structure.

7.2.1 A Cellular Level

Structure.

Smooth muscle is not considered to have sarcomeres, despite it having the same basic structural components of skeletal muscles. Smooth muscle contains much more aactin than does skeletal muscle, and notably does not contain troponin like skeletal muscle does. Smooth muscle networks are connected through intermediate filaments, which usually attach to dense bodies within the cytoplasm. Actin also attaches to dense bodies, maintaining the actin-myosin network within the cell as well. Smooth muscles do not have T-tubules like skeletal muscles. A comparable structure may be calveolae, which do indent into the membrane and seem spatially associated with the sarcoplasmic reticulum. The autonomic nervous system stimulates fibers through neurotransmitter release from varicosities, or bulbous stores of the chemical. The neurons may innervate the muscle fibers, allowing multiple muscle cells to be stimulated at once, or they may be release to a few fibers, which are connected through gap junctions and stimulate the nearby ones, causing a propagating wave to stimulate others. The first case describes a multi-subunit muscle, and the second a single subunit.

Contraction-Relaxation.

Initiation of contraction begins the same as skeletal muscle, in that calcium enters and the concentration is further driven up by sarcoplasmic calcium release. Though, in this case Ca^{2+} may enter either through gap junctions or membrane ion channels. As such, there are many more modes of entry than in skeletal muscle. For example, voltage-gated Ca^{2+} channels may open, but there are also ligand gated channels or stretch-activated channels, adding extra layers of possible regulation. Intracellular differences arise beginning from release from the SR. Firstly, is is no longer a mechanically gated RyR channel which allows its release. The release mechanism is now Ca^{2+} activated RyR release (commonly called Ca^{2+}-induced-Ca^{2+}-release (CICR)), and the IP_3 path. GPCRs activate phospholipase C, driving IP_3 production which binds to SR channels and causes them to open. The IP_3 path is usually considered the greatest way to drive up intracellular Ca^{2+} (or at least, that is what computational biologists seem to think). When Ca^{2+} is available, it binds to calmodulin (CaM), which then binds to the myosin light chain kinase (MLCK). MLCK phosphorylates myosin to increase myosin ATPase activity. Importantly, once contraction occurs, it stays stiff until released by a different mechanism. Because after contraction, no work is being done in the stiff state, smooth muscle is able to stay contracted for long periods. This explains why sphincters in the body are able to stay closed all the time, while one's bicep fatigues after carrying groceries for just a little while. Relaxation begins when Ca^{2+} is either pumped out of the cell through a Ca^{2+}ATPase pump, or sodium transporter. This causes CaM to unbind, myosin light chain phosphatase (MLCP) to dephosphorylate myosin, and the myosin heads to release from actin. Interestingly, diacylglycerol (DAG), another product of the IP_3 path, inhibits MLCP and thereby enhances muscle contraction.

The calcium stored in the SR is maintained in a number of ways. One example being the protein STIM1 responding to lower Ca^{2+} levels within it, moving toward the cell membrane, and activating store operated Ca^{2+} channels, such as Orai1.

7.3 Cardiac Muscle

7.3.1 A Cellular Level

Structure.

Contraction-Relaxation.

7.3.2 Nervous system control

7.3.3 Pacemakers

Chapter 8

The Spinal Cord

8.1 General Structure

A cross section of the spinal cord would reveal meninges just like the brain and skull; an outer layer of dura, then arachnoid, and finally pia mater. Too, there is an outer layer of white matter followed by a grey matter interior, which is centered around a "central canal." The spinal cord has 33 vertebrae, beginning with cervical (C_1 to C_7), then thoracic (T_1 to T_{12}), then lumbar (L_1 to L_5), then sacral (S_1 to S_5), and finally one coccyx (Co_1 to Co_4). The sacral, and especially the coccygeal, vertebrae are fused, so you may see them depicted as a single unit. Adding slightly to the confusion is that there are only 31 pairs of spinal nerves. These are C_1 - C_8, T_1 - T_{12}, L_1 - L_5, S_1 - S_5, and Co_1. **Note: The difference in the nerve location and vertebrae labels is actually quite important, as it is not always that a fracture at some vertebrae leads to nerve damage at the same site**—this is expanded on in later chapters. The spinal nerves exit the spinal cord on either side, i.e., either dorsal or ventral roots. Dorsal entry neurons carry sensory information to the CNS, while ventral exit zones carry information from the CNS to the muscles. Though, importantly, their axons can not necessarily be found on either side, this will depend on where crossing over occurs. The secton of grey matter which connects to the dorsal root is called the dorsal horn, and the same is true for the ventral horn. The lateral horn is in between the dorsal and ventral horns.

Note to self: It might be fun to add a more detailed breakdown at some point, such as the general arrangement of all the nuclei in the spinal cord.

Another note to self: It would be good to integrate more clinical outcomes into this. For example, what happens in the case of SCI at various vertebrae, or in vagotomy, etc.

8.2 Autonomic Nervous System

Pathways of the autonomic nervous system require two neurons, one that originates in the CNS and terminates at a ganglion, and a post-ganglionic neuron which terminates at the tissue of interest. While the circuit is often considered to have only two neurons, in fact many pre-ganglionic neurons synapse onto many post-ganglionic neurons, which means one neuron can affect many target tissues. Adding to this is the mode of release. The connection between neuron and target tissue is called the *neuroeffector junction*. At this junction, transmitters are secrete indirectly from bulbous varicosities into the interstitial space, meaning neurotransmitters can diffuse over a larger area. Fascinatingly, these neurotransmitters are actually often synthesized within the axon / varicosities of neurons.

