6.002x notes

Thomas Backman, exscape@gmail.com

February 10, 2022

Contents

Chapter 1

Formulas and such

1.1 Intro

This chapter contains several common formulas etc., without any further explanation of how they are derived.

1.2 Components in series and parallel

Series resistances add:

$$R_s = R_1 + R_2 + R_3 + \cdots$$

Parallel resistances diminish:

$$R_p = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots}$$

Impedances behave like resistances.

Series capacitances diminish:

$$C_s = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots}$$

Parallel capacitances add:

$$C_p = C_1 + C_2 + C_3 + \cdots$$

Series inductances add:

$$L_s = L_1 + L_2 + L_3 + \cdots$$

Parallel inductances diminish:

$$L_p = \frac{1}{\frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + \cdots}$$

For the special case of two components of the diminshing type (x being a dummy variable):

$$\frac{1}{\frac{1}{x_1} + \frac{1}{x_2}} = \frac{x_1 \cdot x_2}{x_1 + x_2}$$

1.3 Digital

1.3.1 Thresholds

- V_{OL}: the *highest* allowed output voltage for a logical 0 (lower is OK)
- V_{IL}: the *highest* allowed input voltage for a logical 0 (higher would be in the forbidden region)
- V_{IH}: the *lowest* allowed input voltage for a logical 1 (lower would be in the forbidden region)
- V_{OH}: the *lowest* allowed output voltage for a logical 1 (higher is OK)

$$0 \leq V_{OL} < V_{IL} < V_{IH} < V_{OH} \leq V_S$$

1.3.2 Noise margins

$$\begin{aligned} NM_H &= V_{OH} - V_{IH} \\ NM_L &= V_{IL} - V_{OL} \\ \text{Forbidden region} &= V_{IH} - V_{IL} \end{aligned}$$

1.4 MOSFETs

1.4.1 SCS model

$$i_{DS} = \begin{cases} \frac{K}{2}(v_{GS} - V_T)^2 & \text{for } v_{GS} \geq V_T, v_{DS} \geq v_{GS} - V_T \\ 0 & \text{for } v_{GS} < V_T \end{cases}$$

The above covers saturation and cutoff regions only (the SCS model).

1.4.2 SR model

$$i_{DS} = \begin{cases} \frac{V_{DS}}{R_{ON}} & \text{for } V_{GS} \geq V_T \\ 0 & \text{otherwise} \end{cases}$$

Alternatively:

$$R_{DS} = \begin{cases} R_{ON} & \text{for } V_{GS} \geq V_T \\ \infty & \text{otherwise} \end{cases}$$

Used for digital circuits only (in this course).

1.5 State devices / energy storage devices

This section assumes time-invariant devices, i.e. capacitance/inductance is a fixed value, and not a function of time.

1.5.1 Capacitors

The current through a capacitor is a function of the *rate of change* of voltage:

$$i(t) = C\frac{dv(t)}{dt}$$

To find the voltage over a capacitor, we need to know its full history:

$$v(t) = \frac{1}{C}\int_{-\infty}^{t} i(t)dt$$

... or we can simply do it by knowing the current through it between t_1 and t_2, plus the initial voltage:

$$v(t_2) = \frac{1}{C}\int_{t_1}^{t_2} i(t)dt + v(t_1)$$

The energy stored in a capacitor is:

$$E = \frac{1}{2}Cv^2$$

1.5.2 Inductors

The voltage over an inductor is a function of the *rate of change* of current:

$$v(t) = L\frac{di(t)}{dt}$$

To find the current through an inductor, we need to know its full history:

$$i(t) = \frac{1}{L}\int_{-\infty}^{t} v(t)dt$$

... or we can simply do it by knowing the voltage over it between t_1 and t_2, plus the initial current through it:

$$i(t_2) = \frac{1}{L}\int_{t_1}^{t_2} v(t)dt + i(t_1)$$

The energy stored in an inductor is:

$$E = \frac{1}{2}Li^2$$

1.6 First order circuits

General equation for an increasing/decaying exponential in a RC/RL circuit:

$$v = V_S + (V_0 - V_S)e^{-t/RC}$$
$$i = \frac{V_S}{R} + (i_0 - \frac{V_S}{R})e^{-Rt/L}$$

where v = the voltage across the capacitor. For RL circuits, replace V_S by the maximum current and V_0 by the initial current through the inductor. Above is for a thevenin equivalent circuit.

ZIR / Zero Input Response, same as above with $V_S = 0$:

$$V_0 e^{-t/RC}$$

ZSR / Zero State Response, same as above with $V_0 = 0$:

$$V_S(1 - e^{-t/RC})$$

1.7 Second order circuits, impedance, filters

Canonical form of the characteristic equation for second-order circuits; use this to match up the values of α and ω_0 for a circuit:

$$s^2 + 2\alpha s + {\omega_0}^2 = 0$$

1.7.1 For all LC and RLC circuits:

Natural/undamped resonant radian frequency: ω_0 (rad/s)

Damping factor: α (rad/s)

Note that zeta (ζ) is used as a damping factor in many texts; it is defined as

$$\zeta = \frac{\alpha}{\omega_0} \text{ (dimensionless)}$$

The bandwidth $\Delta\omega$, i.e. the *width* of the frequency *band* that is above $\frac{1}{\sqrt{2}}$ times the input amplitude, is given by

$$\Delta\omega = 2\alpha \text{ (rad/s) (measured at } \frac{1}{\sqrt{2}} \text{ points)}$$

$$\text{Quality factor: } Q = \frac{\omega_0}{2\alpha} \text{ (dimensionless)}$$

RLC circuits can be underdamped, overdamped, or critically damped.

$$\text{Underdamped: } \omega_0 > \alpha \text{ or, equivalently, } Q > \frac{1}{2} \text{ or, equivalently, } \zeta < 1$$

$$\text{Overdamped: } \omega_0 < \alpha \text{ or, equivalently, } Q < \frac{1}{2} \text{ or, equivalently, } \zeta > 1$$

$$\text{Critically damped: } \omega_0 = \alpha \text{ or, equivalently, } Q = \frac{1}{2} \text{ or, equivalently, } \zeta = 1$$

When they are *underdamped*, the *damped resonant frequency* ω_d is given by

$$\omega_d = \sqrt{{\omega_0}^2 - \alpha^2}$$

The naming here might be confusing; the *damped* frequency is used for *underdamped* systems. The reason is that the *undamped* frequency is used for systems with no damping whatsoever, i.e. LC circuits with no resistor.

Underdamped RLC circuits are the only kind that oscillate, so the natural frequency is less interesting for overdamped circuits.

1.7.2 Series RLC circuits

$$\omega_0 = \frac{1}{\sqrt{LC}} \text{ rad/s}$$

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \text{ Hz}$$

$$\alpha = \frac{R}{2L} \text{ rad/s}$$

$$\Delta\omega = 2\alpha = \frac{R}{L} \text{ rad/s}$$

$$\text{Period: } \frac{2\pi}{\omega_0} = 2\pi\sqrt{LC} \text{ seconds}$$

$$Q = \frac{\omega_0}{2\alpha} = \frac{L}{R\sqrt{LC}} \text{ (dimensionless)}$$

1.7.3 Parallel RLC circuits

$$\omega_0 = \frac{1}{\sqrt{LC}} \text{ rad/s}$$

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \text{ Hz}$$

$$\alpha = \frac{1}{2RC} \text{ rad/s}$$

$$\Delta\omega = 2\alpha = \frac{1}{RC} \text{ rad/s}$$

$$\text{Period: } \frac{2\pi}{\omega_0} = 2\pi\sqrt{LC} \text{ seconds}$$

$$Q = \frac{\omega_0}{2\alpha} = \frac{RC}{\sqrt{LC}} \text{ (dimensionless)}$$

1.7.4 Frequency- to time-domain conversion

You can find the time-domain behavior of a circuit to sinusoidal input from nothing but a complex amplitude of the form V_x:

$$v_X(t) = |V_x| \cos\left(\omega t + \angle V_x\right)$$

See below for information about how to calculate the magnitude $|z|$ and the angle $\angle z$ of a complex number.

1.7.5 Complex algebra

A few properties of complex numbers that are necessary to know:

$$|a + jb| = \sqrt{a^2 + b^2}$$

$$\angle(a + jb) = \arctan\left(\frac{b}{a}\right) \text{ or, preferably, } \operatorname{atan2}(a, b)$$

$|a + j0| = a$ if $a > 0$; otherwise, the magnitude is just the absolute value $|a|$

$$|0 + jb| = b$$

$$|0 - jb| = b$$

$$\angle(a + j0) = 0$$

$$\angle(0 + jb) = \frac{\pi}{2}$$

$$\angle(0 - jb) = -\frac{\pi}{2}$$

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|$$

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$$

$$\angle(z_1 \cdot z_2) = \angle z_1 + \angle z_2$$

$$\angle\left(\frac{z_1}{z_2}\right) = \angle z_1 - \angle z_2$$

When finding angles, it's best to use a function which is capable of giving correct answers in all quadrants of the unit circle. The "atan2" function was created in many computer languages for this purpose. It is equal

to $\arctan(\frac{b}{a})$ for some inputs, but not all; arctan() cannot differentiate between all quadrants, because $\arctan\left(\frac{-a}{-b}\right) = \arctan\left(\frac{a}{b}\right)$ and likewise, $\arctan\left(\frac{-a}{b}\right) = \arctan\left(\frac{a}{-b}\right)$. Thus, the atan2 function has two arguments, and when using it, the angle of a complex number $a + jb$ is simply $\text{atan2}(a, b)$.
The angle of a fully real number is always 0, and the angle of a fully imaginary number always either $\frac{\pi}{2}$ (for positive imaginary numbers) or $-\frac{\pi}{2}$ (for negative imaginary numbers).
One definition of atan2, if you are not using math software that has it, is:

$$\text{atan2}(b, a) = 2\arctan\left(\frac{b}{\sqrt{a^2 + b^2} + a}\right)$$

However, do note that these notes use an atan2 function that is defined with the variables in the opposite order, i.e. atan2(a, b); specifically, Mathematica's ArcTan[a, b] function.

1.7.6 Impedances

$$\text{Resistor: } Z_R = R$$

$$\text{Capacitor: } Z_C = \frac{1}{j\omega C}$$

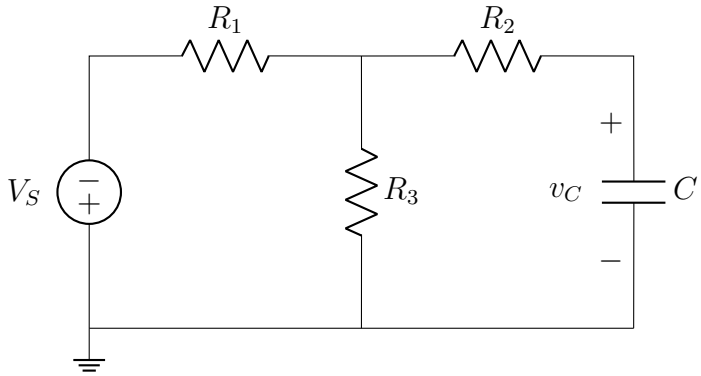
$$\text{Inductor: } Z_L = j\omega L$$

Chapter 2

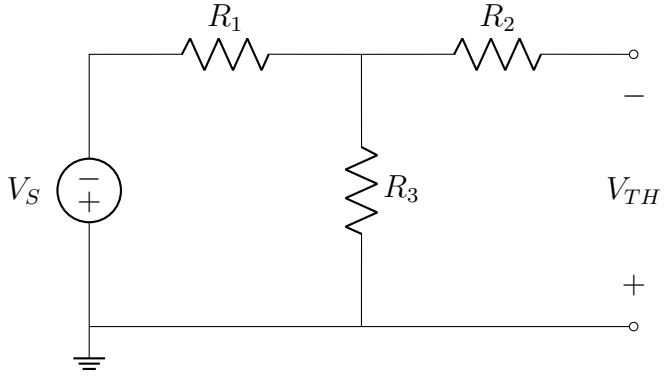
Circuit analysis

2.1 Thevenin equivalent circuits

Say we have an capacitor circuit to analyze:



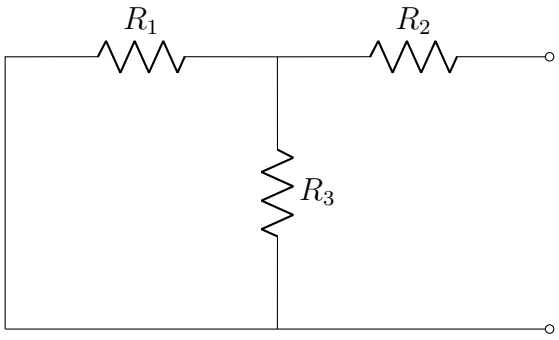
Since this is a linear network, we can simplify it by calculating its *Thevenin equivalent*. Consider the network as seen from the port where the capacitor is attached:



V_{TH}, the open circuit voltage, will be given by the voltage divider formed by R_3 and R_1:

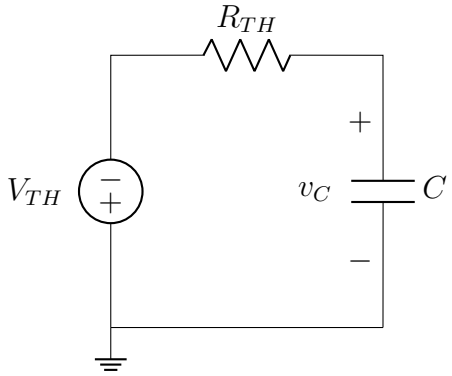
$$V_{TH} = \frac{R_3}{R_1 + R_3} \cdot V_S$$

Since no current flows at the port (for the *open circuit* voltage!), R_2 doesn't contribute at all.
We also want to measure the resistance "looking in" to this port; this will be the Thevenin resistance R_{TH}. To do this, we turn off all *independant* voltage and current sources, by replacing all current sources with *opens* and all voltage sources with *short circuits*.
Leave dependant sources in the circuit!



$$R_{TH} = R_2 + (R_1||R_3) = R_2 + \frac{R_1 \cdot R_3}{R_1 + R_3}$$

Now that we know the Thevenin voltage V_{TH} and the Thevenin resistance R_{TH}, we can replace the circuit with a voltage source of voltage V_{TH} volts in series with a resistor of value R_{TH} ohm, and place the capacitor back into the circuit:



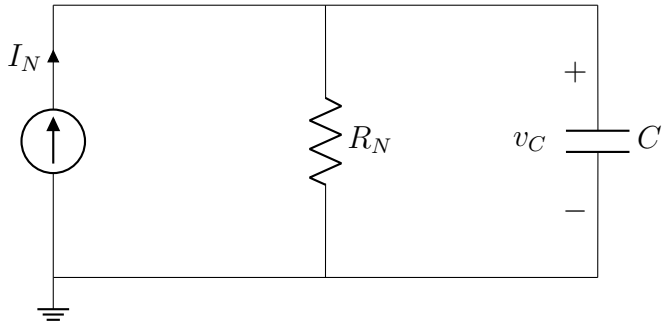
Our previous circuit has now turned into a simple RC circuit, which is easier to analyze. See the chapter on RC circuits.

As a side note, another way of measuring the Thevenin resistance is to short circuit the output node, calculate/measure the short-circuit current (with all sources left intact, of course), and calculate R_{TH} as $\frac{V_{TH}}{I_{SC}}$.

In summary:
• Calculate/measure the open circuit voltage V_{TH} at the port
• Turn off all independent sources (make short circuits of voltage sources, and open circuits of current sources), but leave dependant sources intact
• Calculate/measure the resistance R_{TH} at the port terminal pair
• Replace the original circuit with a series circuit of a voltage source (voltage V_{TH} volts), a resistor (resistance R_{TH} ohm) and the element you want to analyze.

2.2 Norton equivalent circuits

Nortan equivalent circuits are very similar to Thevenin equivalents, but use a *current source* in *parallel* with a resistor rather than a *voltage source* in *series* with a resistor.
To convert a circuit to its Norton equivalent:
• Calcurate/measure the *short circuit current*, i.e. the current that would flow through the output port if we were to short-circuit it. The result is the Norton current I_N.
• Turn off all independent sources (make short circuits of voltage sources, and open circuits of current sources), but leave dependant sources intact.
• Calculate/measure the resistance at the port terminal pair; the result is the Norton resistance R_N.
• Replace the original circuit with a parallel circuit of a current source (current I_N amperes), a resistor (resistance R_N ohm) and the element you want to analyze.



Note that since the method for calculating the equivalent resistance is identical for the Thevenin and Norton methods, $R_N = R_{TH}$. It is easy to convert between one and the other:

$$R_N = R_{TH}$$
$$I_N = \frac{V_{TH}}{R_{TH}}$$
$$V_{TH} = I_N \cdot R_N$$

Chapter 3

Small signal method

3.1 Deriving small signal models

For a device with $i_D = f(v_D)$, the small signal current i_d is given by

$$\underbrace{\left.\frac{\partial f(v_D)}{\partial v_D}\right|_{v_D=V_D}}_{g_m} \cdot v_d$$

where v_d is the small signal input voltage.
In other words, take the partial derivative of the V-I relation, with respect to the voltage. That gives g_m, the transconductance. The transconductance multiplied by the small signal input voltage v_d gives the small signal current.

As an example, a MOSFET in the saturation region has $i_{DS} = f(v_{GS}) = \frac{K}{2}(v_{GS} - V_T)^2$:

$$i_{ds} = \underbrace{\left.\frac{\partial \frac{K}{2}(v_{GS} - V_T)^2}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}}}_{g_m} \cdot v_{gs} = \underbrace{K(V_{GS} - V_T)}_{g_m} \cdot v_{gs}$$

Note the difference between v_{GS} (the total gate-to-source voltage), V_{GS} (the DC bias voltage) and v_{gs} (the small signal / incremental voltage).
Also note that the value of g_m depends not only on the MOSFET parameters K and V_T, but also on the user-chosen DC bias voltage V_{GS}.