8.2.1 Sympathetic Nervous System

Pre-ganglionic neurons of the sympathetic nervous system originate in the hypothalamus or reticular formation and extend down to the mid-sections of the spinal cord, from neuron pairs T_1 to L_2. Sympathetic ganglia are directly beside the spinal cord in a long chain (often called the sympathetic chain). Generally speaking, neurons closer to the T_1 section correspond to organs higher in the body, such as the heart or lungs, while neurons closer to the L_2 section corresponds to those lower in the body, like one's reproductive organs[1]. Pre-ganglionic neurons release acetylcholine onto nicotinic receptors, and post-ganglionic neurons release norepinephrine onto adrenergic receptors.

Norepinepherine is a tyrosine derivative and, like other catecholamines, can be broken down either by monoamine oxidase (MOA) in the neuron's mitochondria, requiring re-uptake, or in the liver by catechol-o-methyltransferase (COMT).

Adrenal Medulla Digression.

It's worth noting that the adrenal medulla secretes epinephrine in a hormone-like manner. However, the adrenal medulla is often considered a collection of pre-ganglionic neurons, meaning epinepherine acts globally like a hormone, but is also a neurotransmitter.

Let us look at the different adrenergic receptors[2]:

Adrenergic Receptors			
Receptor	Location	Sensitivity	Effect
α_1	Most tissues	NE > E	Increase Ca^{2+}
α_2	Gut and pancreas	NE > E	Decrease cAMP
β_1	Heart and kidney	NE = E	Increase cAMP
β_2	Select smooth muscle	NE < E	Increase cAMP
β_3	Adipose	NE > E	Increase cAMP

8.2.2 Paraympathetic Nervous System

Parasympathetic neurons originate from the uppermost and lowermost neurons, either leaving directly from the cranial nerves or from the S_2 to S_4 nerves. Ganglia in the parasympathetic nervous system are close to the target tissue, meaning post-ganglionic neurons are very short compared to the pre-ganglionic ones. One of the most paths being from the vagus nerve (a cranial nerve) that carries the majority of parasympathetic signals to organs including the heart, lungs, liver, stomach, intestines, and pancreas. Like the sympathetic nervous system, pre-ganglionic neurons release acetylcholine onto nicotinic receptors. Notably, parasympathetic post-ganglionic neurons also secrete acetylcholine, but onto muscarinic receptors instead.

[1] Memorizing which nerves respond to which segments is likely not a good use of your time at this stage..., although it might have some clinical use to know which organs may be affected after SCI.

[2] This is ripped directly from Silverthorn.

Cholingergic Receptors		
Receptor	Location	Effect
N_N	Postganglionic neurons	Opens cation channels
N_M	Skeletal muscle	Opens cation channels
M_1, M_3, M_5	Target tissues	Increase Ca^{2+}
M_2, M_4	Target tissues	Decrease cAMP, open K^+ channels

Exceptions.

There are some exceptions to these rules, such as some sympathetic postganglionic neurons that terminate on sweat glands secrete norepinepherine rather than acetylcholine. These are called "sympathetic cholinergic neurons." There are also neurons which secrete none of these, are are called "nonadrenergic, noncholingeric neurons." There are also neurons which secrete multiple types of neurotransmitters.

8.3 Somatic Nervous System

The somatic motor division is exclusively excitatory, and neurons project from the CNS all the way to muscles. Inhibition can occur at the neuronal level (i.e., motor neurons are inibited) but not at the muscle level. Cell bodies of motor neurons are either within the ventral horn of the spinal cord or within the brain. Close to the muscle itself, the axons branches and connects to the muscle at the neuromuscular junction (NMJ), which includes both the target muscle and Schwann cell projections. The "synapse" equivalent is called the *motor end plate*, which is an area on the muscle with a high density of nicotinic receptors. The ECM contains acetylcholinesterase, responsible for breaking down acetylcholine. Notably, the nicotinic receptors on muscle cells are considered N_M, a slightly different version than is found on neurons (N_N). This difference has proven to be curical in examples such as α-bungarotoxin, which binds only to N_M[3].

8.4 Acetylcholine, Nicotinic Receptors, and Muscarinic Receptors

Acetylcholine (which I'll call ACh for this section) is unquestionably one of the most important molecules in the body. It is sythesized from choline (a hydroxyl azanium) and acetyl-CoA. Typical usage goes as follows: ACh is released from vesicles and binds to cholinergic receptors. Upon unbinding, acetylcholinesterase (AChE) breaks it down into acetate and choline. Choline is brought back into the pre-synaptic neuron through sodium cotransporters so that it may be reused. It is then re-combined with acetyl-CoA and repackaged and vesicles.

Funnily enough, nicotinic receptors (nAChRs) have been described as the most well understood membrane receptor[4]. Nicotinic receptors are fast opening non-specific cationic channels. They are expressed throughout much of the major structures in the brain, peripheral nervous system, and skeletal muscle. As such, drugs targeting these receptors are of considerable interest in many neurological

[3]Dee Silverthorn quite possibly wrote the greatest textbook known to humanity, didn't she?

[4]https://www.nature.com/articles/nrd2927

disorders. The binding site for ACh is composed of aromatic amino acids, namely W and Y. This paper describes continual exposure to drugs treating nAChRs as causing their eventual desensitization, while exposure to nicotine increases their expression greatly.

Both of these are of key interest in treating Alzheimer's disease, as loss of cholinergic synapses is one of its features (including significant reduction of both muscarinic and nicotinic receptors). Interestingly, Alzheimer's shows varying loss of nicotinic receptor subtypes across the brain. There is no shortage of drugs currently in development for reversing this.

Myasthena Gravis Digression.