Here's a table of common circuit elements and their small signal equivalents:

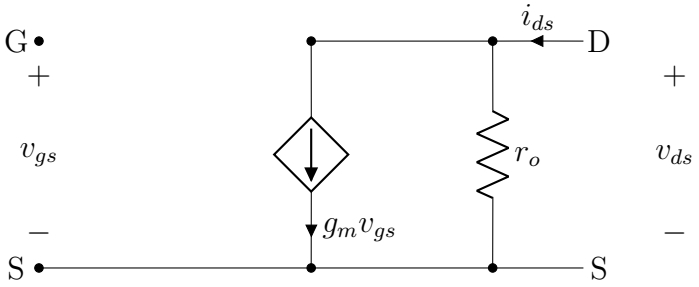
Device	Small signal replacement
Resistor, R Ω	Resistor, R Ω
Voltage source, V volts	Short circuit
Current source, I amps	Open circuit
MOSFET in saturation region	VCCS, $i_{ds} = g_m v_{gs}$, $g_m = K(V_{GS} - V_T)$
MOSFET with gate/drain tied together	Resistor, $\frac{1}{K(V_{DS} - V_T)}\Omega$ (for $v_{DS} > V_T$)

3.2 Multivariable small signal models

In some cases, it might be necessary to make a small signal model of a device where the current depends on more than one variable. An example (that will be used here) in 6.002x is the (probably hypothetical) "NewFET" in the week 6 homework.
In this case, we take the partial derivative of the gate-to-source voltage at the bias point times the small signal voltage *vgs*, *plus* the partial derivative of the drain-to-source voltage at the bias point times the small signal voltage *vds*.
First, the properties of the NewFET:

$$iDS = \begin{cases} 0 & \text{for } v_{GS} < V_T \\ K(v_{GS} - V_T){v_{DS}}^2 & \text{for } v_{GS} \geq V_T \end{cases}$$

The small signal model will look like this:



The small signal current i_{ds} will depend on both v_{gs} and v_{ds}, in this manner:

$$i_{ds} = v_{gs} \underbrace{\left.\frac{\partial i_D}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}}}_{g_m} + v_{ds} \underbrace{\left.\frac{\partial i_D}{\partial v_{DS}}\right|_{v_{DS}=V_{DS}}}_{1/r_o}$$

g_m will be the regular transconductance, calculated the same as with MOSFETs (see the above section):

$$g_m = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}} = \left.\frac{\partial K(v_{GS} - V_T){v_{DS}}^2}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}} = K{V_{DS}}^2$$

r_o will be the *reciprocal* of the partial of i_D with respect to v_{DS}:
(In other words, the partial will give a conductance, and we want a resistance.)

$$\frac{1}{r_o} = \left.\frac{\partial i_D}{\partial v_{DS}}\right|_{v_{DS}=V_{DS}} = \left.\frac{\partial K(v_{GS} - V_T){v_{DS}}^2}{\partial v_{DS}}\right|_{v_{DS}=V_{DS}} = 2KV_{DS}(V_{GS} - V_T)$$

So

$$r_o = \frac{1}{2KV_{DS}(V_{GS} - V_T)}$$

From these equations and the circuit diagram, we see that

$$i_{ds} = \frac{v_{ds}}{r_o} + g_m v_{gs} = v_{ds} \cdot 2KV_{DS}(V_{GS} - V_T) + K{V_{DS}}^2 \cdot v_{gs}$$

Note that, although the expression contains a square term (${V_{DS}}^2$), it is still linear, as the square term is a constant - the bias voltage V_{DS} should not change, or the entire small signal model will be incorrect either way.
So, the result is not quite the simplest of expressions, but when the bias constants are replaced with their actual bias values, the result is of the form

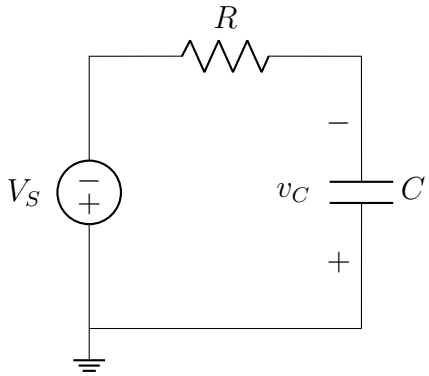
$$i_{ds} = C_1 v_{ds} + C_2 v_{gs}$$

where C_1 and C_2 are constants.

Chapter 4

First-order circuits

4.1 Series RC circuits



We start off by writing down a KCL equation for the unknown node voltage v_C:

$$\frac{v_C - V_S}{R} + C\frac{dv_C}{dt} = 0$$

Rewrite the equation to get it in the form we prefer:

$$RC\frac{dv_C}{dt} + v_C = V_S$$

To solve this first-order linear differential equation, we'll use the method of particular and homogeneous solutions, where we need to find two solutions to the differential equation: the first (the *particular* solution) is *any* solution that makes the equation true:

$$RC\frac{dv_{Cp}}{dt} + v_{Cp} = V_S$$

We see here that if we pick $v_{Cp} = V_S$, where V_S is a constant, thus making $\frac{dV_S}{dt} = 0$, this equation is indeed true; since the first term becomes 0, all that remains is $V_S = V_S$ - which is clearly true!
Thus, we've found the particular solution:

$$v_{Cp} = V_S$$

The next step in this method is to find a solution to the homogeneous equation, where V_S (the "input drive") is zero:

$$RC\frac{dv_{Ch}}{dt} + v_{Ch} = 0$$

We need a function such that its derivative is the function itself times a constant. e^x comes to mind - more specifically, the solution will have some (still unknown) coefficients A and s, such that:

$$v_{Ch} = Ae^{st}$$

We substitute Ae^{st} into the homogeneous equation and end up with:

$$RC\frac{dAe^{st}}{dt} + Ae^{st} = 0$$

We calculate the derivative and replace the $\frac{d}{dt}$ term with it:

$$RCAse^{st} + Ae^{st} = 0$$

Divide both sides by Ae^{st}:

$$RCs + 1 = 0$$
$$RCs = -1$$
$$s = -\frac{1}{RC}$$

We've thus found one of our two constants.
The total solution to the differential equation will be the *sum* of the particular and homogeneous solutions, so the next step is to add them together:

$$v_C(t) = v_{Cp}(t) + v_{Ch}(t)$$
$$v_C(t) = V_S + Ae^{-\frac{1}{RC}\cdot t}$$

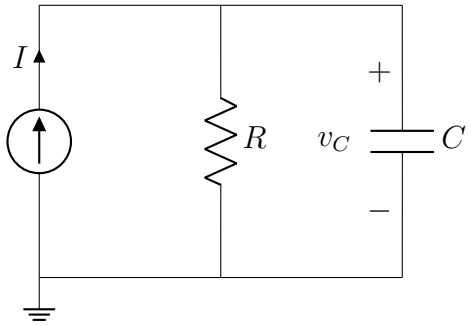
All that remains to do is to find the value of the constant A. To do so, we substitute v_C for the given initual condition $v_C(0) = V_0$, while setting $t = 0$:

$$V_0 = V_S + Ae^0$$
$$V_0 = V_S + A$$
$$A = V_0 - V_S$$

We've thus found the full solution:

$$v_C(t) = V_S + (V_0 - V_S)e^{-\frac{t}{RC}}$$

4.2 Parallel RC circuit with a current source



To save time (and space): the differential equation we end up with is exactly the same as for the series circuit above, with the sole difference that IR replaces V_S, where I is the current source drive current, and R is the parallel resistor's resistance.
Since the resulting equation

$$RC\frac{dv_C}{dt} + v_C = IR$$

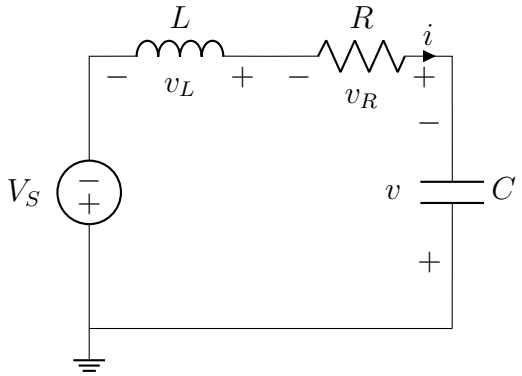
is of the same form as the one for the series circuit, the solution is also the same:

$$v_C(t) = IR + (V_0 - IR)e^{-\frac{t}{RC}}$$

Chapter 5

Second-order circuits

5.1 Series RLC circuits



We know that the current i relates to the capacitor voltage:

$$i = C\frac{dv}{dt}$$

Since this is a series circuit, that current goes through all elements.
By KVL, we can add up the voltage drops around the loop, with the proper sign. I'll go clockwise, and start in the bottom left:

$$-V_S + v_L + v_R + v = 0$$

If we solve for V_S:

$$V_S = v_L + v_R + v$$

Now, let's use the element laws for the inductor and resistor, and substitute them into the above:

$$V_S = L\frac{di(t)}{dt} + Ri(t) + v(t)$$

We know that $i(t) = C\frac{dv}{dt}$, so let's substitute that back in. Let's also differentiate, in case of the inductor, to reduce the mess of nested differentiation operators:

$$V_S = LC\frac{d^2v(t)}{dt^2} + RC\frac{dv}{dt} + v(t)$$

There we go; we now have a second-order, linear, constant coefficient ordinary differential equation.[1]

We'll use the same method to solve it as we did for the first-order ones, namely the method of particular and homogeneous solutions. First we first the particular solution (*any* solution that makes the equation true), then the homogeneous solution (a solution to the equation with the input drive V_S set to 0), and finally we add the two together to find the total solution.

As for the particular solution, just as in the first-order case, we try $v(t) = V_S$ and see that the two derivatives both go to zero (as V_S is a constant), and so we end up with

$$V_S = V_S$$

which tells us that V_S indeed is a particular solution. On to the homogeneous one, i.e. the solution to

$$LC\frac{d^2v(t)}{dt^2} + RC\frac{dv}{dt} + v(t) = 0$$

Again, like in the first-order case, we try the form

$$v(t) = Ae^{st}$$

where A and s are constants we'll have to find later. So, we substitute Ae^{st} for $v(t)$, and differentiate where necessary, which gives us

$$LCAs^2e^{st} + RCAse^{st} + Ae^{st} = 0$$

We can cancel out a ton of the above, by dividing both sides by Ae^{st}:

$$LC\cancel{A}s^2\cancel{e^{st}} + RC\cancel{A}s\cancel{e^{st}} + \cancel{Ae^{st}} = 0$$

What remains is

$$LCs^2 + RCs + 1 = 0$$

Divide by LC throughout:

$$s^2 + \frac{R}{L}s + \frac{1}{LC} = 0$$

We've thus arrived at the *characteristic equation*, which describes most details of the circuit's behavior. There's a canonical form to write such an equation:

$$s^2 + 2\alpha s + {\omega_0}^2 = 0$$

So, in the case of the series RLC circuit, the values of α and ω_0 would have to be

$$\alpha = \frac{R}{2L}$$

and

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

[1] I just love these overly verbose classifications.

α relates to the damping factor of the circuit (sometimes the damping factor, $\zeta = \frac{\alpha}{\omega_0}$, is used) - that is, the ringing second-order circuits can produce will decay faster for a larger value of α. Meanwhile, ω_0 is the *undamped resonance frequency* of the circuit (in radians/second), the same as in undriven LC circuits. However, this might not be the frequency of interest in a RLC circuit - more on that later.

Let's get back to solving the differential equation. Remember, we only found the homogeneous solution - we still haven't figured out the values of s or A.
The next step would be to find the roots of the characteristic equation:

$$s^2 + 2\alpha s + {\omega_0}^2 + 0$$

The quadratic formula will work nicely on this. The two roots are:

$$s_1 = -\alpha + \sqrt{\alpha^2 - {\omega_0}^2}$$
$$s_2 = -\alpha - \sqrt{\alpha^2 - {\omega_0}^2}$$

The full solution to the homogeneous solution will then be

$$v_H = A_1 e^{s_1 t} + A_2 e^{s_2 t}$$

However, we've now found s_1 and s_2, so we can fill them in:

$$v_H = A_1 e^{(-\alpha + \sqrt{\alpha^2 - {\omega_0}^2})t} + A_2 e^{(-\alpha - \sqrt{\alpha^2 - {\omega_0}^2})t}$$

Then, the total solution (prior to finding A_1 and A_2, and also prior to simplifying) is:

$$v(t) = V_S + A_1 e^{(-\alpha + \sqrt{\alpha^2 - {\omega_0}^2})t} + A_2 e^{(-\alpha - \sqrt{\alpha^2 - {\omega_0}^2})t}$$

After some magic[2] and making the definition $\omega_d = \sqrt{{\omega_0}^2 - \alpha^2}$, the total solution ends up looking like this, for the case $\omega_0 > \alpha$, i.e. the underdamped case:

$$v(t) = V_I + K_1 e^{-\alpha t} \cos \omega_d t + K_2 e^{-\alpha t} \sin \omega_d t$$

Evaluating the above at t = 0 with the initial condition $v(0) = 0$ gives us:

$$0 = V_I + K_1$$
$$K_1 = -VI$$

Evaluating at t = 0 with the other initial condition, $i(0) = C\frac{dv}{dt} = 0$, gives, after the differentiation and substitution for $t = 0$:

$$0 = -K_1 \alpha + K_2 \omega_d$$

We know that $K_1 = -V_I$:

$$0 = V_I \alpha + K_2 \omega_d$$
$$K_2 \omega_d = -V_I \alpha$$
$$K_2 = -\frac{V_I \alpha}{\omega_d}$$

Thus, finally, the full equation that governs the capacitor voltage of the series RLC circuit, in the underdamped case ($\omega_0 > \alpha$) is:

[2]Sorry, but I don't quite follow the entire process myself yet. Since this part was cut out from the lectures in order to save time (after 14 videos going through this process so far), I will do the same.

$$v(t) = V_I - V_I e^{-\alpha t} \cos \omega_d t - \frac{V_I \alpha}{\omega_d} e^{-\alpha t} \sin \omega_d t$$

Using the scaled sum of sines trig identity, we can turn this cosine minus sine business into a single cosine, and end up with:

$$v(t) = V_I - V_I \frac{\omega_0}{\omega_d} e^{-\alpha t} \cos\left(\omega_d t - \arctan\left(\frac{\alpha}{\omega_d}\right)\right)$$

While the equation is rather involved, it's still very clear that the two main features are a cosine multiplied by a decaying exponential, which will give us exactly the kind of waveform we see in a damped second-order system.
However, another extraordinary detail hides in the equation. Note how the capacitor voltage is V_I minus V_I times a bunch of things that is always positive... except for the cosine. In fact, if V_I is 5 volts, this circuit will oscillate between 0 and 10 volts - yes, the capacitor voltage will be twice as high as the input step! Rather incredible.

5.1.1 Summary

Since this section was by *far* the longest of this document so far, I thought a summary of the relevant formulas could be useful. Remember that most of these only apply to series RLC circuits, though some definitions (such as ω_d) are universal.

Where v(t) denotes the capacitor node voltage in the circuit shown (far) above - also, remember that this is for the *underdamped* case, i.e. $\omega_0 > \alpha$:

$$v(t) = V_I - V_I \frac{\omega_0}{\omega_d} e^{-\alpha t} \cos\left(\omega_d t - \arctan\left(\frac{\alpha}{\omega_d}\right)\right)$$

where

$$\alpha = \frac{R}{2L} \text{ rad/s}$$

$$\omega_0 = \frac{1}{\sqrt{LC}} \text{ rad/s}$$

$$\omega_d = \sqrt{{\omega_0}^2 - \alpha^2} \text{ rad/s}$$

Other formulas that may be useful are:

$$Q = \frac{\omega_0}{2\alpha} \text{ (dimensionless)}$$

$$\text{Natural period} = \frac{2\pi}{\omega_0} \text{ seconds}$$

$$\text{Natural frequency} = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{LC}} \text{ Hz}$$

The quality factor, Q, can (for now) be thought of as the approximate number of oscillations that will occur before the ringing dies out due to the damping, though the actual meaning is more precisely defined.

One thing that is important to note, and that in truth should have been brought up in more detail, is that for the special case where $R = 0$, the circuit becomes a pure LC circuit, which will (in theory) oscillate forever. In practice, of course, parasitic resistances will make sure that doesn't happen - unless the circuit is superconducting.
Why does it oscillate forever? Well, it's rather intuitive, at least if you accept that it will oscillate at all: it oscillates until the energy stored in the circuit is small enough that we see the voltages in the circuit to

be constants. The only way energy stored is reduced at all is by resistances; without them, it rings forever. Mathematically, this is easy to see from the above equations: $Q = \frac{\omega_0}{2\alpha} \to \infty$ as $2\alpha = \frac{R}{L} \to 0$.

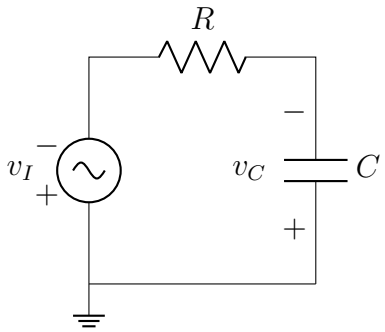
Chapter 6

Sinusoidal Steady State, Impedance and Filters

6.1 Sinusoidal Steady State

Instead of showing the entire, rather long path from a drawn circuit to the impedance model, I will take considerable shortcuts by simply skipping some parts. The reason is that I don't find it that important to remember fully; any good book on the subject should show it, though. I'll try to cover what *is* important to remember, however!

Thus, the beginning of this chapter might be a bit hard to follow. However, this isn't an electronics book (which I would be very underqualified to write) - I assume that all readers already know most of this, and just wanted a little refresher.
We'll try to analyze this circuit, for sinusoidal input:



where v_I is a sinusoidal input, e.g. $V_I \cos(\omega t)$.
So, let's get started. We write down the differential equations governing the dynamics of this circuit:

$$RC\frac{dv_C(t)}{dt} + v_C(t) = V_I \cos(\omega t)$$

If this were anything like the time-domain RC circuit analysis for step inputs, we'd now go and find particular and homogeneous solutions for the circuit, add them, and solve for constants using initial conditions. However, this would not be a very nice analysis due to the massive amount of trigonometry that would turn out to be involved due to the sinusoidal drive. Instead, we'll go down a different path, and see that doing so brings us some rather massive advantages in simplicity - especially later on, when we've derived and understood the impedance model, that removes the need to use differential equations for many kinds of analyses!

So, instead of analysing the circuit for the input signal $V_I \cos(\omega t)$, we'll analyse it for the input $V_i e^{st}$.
Our first analysis step is to attempt to find a particular solution that works with the given differential equation, which we've now modified (by choosing a different input) to be:

$$RC\frac{dv_C(t)}{dt} + v_C(t) = V_i e^{st}$$

Let's try a solution of the form $V_p e^{st}$, and substitute that in:

$$RC\frac{dV_p e^{st}}{dt} + V_p e^{st} = V_i e^{st}$$

Let's carry out the differentiation:

$$RCV_p s e^{st} + V_p e^{st} = V_i e^{st}$$

Let's cancel out the e^{st} terms that are common to all terms:

$$RCV_p s\cancel{e^{st}} + V_p\cancel{e^{st}} = V_i\cancel{e^{st}}$$

$$RCV_p s + V_p = V_i$$

Rearrange, then solve for V_p:

$$V_p(RCs + 1) = V_i$$
$$V_p = \frac{V_i}{1 + sRC}$$

However, remember that we crossed out all the e^{st} terms. The full particular solution is

$$V_p e^{st} = \frac{V_i}{1 + sRC} e^{st}$$

Now, let's assume that $s = j\omega$, where j is the imaginary unit $j = \sqrt{-1}$. We do the substitution, and get the particular solution

$$\text{Particular solution } = \frac{V_i}{1 + j\omega RC} e^{j\omega t}$$

Note that because this chapter deals with sinusoidal *steady state* analysis, we will ignore the homogeneous solution as it relates only to transients, i.e. what happens in the circuit *before* it reaches steady state.
Now we have a complex number. $V_p = \frac{V_i}{1 + j\omega RC}$ is the *complex amplitude*, which is an extremely useful variable. We'll see later that it can describe the entire frequency-domain behavior of the circuit to sinusoidal input, *and* most time-domain behavior as well.