Myasthena Gravis, an autoimmune disorder, is an example of an actylcholine dysregulation caused muscle disorder. As a digression within a digression, autoimmune disorders most often target endocrine organs, and the belief is that they are to protect against mutations causing hypersecretion[5]. Upon first glance, you may think that Myasthena Gravis evades this generality, but in fact it does not, as it is well associated with patients that dually have a thymoma (tumor of the thymus). Immune cells begin attacking cells with ACh recepetors, tagging them with antibodies. Therefore, Myasthena Gravis presents itself usually in the weaker muscles, such as those that control the eyes, as they have less ACh receptors and therefore a diminished response to signals. Thus, patients may have drooped eyelids (ptosis), double vision (diplopia), or trouble following moving objects with their eyes. The disease worsens with increased activity, as this causes more antibodies to be released onto ACh receptors—but improves on rest. Fascinatingly, men and women "get" the disease at different times (women typically under 40, while men over 60).

Antibodies may also target other genes (such as MUSK or LRP4) which are involved in ACh receptor localization, or other forms of regulation.

8.5 Central Pattern Generation

Defining the central pattern generator is still somewhat contested. Computational models have been used to explore its existence[6], and for now we should take a few unifying assumptions: **(1)** the central pattern generator exists and is used to generate rhythmic moving, such as walking, **(2)** in many cases it is initiated by the CNS, but aside from that is largely devoid from CNS input, and **(3)** the CNS can work in tandem to compute integrated moving that requires coordination and balance.

[5] https://www.cell.com/immunity/pdf/S1074-7613(20)30180-1.pdf
[6] https://www.nature.com/articles/s41598-021-91714-1

Part IV

Spinal Cord Injury

Maybe it sounds like I'm just spouting moral platitudes. But from a vagabond like me, it's not that. I can't begin to tell you how lonely I feel when I come across a beautiful view, then suddenly realize there's no one to enjoy it with me.
– Musashi by Eiji Yoshikawa

Chapter 9

The Injury Itself

Spinal cord injury is composed of the primary injury, prototypically, but not restrictively, due to some kind of high impact action. This is usually unpredictable and highly variable. Secondary injury, resulting from inflammation, oxidative stress, and other biological responses is much more predictable and potentially lends itself better to therapeutic intervention.

The lesion's composition is categorized in three ways: **(1)** the non-neural core, **(2)** the astrocytic scar around the core, and **(3)** the spare reactive neural tissue. In the mix of immune cell influx and scar formation, no neural cells can survive at the center of the lesion. On a neuronal level, the rostral end retracts in a process of Wallerian degeneration. The caudal end dies away. Growth from the cell body is limited both by the damaged cell's biochemistry and by the physical barriers which now present themselves in front of the axons. The physical barriers that immediately succeed injury are often called damaged axon-glia complexes (AGCs). Discussed further later, immune cell influx causes astrocytes to form a scar, meant to save the spare surrounding neural tissue, which is composed of both glia and neurons.

The traditional aim of treatment is to bridge the corticospinal tract with distant neurons through a therapeutic combination of inhibiting anti-regenerative and promoting regenerative factors. As I have commented many times, getting neurons to regenerate alone is insufficient in many cases, as reformation of the correct synapses will not necessarily follow. Forcing axon regeneration alone is, incidentally, not too hard—one can pump neurons full of metabolites or simply implant stem cells. The issue being that they do not know where to grow to. One possible route to solving this is remodeling neural circuits using interneurons to bridge these connections. There are also attempts to use biomaterials to simulate a pro-regenerative environment, hopefully enhancing plasticity of the circuits.

9.1 Cell Specific Responses

The discussion, for the moment, will mostly use information gathered from Hu *et al.* 2023.

9.1.1 Immune Response

As SCI breaks the spine-blood barrier, influx of immune cells can cause further damage. Evidently, the nature of the immune response being helpful or harmful is still largely contested.

Neutrophils.

Neutrophils compose part of the immediate response to injury, which are recruited by cytokines and chemokines secreted by cells damaged in the primary injury. They essentially initiate the secondary injury, and reach their peak around 1 DPI. Like most cells, the role of neutrophils cannot be characterize as solely pro- or anti-regenerative. While a high influx of neutrophils is associated with poor patient outcomes, so too are neutrophils associated with guiding macrophages to damaged tissue, suggestive of better recovery.

Microglia.

In mouse SCI models, it seems that there are two peaks of microglial activity. The time course is remarkably long and disparate, reported 7 DPI and 60 DPI. Microglia can either promote inflammation, thereby worsening the secondary injury (called the M1 phenotype) or decrease inflammation, and promote repair (called M2). It is likely that this response depends on the subtype of microglia, which varies depending on the environment. Regardless, it is true that the earlier one treats SCI, the more likely one is to avoid negative microgial effects. Fascinatingly, in a neonatal setting microglia are able to heal SCI almost entirely through their role secreting fibrinogen, which is able to connect damaged axons back together.

Macrophages.

Macrophages are considered to be the dominant immune cell located around the injury site. Microglia, conversely, are scattered around the borders of the injury. Depending on the type of glial scar that is formed, different types of macrophages have been found. Macrophages mediate the corralling[1] of cells around the injury site. The phagocyotitic abilities of macrophages are of key importance, as loose fragments of cells must be removed, and microglia are incapable of keeping up such a high demand for removal. As a large part of this includes the destroyed mylein of oligodendrocytes, macrophages uptake great amounts of lipids. This can result in the formation of lipid droplets, which causes macrophages to become "foamy." This foamy phenotype impairs further repair.

Lymphocytes.

The adaptive immunity is fairly universally regarded as harmful to regeneration (with some exceptions, of course). T cells further break down the spine-blood barrier and increase immune cell invasion. Evidently, T cell entry is also a major sourse of neuropathic pain in SCI patients. In general, it seems established that the overall immune response after injury impairs further regeneration, and a good example of this can be found here[2].

9.1.2 Neural Response

Apoptosis of neurons initiates at around 4 HAI, but peaks at only 8 HAI[3].

[1] Corralling is a term used to describe the formation of a barrier around the injury, preventing further injury. It is composed astrocytes, and other cells, and is important in repair.

[2] https://www.science.org/doi/full/10.1126/science.abd5926

[3] It is worth noting that these times are likely quite inaccurate, or very injury-type specific. But, they do give a good indication of the approximate timeframe—such as that the majority of this apoptosis occurs within the first day or so after injury.

Interneurons.

Fascinatingly, it was shown that the ability of neonatal mice to fully recover from SCI was due, in part, to interneurons maintaining excitatory conditions. In adult mice, these interneurons switch to inhibitory after SCI, which dampens signals to motor neurons. A paper investigating this can be found here[4]. One must wonder, can you electronically mimic the excitatory interneurons in fully grown mice? Similar approaches have been done therapeutically, such as with potassium-chloride cotransporter-2 (KCC2) agonist CLP290[5] which seems to dampen the overexcited, inhibitory interneurons. Perhaps one could simply use something like DBS (or in this case, DSS?) on these interneurons.

Astrocytes.

Astrocytes, being the dominant supportive cell, plays an essential role in SCI. After injury, astrocytes form a physical barrier that is supposedly intended to limit the secondary injury. This occurs after astrocytes become activated, and are helpful in the initial stages but later form a glial scar, impairing regeneration. Astrocytes may either be activated by inflammation, causing them to be neurotoxic (called A1 cells) or by ischemia, causing them to be neuroprotective (A2 cells). The first transformation occuring through the NF-κB path, and the second through STAT3. Microglia are said to be the greatest contributor to activating astrocytes through release of signaling molecules. Another important factor is type 1 collagen upregulation, which results in astrocytic adhesion through cadherin, causing activation and eventual scar formation.

The astrocyte scar is surprisingly thin, only a few layers of cells. Though, its importance is not to be underestimated. When ablated, mice with SCI were worse off by almost every metric. A cornerstone paper on the topic seems to be here[6]. While scar tissue is primarily astrocytic in origin, it is worth mentioning that pericyte derived scar tissue (sometimes called the fibrotic scar) too play a role. Their positive roles include boosting tissue integrity, but so too do they seem to block axon regeneration as a physical barrier.

Oligodendrocytes.

Oligodendrocytes reportedly begin apoptosis around 1 DPI and it peaks around 8 DPI. Oligodendrocyte precursor cells (OPCs) are may differentiate into oligodendrocytes or Schwann cells after SCI. OPCs have been show to remyelinate neurons after SCI, but fascinatingly, it has been shown that locomotor recovery after SCI does not necessarily require remyelination by oligodendrocytes[7]. Though, plenty of other research suggests it is required—so it is likely context dependent.

9.2 Side comments

This section serves to allow me to comment random things until I find a place for them later.

1. Non-neuronal cells secrete a collagen protein (Cthrc1), which is pro-regenerative and affects the ECM[8].

[4] https://www.nature.com/articles/s41593-022-01067-9
[5] https://www.sciencedirect.com/science/article/pii/S009286741830730X
[6] https://www.nature.com/articles/nature17623
[7] https://www.nature.com/articles/s41467-018-05473-1
[8] https://www.cell.com/developmental-cell/pdf/S1534-5807(20)30984-9.pdf

2. The ependyma is a source of stem cells, which contributes to glial scar formation after SCI. Connexin (gap junction) signaling between ependymal cells is supposedly a factor involved in their differentiation.

Chapter 10

In The Clinic & Therapeutic Approaches

10.1 Clinical Presentation

Death rates from SCI are still as high as 20% in some countries. There is a strong age dependence to this, as the probability of walking again after SCI in those older than 50 is much lower than those under.

Diagnosing Injury.

Immediate diagnosis is done through scanning, such as X-ray or CT scans, combined with general neurological exams, including voluntary or involuntary motor control tests. An alternative technique in neural evaluation, if the patient is not responsive, is electrophysiological recordings (either through EEG or EMG). X-rays are often used to immediately see large fractures, and follow-up CT scans for investigating the more possible hairline fractures. CT angiography may be used to investigate vascular destruction. More soft-tissue damage follow-up is done through MRIs. Interestingly, a trade-off exists in an MRI vs. immediate surgery. An MRI may be used to detect non-obvious issues, like disc herniations away from the primary injury site—which, if not fixed, will cause further degeneration. However, doing an MRI delays one's ability to decompress the spine.

Classifying Injury.

You will often hear injuries as being complete or incomplete. Another helpful distinction is discomplete, where the injury is considered clincially complete, but one can still observe connections through electrophysiology. The main method of classifying injury and tracking progress of patients is called the American Spinal Injury Association (ASIA) impairment scale. ASIA scores are broken into the following categories:

1. Grade A: Complete impairment, where there is no motor or sensory information being transmitted below the injury site[1].
2. Grade B: Incomplete impairment, where there is no motor information being transmitted, but some sensory information is preserved.
3. Grade C: There is some motor activity preserved, but more than half of the key muscles are too weak to move against gravity (Grade 3 muscles).

[1] It is not clear to me if this is measured by EMG or movement.

4. Grade D: A fair amount of motor activity is preserved, where at least half of the key muscles are above muscle Grade 3.