However, remember that our goal was to find the response to sinusoidal input, not exponential input as we've actually done! What is the connection?
The answer lies in Euler's formula, which states

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

which can be proved e.g. via Taylor series.
Thus, we can take the real part of the exponential input, to yield the input we sought after:

$$V_i \cos(\omega t) = Re\left[V_i e^{j\omega t}\right]$$

where Re[z] gives the real part of the complex number z, i.e. $Re[x + jy] = x$.
Thus, we can use a "inverse superposition" argument: for linear circuits, the response to the input $Re[z]$ should be of the form $Re[f(z)]$. That is, the real part of the output should be a function of the real part of the input.

6.1.1 Time-domain analysis

I promised earlier that V_p could describe the time domain behaviour of the circuit. Let's see how! Also note that we will soon see a much faster and easier method of writing V_p down, than to work with the differential equations for a circuit.

The relation between V_p and the time-domain behavior of the circuit is:

$$v_C(t) = |V_p| \cos(\omega t + \angle V_p)$$

Note that we derived the time-domain behavior for the circuit simply by taking the magnitude and angle of the complex amplitude V_p. However, the above equation only specifies the relation, without using circuit values (i.e. R, C and V_i). Let's calculate the magnitude and phase angles.

$$|V_p| = \left|\frac{Vi}{1 + j\omega RC}\right| = \frac{|V_i|}{|1 + j\omega RC|}$$

The magnitude of V_i is simply the absolute value, and so we'll leave it as-is and assume it is positive.
The denominator is a complex number, and so we calculate the magnitude as always with complex numbers:

$$|a + jb| = \sqrt{a^2 + b^2}$$
$$|1 + j\omega RC| = \sqrt{1 + (\omega RC)^2}$$

Combining the two answers:

$$|V_p| = \frac{Vi}{\sqrt{1 + (\omega RC)^2}}$$

Thus step one is complete. Next, let's find the phase angle.
The angle of divided complex numbers subtracts (while magnitude divides). The angle of V_i, being fully real, is 0. (The angle of a fully imaginary number is $\frac{\pi}{2}$, since the imaginary axis is the up/down axis, and $\frac{\pi}{2}$ is the number of radians to rotate from the real (x) to the imaginary (y) axis.)

$$\begin{aligned}\angle V_p &= \angle V_i - \angle(1 + j\omega RC) \\ &= 0 - \arctan(1, \omega RC) \\ &= -\arctan(1, \omega RC)\end{aligned}$$

... where arctan is the two-argument arctan function. $\arctan(x, y)$ is equal to $\arctan(\frac{y}{x})$ for some inputs, but not all, which is why the two-argument one (also known as "atan2" in many programming languages) is used; the regular form can't differentiate between quadrants uniquely.

Thus, combining the time-domain equation with the now-known values of $|V_p|$ and $\angle V_p$:

$$v_C(t) = |V_p| \cos(\omega t + \angle V_p)$$
$$v_C(t) = \frac{Vi}{\sqrt{1 + (\omega RC)^2}} \cos(\omega t - \arctan(1, \omega RC))$$

As for the frequency response, we can also calculate and plot the magnitude of the *transfer function*, i.e. the ratio between the capacitor voltage and the input voltage.

$$\left|\frac{V_p}{V_i}\right| = \left|\frac{1}{1 + j\omega RC}\right| = \frac{1}{\sqrt{1 + (\omega RC)^2}}$$

... and the same goes for the phase of the transfer function:

$$\angle\frac{V_p}{V_i} = \angle\left(\frac{1}{1+j\omega RC}\right) = \angle 1 - \angle(1+j\omega RC) = 0 - arctan(1,\omega RC) = -arctan(1,\omega RC)$$

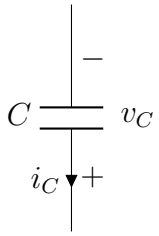
Now that we know how to find the magnitude and phase of the complex amplitude rather easily (compared to solving the differential equations!), let's see if we can find an easier way to writing the expression down in the first place, so that we can solve linear circuits with resistors, capacitors and inductors in no-time.

6.2 The Impedance Model

For the resistor, Ohm's law applies equally well to the complex amplitudes of voltages and currents as it does to the usual inputs:

$$V_r = I_r \cdot R$$

However, the interesting things turn up when we look at the capacitor and inductor. Let's consider a lone capacitor and its relation between current and voltage.



We know from earlier that

$$i_C = C\frac{dv_C}{dt}$$

But what happens when we use complex amplitudes instead? Say the current is of the form $I_c e^{st}$ and the voltage across it of the form $V_c e^{st}$ - what then? Well, let's see. We start by substituting the values:

$$I_c e^{st} = C\frac{dV_c e^{st}}{dt}$$

Perform the differentiation:

$$I_c e^{st} = CV_c s e^{st}$$

Cancel out the e^{st} terms from both sides, and solve for V_c:

$$I_c \cancel{e^{st}} = CV_c s \cancel{e^{st}}$$

$$I_c = CV_c s$$

$$V_c = \frac{1}{sC} I_c$$

Hmm, it looks like they have a very simple relation indeed! To simplify even further, let's set $Z_c = \frac{1}{sC}$:

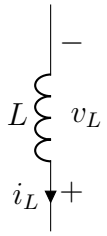
$$V_c = Z_c I_c$$

Incredible! The voltage is proportional to the current through this Z_c variable, just like a resistor's voltage drop is $V = IR$!

Z_c is known as the capacitor's *impedance*, and is frequency dependant: remember that $Z_c = \frac{1}{sC}$, and $s = j\omega$, where ω denotes the angular frequency of the input sinusoid. Therefore, we see that the higher

the frequency, the lower the impedance.
What this means is that the capacitor will have high impedance - which is similar (but not equal!) to high resistance - to low-frequency signals, to the point where it completely blocks DC ($\omega = 0$), but has a low impedance for high-frequency signals. As the frequency increases, the capacitor impedance decreases, and as $\omega \to \infty$, the impedance goes to zero and the capacitor acts as a short circuit.

Let's see how an inductor deals with this.



Again, we know that

$$v_L = L\frac{di_L}{dt}$$

Let's see what happens for complex input, where $v_L = V_l e^{st}$ and $i_L = I_l e^{st}$:

$$V_l e^{st} = L\frac{dI_l e^{st}}{dt}$$

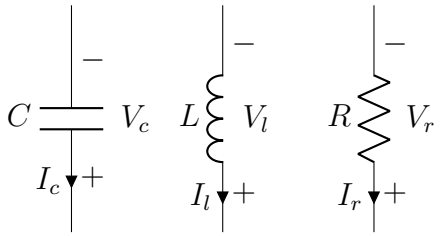
Same as before; we perform the differentiation, cancel out the e^{st} terms, and solve for V_l:

$$V_l e^{st} = L\frac{dI_l e^{st}}{dt}$$
$$V_l e^{st} = LI_l s e^{st}$$
$$V_l \cancel{e^{st}} = LI_l s \cancel{e^{st}}$$
$$V_l = LI_l s$$
$$V_l = I_l \cdot sL$$

Again, the voltage is related to the current through a simple algebraic expression! The impedance of the inductor is $Z_L = sL$, where (again) $s = j\omega$.

$$V_l = Z_L I_l$$

This is an incredible powerful technique, as we'll soon see. We can now write down the frequency (and time) behaviour of circuits by inspection, without writing a single differential equation!
Impedances not only look like resistances (impedance times current equals voltage), but can be used in the same way, to create things like voltage dividers using capacitors and inductors! This is the basis of *filters*, which we'll look at soon.
So, in summary, the relevant models are:



Where the relations between voltage amplitudes and current amplitudes are given as

$$V_r = I_r Z_R$$

$$V_c = I_c Z_L$$
$$V_l = I_l Z_L$$

... and the impedences are given as

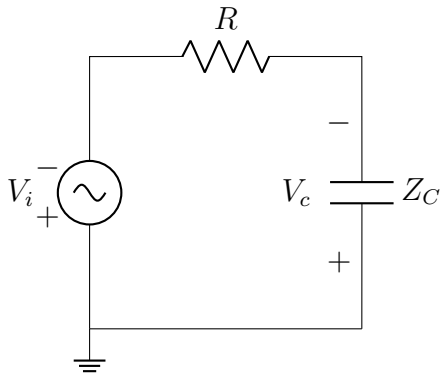
$$Z_R = R$$
$$Z_C = \frac{1}{j\omega C}$$
$$Z_L = j\omega L$$

These impedances can be used exactly as resistances. For example, in a series RC circuit, the current through the circuit is given by

$$\frac{V_i}{R + Z_C}$$

just as you would expect.

6.2.1 RC circuit example



Let's analyze the above circuit, to find the voltage across the capacitor. The impedance is $Z_C = \frac{1}{sC}$, and we can use that to create a voltage divider relationship:

$$V_c = V_i \cdot \frac{Z_C}{Z_C + R} = V_i \cdot \frac{\frac{1}{sC}}{\frac{1}{sC} + R}$$

Let's multiply out by $\frac{sC}{sC}$ to clean things up:

$$V_c = V_i \cdot \frac{1}{1 + sRC}$$

This is another HUGE result; look at the denominator: what he have there is the characteristic equation for the series RC circuit! We can use that to find everything we want to know about the time-domain behavior of the circuit, such as the natural frequency, bandwidth, damping factor and more - without a single differential equation; not to mention that the same technique works in more complex circuits, such as a series-parallel combination of resistors, capacitors and inductors!