5. Grade E: There is no impairment at all!

Importantly, the injury site is often designated by the vertebrae that was fractured, but symptoms are due to the nerve pair that is damaged, which may be at a different location than the primary site of bone damage[2]. Note: this review had quite a bit more on clinical complications and is definitely worth going through in more depth! This discrepancy seems to be exacerbated the more caudal you go. Injury in the cervical portions can lead to severe bradycardia and hypotension, due to dysregulation of brain-heart communication, particularly regarding baroreceptor feedback. Too, damage to the vagus nerve can occur here, leading to dysregulation of most organs. Injury in the thoracic part may have widespread affects on the symathetic nervous system due to damage both of the spinal cord nerves and the nearby ganglia. Interestingly, an unconsidered byproduct of lower thoracic SCI is damage to motor signals to the legs. The main focus of such being that one loses their ability to walk, but accompanying this is reduced venous return, as veins rely on muscle movement to get blood back to the heart. Dampening of CNS-cardiovascular system communication seems to be one of the primary indicators of poor prognosis.

In evaluating the severity of the injury, a *spinal shock*, marked by temporary paralysis, may muddy the waters. While one may temporarily lose their reflexes, it can sometimes be regained soon later. However, the ability to define this state, and its duration, remains problematic. *Neurogenic shock* manifests similarly, but the cause is hypotension after SCI. This may be caused by hypovolaemia from blood loss, or pooling of blood due to reduced venous return. This occurs most often in SCI above T_6, as it is these sympathetic nervous which maintain vascular tone.

A few named pathologies exist, such as Central Cord Syndrome. This is the most common incomplete SCI. Often, this occurs in elderly patients who fall and already had some form of spondylosis[3]. It is marked by more damage to upper extremeties and possible incontinence. Brown-Séquard Syndrome occurs from penetrating SCI, such as a stab wound. It is usually characterized by sensory loss.

10.2 Treatments

For this section, I will largely ignore the uninteresting, obvious approaches to axon/neuron regeneration. This includes the often successful implantation of stem cells and or treatment with various neurotrophic factors. While these represent important measures in the field, they are largely intuitive and still face the broad issue of improper circuit reformation.

10.2.1 Electrical Stimulation

Fascinatingly, electrical stimulation has been used in conjunction with physical therapy in the past with good resuts[4]. The reasons may be that this promotes stem cell differentiation[5], or disrupts inhibitory interneuron signaling. The optimal electrical application for differentiation has been explored

[2] https://www.nature.com/articles/nrdp201718
[3] Weathering of the vertebrae.
[4] https://www.nejm.org/doi/pdf/10.1056/NEJMoa1803588?articleTools=true
[5] https://www.mdpi.com/2073-4409/11/5/846

extensively[6]. One may wonder if the benefits of BSI are in the interface itself, or simply the stimulation. Combining stem cell implantation and electronics is, likely, the future.

A slightly different rose by the same name is functional electrical stimulation (FES). Many trials have shown improvement in patients treated with either external stimulation or internal stimulation.

Light.

Light stimulation feels like a footnote in the electrical modulation story, to me. Though, if one wanted to control different neurons or enzymes on an alternate time course, optogenetic activation may be an option. The obvious issue being that one does not have genetic access to patients, and therefore would need to design (likely very complicated) targeted therapeutics.

Sound.

Another footnote is ultrasound simulation. In this case, it will be low intensity focused ultrasound. Some approaches have seen altered gene expression, but perhaps a more promising one is modulating mechanosensitive channels as was shown here[7]. Notably, this paper found that many mechanically activated channels are affected in ultrasound, including Piezo, and many of the Trp family proteins.

Magnetics.

I would be extremely curious to know if the magnetic field itself has any unique properties beyond its manipulation of the electric field. Still, too, the story is the same. Some seem to enhance channel activation, while others expression. Interestingly, transcranial magnetic stimulation (TMS) has been used as a treatment with some success. Incomplete spinal cord injury has seen improvements from TMS.

10.2.2 Biomaterials

The overall goal in the use of biomaterials is to block a worsened immune response, scar formation, and promote neuron activity. Adding promise to stem cell implantation is the use of biomaterials that enhance proper network reformation[8]. Theoretically, a perfect biomaterial could be a substrate preferable for neuron growth, contain molecules that inhibit the immune response, neurotrophic factors that enhance stem cell differentiation and recruitment, and ion channel agonists. Notably, to date there have been no major publications where a "cocktail" like this has been successful. These sorts of things are usually made from hydrogels, collagens, or select inorganic fibers.

An open question is how one could leverage biomaterials to help clear damaged parts of neurons/cells that would normally be cleared by phagocytosis. Perhaps, one could add materials that are easily oxygenated to dampen the blow of ROS.

To date, implantation of biomaterials have been relatively lackluster in treating patients. While some regeneration scaffolds have proven to improve some neurological function, no patient has regained motor function.

[6] https://www.frontiersin.org/articles/10.3389/fbioe.2021.591838/full
[7] https://www.nature.com/articles/s41467-022-28040-1
[8] https://pubs.rsc.org/en/content/articlepdf/2022/bm/d1bm01744f

10.2.3 Drug Treatment

Drug treatment primarily follows the same paths, being reduction of inflammation and neuroprotection. Methylprednisolone (MP) is the only drug approved to treat SCI and works through reducing inflammation. Notably, some side effects have been observed and therefore MP has fallen out of favor for treatment.