6.2.2 Impedance Model Summary

Before trying a few more complex circuits, let's summarize the method we'll use, and clarify the meaning of a few variables.

1. Replace sinusoidal sources by their complex amplitudes. For example, $v_I \cos(\omega t)$ would simply become V_i, which represents a complex amplitude (capital letter, small subscript).

2. Replace the elements with their impedance models; resistors don't change, capacitors turn in to $Z_C = \frac{1}{sC}$ "resistors" and inductors into $Z_L = sL$ "resistors", where $s = j\omega$.

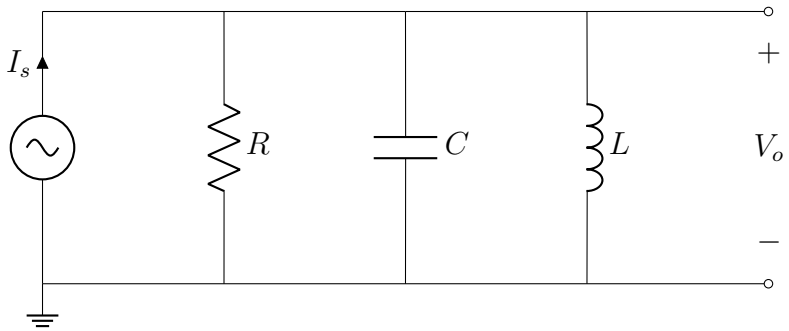
3. Calculate the complex amplitudes for the nodes that matter, using the normal circuit analysis methods e.g. the node method, Thevenin equivalents, etc, treating impedances exactly as resistances.

4. If necessary, find the time-domain behavior from the complex amplitudes: if you have the complex amplitude V_a, the time-domain behavior of that amplitude is described by

$$v_A(t) = |V_a| \cos(\omega t + \angle V_a)$$

6.2.3 Finding the characteristic equation

What we generally want to do here is to write the expression as a proper polynomial: no $\frac{1}{s}$ terms are allowed, and the denominator must be of a higher degree than the numerator. Note that there may be s terms in the numerator, which are then to be ignored, if the polynomial is in proper form.
As an example, let's look at the *parallel* RLC circuit, to avoid overfocusing on the series one.



Note that the source is a sinusoidal *current source*.
The impedance Z seen by the current source is given by

$$Z = R \,||\, \frac{1}{sC} \,||\, sL = \frac{1}{\frac{1}{R} + \frac{1}{sL} + sC}$$

Let's start off by multiplying everything by sL, to get rid of the improper s:

$$Z = \frac{sL}{\frac{1}{R}sL + 1 + s^2LC}$$

Next, divide by LC to get s^2 by itself (as usual, we want the form $s^2 + 2\alpha s + {\omega_0}^2$):

$$Z = \frac{\frac{1}{C}s}{\frac{1}{RC}s + \frac{1}{LC} + s^2}$$

$$Z = \frac{\frac{1}{C}s}{s^2 + \frac{1}{RC}s + \frac{1}{LC}}$$

Are we done? Well, the s^2 is on its own, we have a term multiplying s, and a term independant of s. Also, we have nothing on the form $\frac{1}{s}$[1], and the denominator is of higher degree than the numerator. We're done!

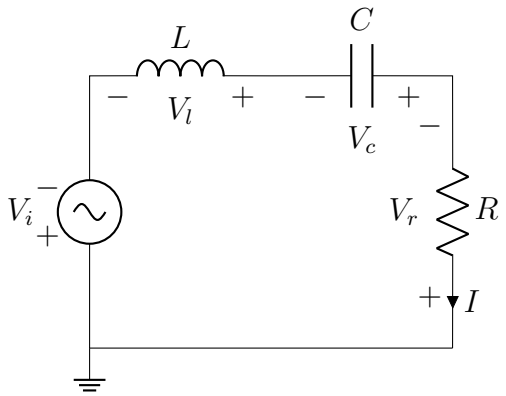
Thus, comparing the denominator with the canonical form $s^2 + 2\alpha s + {\omega_0}^2$, we see that, for the parallel RLC circuit:

$$\alpha = \frac{1}{2RC} \text{ rad/s}$$

[1] Except the entire denominator, of course.

$$\omega_0 = \frac{1}{\sqrt{LC}} \text{ rad/s}$$
$$f_0 = \frac{1}{2\pi\sqrt{LC}} \text{ Hz}$$
$$\Delta\omega = 2\alpha = \frac{1}{RC}$$
$$Q = \frac{\omega_0}{2\alpha} = \frac{RC}{\sqrt{LC}}$$

6.2.4 Series RLC circuits



Say we want to find the voltage amplitude across the resistor for this circuit. Doing so is very easy; we use a voltage divider again:

$$V_r = V_i \cdot \frac{R}{R + Z_C + Z_L} = V_i \cdot \frac{R}{R + \frac{1}{sC} + Ls}$$

Easy! We can clean it up a bit by multiplying by $\frac{s}{L}$ (why did I pick $\frac{s}{L}$? Because I know what will happen.):

$$V_r = V_i \cdot \frac{\frac{R}{L}s}{\frac{R}{L}s + \frac{1}{LC} + s^2}$$

Note that we have *again* found the characteristic equation of this circuit (in the denominator), again without writing a differential equation!
Also, if we want the time domain behavior, we substitute $s = j\omega$ into the first equation above, rationalize the denominator by multiplying top and bottom by the complex conjugate of the bottom, and take the magnitude and phase:

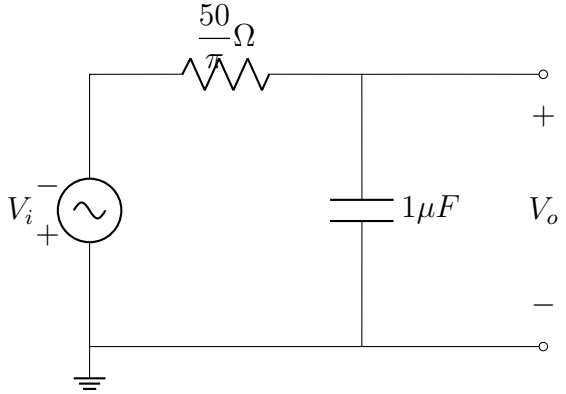
$$v_R(t) = |V_r| \cos(\omega t + \angle V_r) = \left| V_i \cdot \frac{j\omega RC}{1 - LC\omega^2 + j\omega RC} \right| \cos\left(\omega t + \angle \left(V_i \cdot \frac{j\omega RC}{1 - LC\omega^2 + j\omega RC} \right)\right)$$
$$= \frac{V_i \omega RC}{\sqrt{(1 - LC\omega^2)^2 + (\omega RC)^2}} \cos\left(\omega t + \left(\frac{\pi}{2} - \arctan(1 - LC\omega^2, \omega RC)\right)\right)$$

... where the magnitude is given by the magnitude of $jV_i\omega RC$ divided by the magnitude of the denominator, a complex number: such a magnitude is given by $\sqrt{a^2 + b^2}$ where a is the real part and b the imaginary part.
The angle is given by the angle of the numerator *minus* (not divided by) the denominator; the angle of the numerator is $\frac{\pi}{2}$ because $jV_i\omega RC$ is not complex, but fully imaginary, and thus located 90 degrees ($\frac{\pi}{2}$ radians) rotated from the origin. The angle of the denominator is given by $\arctan(a, b)$ where a and b are the real and imaginary parts, respectively.

6.3 Filters

Filters are a great application of the impedance model. Filters are circuits (some simple, some complex) that attenuate certain frequency bands. For example, a low-pass filter removes high frequencies, but lets low frequencies through.

Here is a simple RC low-pass filter (the part values might be unrealistic, but they make for round numbers in the results):



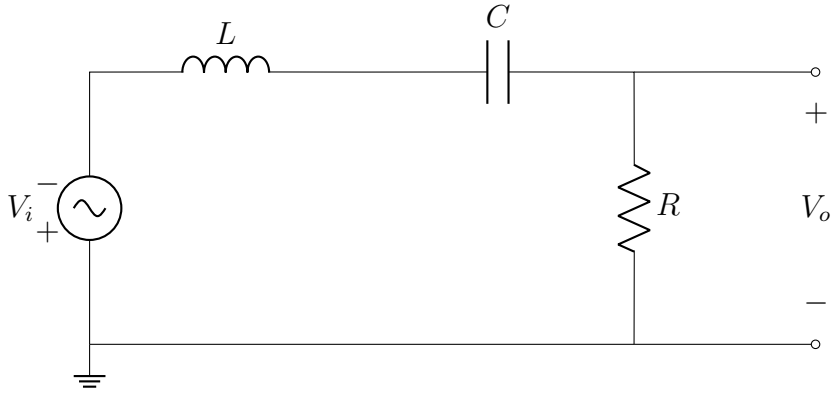
Filter characteristics are often shown as a pair of plots, known as Bode plots. One shows the magnitude as a function of frequency, and the other the *phase* as a function of frequency.
For example, see these Bode plots of a low-pass filter:

The filter above has a *break frequency* (also known as corner frequency, -3 dB point or $\frac{1}{\sqrt{2}}$ point) of 10 kHz, which means frequencies higher than 10 kHz will get attenuated. As you can see, however, the attenuation starts off lower than 10 kHz. The break frequency is the frequency where the magnitude drops to 3 dB below unity (1x the input), which equals $\frac{1}{\sqrt{2}}$ times unity, here shown by a pair of dotted lines. The break frequency is also where the phase shift of the signal is exactly -45 degrees (for this circuit).