10.2.4 Surgery

I think you'll find that there is a disappointing lack of options—signaling the primitive nature of neurosurgery! Anyway, surgical intervention aims to restabilize the spinal cord as quickly as possible, particularly through decompression. Early surgery seems indicative of shorter ICU stays and better neurological recovery. The first day or so post SCI is the critical time. Interestingly, even after decompression, the pressure within the spine remains high due to fluid build-up within the dura matter. This makes blood reperfusion more difficult, leading to more problems. While durotomy is often a complication of surgery due to progressive CSF leakage after operation, in this case it can be helpful to lessen spinal pressure, which there evidently is a long tradition of[9]. Duroplasty is a more modern and sophisticated alternative, and can allow opening of the dura matter without as much risk[10].

Incision directly into the spinal cord itself, myelotomy, has also been done with some success. The belief is that it helps drain some of the harmful dying tissue. There seems to be time dependence in this, where if performed too late after injury it will simply reinvigorate inflammation.

10.2.5 Rehabilitation

As you would intuit, exercise is the most common technique, as it preserves muscle mass and promotes circuit reorganization. Another rehabilitative technique is pumping in a significant amount of oxygen, as ischemia occurs after injury. BSIs have also become more popular. Fascinatingly, decoding of handwriting has been used to generate text[11]. Though, these seem to require deep access to the brain.

10.3 Integrated Approach

So it seems to me that a good course of treatment would be something like:

1. Immediately after injury decompress the spine and implant a hydrogel with a good restorative milieu, and immunosuppressants.
2. As microglia peak around a week after injury, a few days later the hydrogel can be removed and stem cell implantation can begin.
3. Throughout this time, electrical stimulation can be applied.
4. Some time after this, rehabilitation through physical therapy can begin.

[9] https://www.sciencedirect.com/science/article/pii/0020138388901325
[10] https://www.liebertpub.com/doi/full/10.1089/neu.2014.3668
[11] https://www.nature.com/articles/s41586-021-03506-2

Part V

Brain-Spine Interfaces

When we stand at the edge of the ocean, we can not understand its vastness. We know only that it is big. Ittō Ittosai is big. And what about Kojirō?
– Reworded from Vagabond by Takehiko Inoue

10.4 Perspective

There is an interesting part of Yoshikawa's Musashi which details how a vagabond had a poor reputation, despite being an incredible swordsman and a generally giving and selfless person. Their reputation came from those who had long-standing, baseless grudges against them, or those that disagreed with their rise to fame, either out of jealousy or other. A monk in the story comments that a vagabond under 30 should almost never have a pure reputation, as they couldn't possibly be established enough to be beyond critique or ridicule. It is implied that for one to accomplish anything great, it is inherent that one must take risks, be bold, and likely ruffle feathers along the way. Irrespective of one's identity, it is certain that there will be those that dislike them or their approach.

I mention this specifically in this section because it is likely that in advancing technology, you will encounter people who want to belittle your ideas. It is only natural that those who have worked in a field for a long time may be inclined to retain their current way of life, thereby pushing against unseen advancement. The canonical example being those surgeons that resisted the endoscope. I add this to say that it is not unlikely that you will do all the right things, but still find those who disapprove.

Pay them no mind, or better yet, learn from them. Perhaps if you learn from all of those that are unwilling to learn from you: your growth will be unmatched, and real progress can be made. Best of luck.

Chapter 11

Biochemistry

While one may find it odd to begin an electronics discussion with a chapter on biochemistry, the fact of the matter is an integrated approach is the only way. As such, to be able to interface with the electrophysiology of neurons, one must fully understand it.

11.1 Ion Channels

Chapter 12

The Technology of SCI

12.1 Historical Dealings

A proverb in a writing called *Testament* by Daitō Kokushi reads:

> I beg you, try to find the fundamental source.
> ...
> Like our great predecessors,
> Do not merely pinch off the leaves
> Or concern yourselves only with the branches.

Using electronics to study neural circuits has a long and storied history. One can think of Hodgkin and Huxley, whose ubiquitous nature avoids a need for citation, using an operational amplifier to clamp the voltage of giant squid axons to learn of their properties. In the case of solving spinal cord circuits, so too have electronics long been used[1].

Restoration of the signals which pervade severance in spinal cord injury through therapeutics has been equally explored. For example, drugs such as clonidine, α_2 receptor agonist, can promote walking in cats with spinal cord injury[2]. This effect can be blocked completely by yohimbine, an α_2 receptor antagonist[3].

The ability to generate walking-like movements from electrical stimulation of the lower spine, the T10-L1 region, has been long known[4,5]. In these instances, "tonic" stimulation was able to generate step-like movements. Although the term "tonic" is dubious because other papers by the same lab identified frequency dependence in their stimulation[6]. While one can generate step-like movements at once frequency, continued extension can be generated with another.

[1] https://journals.physiology.org/doi/epdf/10.1152/jn.1943.6.2.111
[2] https://journals.physiology.org/doi/full/10.1152/jn.1998.79.6.2941
[3] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2278596
[4] https://www.sciencedirect.com/science/article/pii/B9780444521378000188
[5] https://journals.sagepub.com/doi/pdf/10.1177/1073858417699790
[6] https://link.springer.com/article/10.1007/s00422-004-0511-5

12.2 Electrostimultaion and Accompanying Cells

Specific Cells.

Using a simplified device similar to what will be discussed in the upcoming Lorach *et al.* 2023 section, the Courtine group used electrical therapy to re-establish walking in a number of SCI patients[7]. Their paper goes into depth on the cell-type specific response to this electrical modulation. They called this model epidural electrical stimulation (EES) + rehab, or $\text{EES}^{\text{REHAB}}$. They first approached by measuring metabolic consumption in the spinal cord before and after $\text{EES}^{\text{REHAB}}$ using PET scanners, and found that metabolism actually **decreases** in the face of $\text{EES}^{\text{REHAB}}$, despite regaining the ability to walk. Therefore, they hypothesized that $\text{EES}^{\text{REHAB}}$ "cleans up" the circuits a bit, causes activity to become targeted for walking.