There are 4 main types of filters: low-pass, high-pass, band-pass and band-stop. The names explain them quite well; low-pass filters let low frequencies through, but blocks higher ones, while high-pass is the exact opposite. Band-pass lets a *band* of frequencies through, but blocks both lower and higher frequencies. Band-stop are opposite of band-pass; they let everything *except* a band of frequencies through.

6.3.1 Series RLC band-pass filter

Since producing good-looking Bode plots for this document takes a while with my current process, I'll only make a few, by focusing on a RLC band-pass filter.



First, let's think about how this filter works intuitively.
The inductor will let low frequencies through, but block high frequencies. Thus, at high frequencies, most

of the circuit's voltage drop will be across the inductor, so the filter blocks them.
The capacitor will let the high frequencies pass with low impedance, but instead block low frequencies. However, there is a band of frequencies somewhere in the middle that they both let through, so this is a *band-pass* filter.

Since this is a series circuit, we can yet again use a simple voltage divider relation to write the transfer function $\frac{V_o}{V_i}$:

$$\frac{V_o}{V_i} = \frac{R}{R + j\omega L + \frac{1}{j\omega C}}$$

(Side note: the transfer function, also system function, is often denoted H(s) or H(jω).)
Let's clean it up a bit by multiplying through by $j\omega C$.

$$\frac{V_o}{V_i} = \frac{R}{R + j\omega L + \frac{1}{j\omega C}} \cdot \frac{j\omega C}{j\omega C}$$

$$\frac{V_o}{V_i} = \frac{j\omega RC}{1 - \omega^2 LC + j\omega RC}$$

We can now calculate the magnitude and phase of the transfer function, and plot it.

$$\left|\frac{V_o}{V_i}\right| = \frac{\omega RC}{\sqrt{(1 - \omega^2 LC)^2 + (\omega RC)^2}}$$

$$\angle\frac{V_o}{V_i} = \frac{\pi}{2} - \arctan\left(1 - \omega^2 LC, \omega RC\right)$$

Here are the Bode plots for this circuit for $L = 1\mu H$, $C = 1\mu F$ and $R = 20\Omega$:

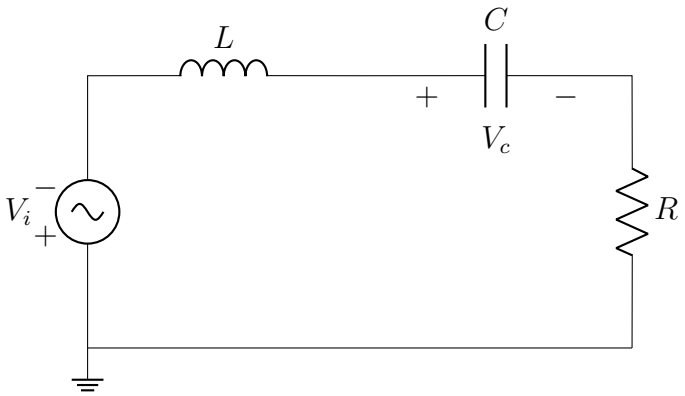
The dotted red line is at $\frac{1}{\sqrt{2}}$. The dotted vertical lines are, from left to right, the low-end break frequency f_1, the resonant frequency $f_0 = \frac{1}{2\pi\sqrt{LC}}$ and the high-end break frequency f_2.
Note that the f notation instead of ω indicates that the frequencies are in hertz, rather than rad/s.

The dotted red lines are at 45, 0 and -45 degrees. The dotted vertical lines are, from left to right, the low-end break frequency f_1, the resonant frequency $f_0 = \frac{1}{2\pi\sqrt{LC}}$ and the high-end break frequency f_2.

These plots show two break frequencies, at roughly 7.94 kHz and 3.19 MHz, respectively. (I have no idea where such a filter could be useful, however; the circuit values could've been better chosen for this example!)
The break frequencies can be found by setting the magnitude equation equal to $\frac{1}{\sqrt{2}}$ and solving for ω (keep in mind that ω is in rad/s; the plots are in hertz).

6.3.2 Q and "peakiness"

Something interesting happens if we observe the capacitor voltage of a high-Q RLC circuit near the resonant frequency.



The above Bode plot shows the magnitude of the capacitor voltage/input voltage ratio for the values $R = 7.5\Omega$, $L = 0.75$ mH, $c = 400$ nF, which causes $f_0 = \dfrac{1}{2\pi\sqrt{0.75 \cdot 10^{-3} \cdot 400 \cdot 10^{-9}}} \approx 9.19$ kHz and $Q = \dfrac{0.75 \cdot 10^{-3}}{7.5\sqrt{0.75 \cdot 10^{-3} \cdot 400 \cdot 10^{-9}}} \approx 5.7735$.
The dotted red lines are at $\dfrac{1}{\sqrt{2}}$ and Q, while the dotted grey line is at $f_0 \approx 9.19$ kHz.

This phenomenon is actually not as strange as it may appear. Q is the ratio of stored energy in a circuit to the energy dissipated per cycle, so at high Q values, more energy is added by the source than is dissipated by the resistor, so the voltage builds up over time, to be greater than the input amplitude (specifically Q times greater).
Let's derive this mathematically. Again, this was for the voltage over the capacitor, so we need to calculate the magnitude of that. Yet again we used a voltage divider relation:

$$\frac{V_c}{V_i} = \frac{\frac{1}{sC}}{\frac{1}{sC} + sL + R}$$

Multiply across by sC to clean up:

$$\frac{V_c}{V_i} = \frac{1}{1 + s^2LC + sRC}$$

Let's divide through by LC, too:

$$\frac{V_c}{V_i} = \frac{\frac{1}{LC}}{s^2 + \frac{R}{L}s + \frac{1}{LC}}$$

There's the characteristic equation again. We know (from memory, but also from comparing the above to $s^2 + 2\alpha s + {\omega_0}^2$) that $\dfrac{1}{LC} = {\omega_0}^2$, and the same goes for $2\alpha = \dfrac{R}{L}$.

$$\frac{V_c}{V_i} = \frac{{\omega_0}^2}{s^2 + 2\alpha s + {\omega_0}^2}$$

$s = j\omega$; however, let's evaluate the above expression at the resonant frequency, i.e. $\omega = \omega_0$, so we make the substitution for $s = j\omega_0$:

$$\frac{V_c}{V_i} = \frac{{\omega_0}^2}{(j\omega_0)^2 + 2\alpha j\omega_0 + {\omega_0}^2}$$

j^2 is -1, so:

$$\frac{V_c}{V_i} = \frac{{\omega_0}^2}{-{\omega_0}^2 + 2\alpha j\omega_0 + {\omega_0}^2}$$

Each term has an ω_0, so we can cancel them out.

$$\frac{V_c}{V_i} = \frac{\omega_0}{-\omega_0 + j2\alpha + \omega_0}$$

Also, note that we have $-\omega_0 + \omega_0$ remaining, so we remove those as well:

$$\frac{V_c}{V_i} = \frac{\omega_0}{j2\alpha}$$

We then take the magnitude of this expression, that (remember what we're doing!) shows the ratio of the capacitor voltage to the input, at the circuit's resonant frequency. The only thing that happens is that j disappears, as its magnitude is 1:

$$\left|\frac{V_c}{V_i}\right| = \frac{\omega_0}{2\alpha} = Q$$

Yup! It turns out that the Bode plot is correct: the magnitude of the capacitor voltage at resonance is exactly Q times greater than V_i!

6.3.3 Selectivity and bandwidth

The selectivity of a filter is given by

$$\frac{\omega_0}{\Delta\omega}$$

where $\Delta\omega$ is the *bandwidth* of the filter:

$$\Delta\omega = \omega_2 - \omega_1$$

where ω_1 and ω_2 are the $\frac{1}{\sqrt{2}}$ points of the filter (see above).
For the band-pass filter in section 6.3.1 (just prior to the "peaky" high-Q low-pass filter):

$$\omega_0 = \frac{1}{2\pi\sqrt{LC}} = 159.2 \text{ kHz}$$
$$\omega_1 = 7.938 \text{ kHz}$$
$$\omega_2 = 3.191 \text{ MHz}$$
$$\Delta\omega = \omega_2 - \omega_1 = 3.183 \text{ MHz}$$
$$\frac{\omega_0}{\Delta\omega} \approx 0.05$$

Let's try to find $\frac{\omega_0}{\Delta\omega}$ in terms of circuit parameters, e.g. R, C, L and so on.
To begin with, we set the magnitude equation equal to $\frac{1}{\sqrt{2}}$. We solve it, and take the two positive solutions, ω_2 and ω_1, and subtract them. The result is the bandwidth, $\Delta\omega$.
Of course, we already know that for a series RLC circuit, $\omega_0 = \frac{1}{\sqrt{LC}}$.

$$\frac{\omega RC}{\sqrt{(1-\omega^2 LC)^2 + (\omega RC)^2}} = \frac{1}{\sqrt{2}}$$

The solutions ω_1 and ω_2 to the above are quite messy, but when subtracted, the end result is very simple:

$$\Delta\omega = \omega_2 - \omega_1 = \frac{R}{L}$$

Nice! Let's confirm that the previous result for $\Delta\omega$ matches this one:

$$\Delta\omega = \omega_2 - \omega_1 = 3.191 \text{ MHz} - 7.938 \text{ kHz} = \frac{20\Omega}{2\pi \cdot 1\mu H}$$

Note that we need to additionally divide by 2π to convert from rad/s to hertz. And indeed, the above is true, excepting for rounding errors.