To investigate which specific neurons undergo changes, they used a mouse model and performed RNAseq. They found excitatory interneurons as being enriched upon $\text{EES}^{\text{REHAB}}$. These neurons were identified by the markers $Vsx2$ and $Hoxa10$. These neurons projected exclusively to the ventral spinal cord and formed synapses with varying neuron types (glutamatergnic, GABAnergic, and cholinergic in no prefered proportion). The synapses were found in dense appositions (i.e., many synapses side-by-side) and SCI caused a reduction in such appositions. Ablating/inactivating these neurons alone in non-SCI mice did not halt their ability to walk. However, inactivating them did halt walking under $\text{EES}^{\text{REHAB}}$. Therefore, it is hypothesized that these neurons are specific to the injury condition and required for recovery after SCI.

The main takeaway from this paper, in my opinion, is simply the recapitulation that excitatory interneurons play a key role in recovery from SCI. This has been established now through many routes, and is certainly of considerable interest in the future. There are a few open questions, with the biggest being: how exactly does electric modulation drive circuit reformation through these interneurons?

12.3 Reading Thoughts

At present, I am unsure if reading one's thoughts is a requirement at all for overcoming SCI. I am unconvinced that electrical stimulation controlled by one's brain is better than electrical neurorehabilitation—or if it is better by any measurable margin. In investigating this, I believe the limiting factor in bridging from the stimulation applied by Tator and Minassian is a reliable way to read ones objective through their brain waves. The Lorach *et al.* 2023 solved this using the WIMAGINE system[8]. But, let's begin by discussing the options and a bit of history.

12.3.1 Different Reading Devices

As you would imagine, there is a tradeoff between the resolution and the damage you will confer upon one's brain. This is described well by Schwartz in 2006, so notably technology may have advanced since then[9]. The four classes of devices are:

1. Electroencephalography (EEG)

[7] https://www.nature.com/articles/s41586-022-05385-7
[8] https://www.frontiersin.org/articles/10.3389/fnins.2019.00847/full
[9] https://www.sciencedirect.com/science/article/pii/S0896627306007264

2. Electrocorticography (ECoG)
3. Local field potentials (LFPs)
4. Single Unit AP

EEGs, naturally, are non-invasive and have a reading range on the order of a few centimeters. ECoG is on the tenths of centimeters level, while LPS is millimeters level. Single unit AP is as the name describes, and reads individual action potentials. EEGs sit far above the cortex, outside of the skull, and ECoG come in contact with it. Both LPS and single unit require the electrodes to be buried within the cortex itself, thus causing damage to the brain tissue. Therefore, EEG is typically preferable, as it does not require surgery. It is highlighted by Schwartz that many negatives can occur long term from invasive procedures, such as degeneration, volume displacement, or glial encapsulation. Therefore, it is for this reason that I presume the ability to read ones objective was the limiting factor in progressing the field.

For the moment, I will avoid describing individual technologies between the four groups, as it is likely that the tech progresses greatly each year, and it would be pointless to attempt to describe the reading ranges of any without being totally up-to-date. Instead, we will discuss the WIMAGINE in depth here.

Soft vs. Hard Electronics

The importance of the electronic's physical form has been considered by the Courtine group as well[10]. In their case, they looked specifically at the electrode paddly inserted into the spinal cord. Notably, they did not use this device in their recent, Lorach *et al.* 2023 paper. One problem I personally find is that in using hard electronics, one immediately resolves to doing massive cranioplasty in order to add their device. If one goes the soft-only electronics route, perhaps you can insert the electronics beneath the skull itself.

WIMAGINE ECoG.

As the name suggests, the device requires the skull to be torn away and must sit directly upon the cortex. It features 64 electrodes, and a human patient would need to have two implanted in order to read both sides of one's motor cortex.

The device was tested longest in sheep, which was a 10 month trial. Surprisingly, at the end of the 10 month trial, through GFAP staining, they still found a great deal of glial migration/build-up around the site. They did not quantify this, but in my personal opinion, this does not bode well for our patient—as they will have to spend a lifetime with the device. Secondly, calcification of the dura matter over the device had begun. This is not inherently surprising, and perhaps is a good sign that the bone was not irreversibly damaged. However, as the device is wirelessly charged, and wirelessly transmits brain information, one must wonder how years of calcium buildup may impact the ability to send or receive information. Too, if the technology needs a dust-up, the surgeons will have to re-destroy these calcified layers. There is quite a lot of fuss made about the fact that the device is wireless, and how it was a decision made to best serve patient comfort. But, in my opinion, it may spell out more long-term harm than good. It is not clear to me yet whether the 8×8 resolution of the electrodes is sufficient to delineate all necessary information. Though, this is the trade-off one must

[10] https://www.science.org/doi/10.1126/science.1260318

make in choosing between EEG, ECoG, and LPS. The WIMAGINE system digitizes information with 12 bits of resolution. Data is processed through pwelch spectral analysis. From what I can gather, this is simply another implementation of the FFT.

12.4 Lorach et al. 2023

This is the most recent iteration of an ongoing project by the Courtine group.

12.4.1 Overview

The group used surgical implantation into the skull, over the motor cortex, to record patient movements and wirelessly transmit this to the spinal cord. The technology used to capture these electrocorticographic (ECoG) signals is the WIMAGINE. The design features two antennas, and the second of which transmits motor signals that are to be decoded and sent to a pulse generator. The pulse generator is ACTIVA RC, the same that is used in deep brain stimulation (DBS) or pacemakers in the heart. There is no mention of having the patient replace this in future years, but my assumption is that a new battery will require a new surgery in years to come.

12.4.2 Shortcomings

The three shortcomings explicitly mentioned in the **Introduction** are that **(1)** one using this BSI must have motion sensors on in order to compensate (a sort of PID controller, perhaps), **(2)** that the patient's movement was not perceived as natural, and **(3)** that there was still considerable hurdles in traversing variable terrain. Many more shortcomings exist than this, and we will go through them below.

Motion Sensors

I suppose it is obvious why one would use a motion sensor rather than an EMG—as it would be too difficult to discern whether the motion itself, or strength of contraction, is too high using an EMG alone. Still, though, I wonder if this is a feasible alternative. Perhaps one can simply predict motion based on EMG patterns.

The largest drawback, in my opinion, is their desire to make it quickly programmed/calibrated. As with DBS, one would actually prefer longer calibration time. That is, one would be happy to wait a few extra hours or days if it makes their ability to walk improve.

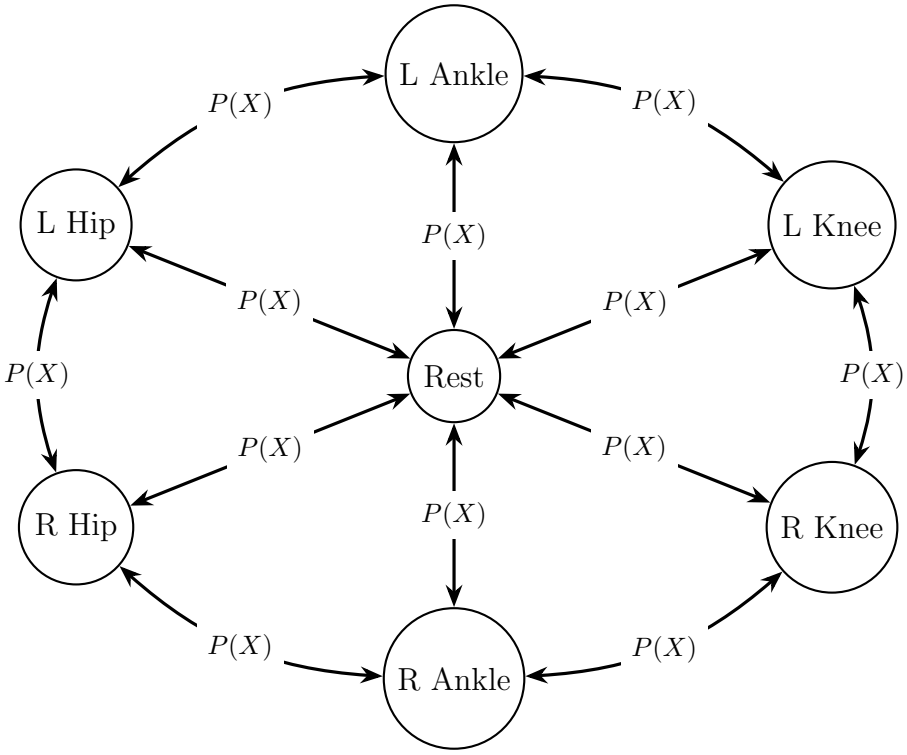
Similarly, they commented that the patient tested was able to walk without the BSI turned on (using crutches) after some time. This signals three things to me, **(1)** this patient's injury was not, per se, insurmountably severe, **(2)** the plastic abilities of a damaged spinal cord may be better than we had thought—especially since the patient's injury predated this treatment by 10 years, **(3)** that there was likely some important tuning being done by the brain-spine endogenous connection (i.e., their BSI was not doing all of the work), and **(4)** in some cases, no BSI is likely needed at all, and electrical stimulation may be sufficient. This in itself is a quite interesting, alternative patient state. For example, one can imagine a scenario where a patient's brain-spine communication is deeply dampened, but not entirely ceased by damage. Therefore, one would only need amplification of signals as opposed to

complete generation of new ones. There may be some of that here. Another comment to make is that their signals are incredibly un-sophisticated, as presented in this paper. That is, they use 16 random electrodes, which do not have connections to synapses themselves. Perhaps this design is to hijack a more reflex-based approach to walking—hence why one can not traverse variable terrain that goes beyond reflexive movement.

A comment regarding the surgery itself: The patient's are having two large, 50mm arrays placed into their skull. Naturally, most would agree to this in order to restore their ability to walk. However, it surprises me that they chose to make the device wireless, given that the surgery is already quite invasive.

Seven States

Data is processed and used to edit the probability distribution of a Markov chain, which progresses the electrode paddle through seven defined states defined below.



Notably, they did not explicitly mention that one of the progressions could be from state to self state, but one can only assume this is true. So too, only seven states seems to limit the mobility of a patient.

Part VI

Project Proposal

Man takes up the sword in order to shield the wound in his heart sustained in a far-off time beyond remembrance. Man wields the sword so that he may die smiling in some far-off time beyond perception.
– Berserk by Kentaro Miura

Chapter 13

Outline

This section will largely be messy—likely for the coming years. It is for me to jot down ideas I have through reading.

1. Inhibit L-type VGCCs
2. Optimize stiffness of paddles
3. Optimize ECoG

Inhibit L-type VGCCs.

It is well known by now that electrical stimulation combined with rehabilitation improves patient's ability to walk after SCI.

Thank you. There's one thing I'd like to leave you with:

The world is always full of the sound of waves.

The little fishes, abandoning themselves to the waves, dance and sing and play, but who knows the heart of the sea, a hundred feet down? Who knows its depth?
– Musashi by Eiji Yoshikawa