However, there is an additional thing to note here. Remember that $2\alpha = \frac{R}{L}$ for a series RLC circuit! Thus the bandwidth $\Delta\omega$ is equal to 2α.

Further, remember the definition of Q for the series RLC circuit:

$$Q = \frac{\omega_0}{2\alpha}$$

We've just shown that $\Delta\omega = 2\alpha$, so the selectivity of a circuit is given by Q!

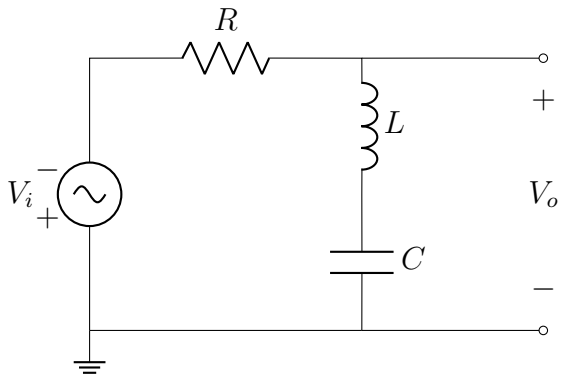
Or, in circuit parameters:

$$Q = \frac{L}{R\sqrt{LC}}$$

6.3.4 Designing filters

Instead of simply analyzing existing circuits, let's design a band-stop filter. Let's say we want it to stop frequencies in the range 1 - 10 kHz. In other words, we want the break frequencies to be at 1 and 10 kHz, respectively, which gives a bandwidth $\Delta\omega$ of 9 kHz.

Again, let's approach this intuitively. Clearly, a resistor combined with either a capacitor or inductor won't cut it, since such a circuit can only block either low or high frequencies (as the (ideal) resistor blocks all frequencies equally), not any sort of a combination. Thus we need either a RLC circuit, or some circuit consisting of multiple capacitors/multiple inductors. One solution would indeed be an RLC circuit, where we take the output over the LC pair rather than over the resistor as we did in the band-*pass* filter above:



In this circuit, the LC voltage will be high either when the frequency is low (the capacitor drops a lot of voltage) or high (the inductor drops a lot of voltage), but low at some frequency band in between. Great! Let's figure out the part values we want, and all that. First, we need to figure out the circuit's transfer function:

$$\frac{V_o}{V_i} = \frac{sL + \frac{1}{sC}}{R + sL + \frac{1}{sC}}$$

where $s = j\omega$. We multiply through by sC to get rid of the nested fractions:

$$\frac{V_o}{V_i} = \frac{s^2LC + 1}{1 + sRC + s^2LC}$$

Hmm, let's clean up that s^2 term in the denominator by dividing by LC throughout.

$$\frac{V_o}{V_i} = \frac{s^2 + \frac{1}{LC}}{\frac{1}{LC} + \frac{R}{L}s + s^2}$$

There's the characteristic equation for this circuit (in the denominator). If we match it up with the canonical form $s^2 + 2\alpha s + {\omega_0}^2$, we find that

$$2\alpha = \frac{R}{L}$$
$$\alpha = \frac{R}{2L}$$
$${\omega_0}^2 = \frac{1}{LC}$$
$$\omega_0 = \frac{1}{\sqrt{LC}}$$

This isn't really news, since we knew the values of α and ω_0 for the series RLC circuit already, but it's nice to know we can find them easily. Let's find out the values of Q and $\Delta\omega$, the bandwidth:

$$Q = \frac{w_0}{2\alpha} = \frac{\frac{1}{\sqrt{LC}}}{\frac{R}{L}} = \frac{L}{R\sqrt{LC}}$$
$$\Delta\omega = 2\alpha = \frac{R}{L}$$

We can use these values to find the part values we need. ω_0 is the *center frequency* of the filter, so if we calculate the center frequency we want and set that equal to ω_0, that should be a start. After that, we can set $\Delta\omega$ equal to the bandwidth we wanted.
Two equations, three unknowns, but to make our lives easier we can simply choose one value ourselves.

Let's calculate the center frequency. This isn't quite as simple as taking the arithmetic mean $\frac{f_1 + f_2}{2}$, which would give us $\frac{1000 + 10000}{2} = 5500$ Hz. Instead, we need to take the *geometric mean*, which is defined as $\sqrt[n]{a_1 \cdot a_2 \cdots a_{n-1} \cdot a_n}$ - that is, multiply the numbers together, and take the nth root of the product, where n is the count of numbers. So, the center frequency we want is computed as:

$$\omega_1 = f_1 \cdot 2\pi = 1000 \cdot 2\pi$$
$$\omega_2 = f_2 \cdot 2\pi = 10000 \cdot 2\pi$$
$$\omega_0 = \sqrt{w_1 \cdot w_2} = \sqrt{(1000 \cdot 2\pi)(10000 \cdot 2\pi)} = 2000\pi\sqrt{10}$$

Now that we have one of the two values we wanted, let's find the bandwidth we want. That one's easy - it's just the difference between the higher and the lower corner frequencies, in this case 10 kHz and 1 kHz, respectively.

$$\Delta\omega = \omega_2 - \omega_1 = 10000 \cdot 2\pi - 1000 \cdot 2\pi = 9000 \cdot 2\pi$$

Now all that remains is to solve the system of equations we get for these two, setting our desired ω_0 equal to the circuit equation, and doing the same for the bandwidth:

$$\frac{1}{\sqrt{LC}} = 2000\pi\sqrt{10}$$
$$\frac{R}{L} = 9000 \cdot 2\pi$$

If we solve these, we get

$$L = \frac{R}{18000\pi},\ C = \frac{9}{20000\pi R}$$

So, for $R = 680\Omega$ (a chosen value, not a calculated one), the values are roughly $L = 12.025$ mH and $C = 210.65$ nF.
Here are the Bode plots for the resulting filter:

As with the previous plots, the dotted red line is at $\frac{1}{\sqrt{2}}$ and the dotted grey lines are at the break frequencies, i.e. 1 and 10 kHz.
The Q of this filter is approximately 0.3514.

Note the discontinuity in the plot, due to us restricting the valid range of phase values between +90 and -90 degrees.
Without the discontinuity, the phase plot looks like this:

It looks like our filter works! However, it's not perfect; the inductor value is relatively high. We could get the same filter with other values, but the capacitor value would have to increase instead.
There are other ways to make band stop filters. In particular, op amps will make it easier.
All in all, however, it does pretty much work - not bad for a first-ever try in filter design, especially considering that low-pass/high-pass filters with only two components are quite a bit easier.

6.4 Summary of SSS, impedance and filters

This chapter dealt with the steady state workings of circuits with sinusoidal input. We learned that, by representing this as an exponential input, we can turn the differential equations governing these circuits into simple algebraic relations: *the impedance model*. In doing so, we not only got access to easy tools for frequency domain analysis of circuits, but also a few that also help us in the time domain, such as the equation for converting a complex amplitude, used in the impedance model, to a time domain representation of the circuit:

$$v_X(t) = |V_x| \cos(\omega t + \angle V_x)$$

The impedances (denoted by Z) for the three common passive circuit elements[2] are:

$$\text{Resistor: } Z_R = R$$

$$\text{Capacitor: } Z_C = \frac{1}{j\omega C}$$

$$\text{Inductor: } Z_L = j\omega L$$

The most important parameters for filters are:
(Where specific circuit parameters are used, they are for the series RLC circuit. The parameters themselves (α, $\Delta\omega$ etc.) apply to all circuits, however.)

$$\text{Resonant angular frequency: } \omega_0 = \frac{1}{\sqrt{LC}} \text{ rad/s}$$

$$\text{Resonant frequency: } f_0 = \frac{1}{2\pi\sqrt{LC}} \text{ Hz}$$

$$\text{Damping factor: } \alpha = \frac{R}{2L} \text{ rad/s}$$

$$\text{Bandwidth: } \Delta\omega = \omega_2 - \omega_1 = 2\alpha = \frac{R}{L} \text{ rad/s}$$

$$Q = \frac{\omega_0}{2\alpha} \text{ (dimensionless)}$$

Remember that Q can be thought of as multiple things. For a frequency-domain analysis, it is indicative of the how selective the filter is. High Q means high selectivity.
For time-domain analyses, Q indicates how long an underdamped system will ring; also, systems with $Q < \frac{1}{2}$ are overdamped and will not ring at all.

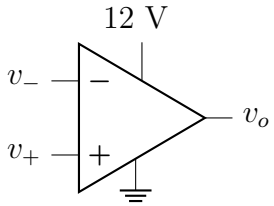
[2] There are four such basic elements; the fourth is the memristor, which as of this writing is still under development, with commercial release slated for 2013.

Chapter 7

Operational Amplifiers

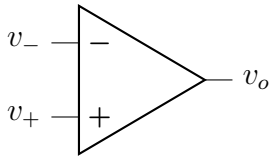
7.1 Intro

Operational Amplifiers, or *op amps*, are an amplifier abstraction; that is, we can use it without knowing about its insides.
The schematic symbol looks like this:



Above is a *single-supply* op amp, that can output voltages between 0 and 12 volts, with respect to ground. *Dual-supply* op amps are also common; in that case, you'd replace the ground above with -12 volts, and have the ground (which is then not directly connected to the op amp) where the two rails "meet", i.e. output range is from 12 volts above ground all the way down to 12 volts below ground.
That way, the op amp can output negative voltages (again with respect to the circuit ground), which is often required when dealing with signals such as audio, which are ground referenced and symmetric around the 0 volt axis.

The ideal op amp has infinite input impedance, and zero output impedance. Thus, no current flows into the input terminals; all current from the output terminal comes from the power terminals.
In many cases, we don't really care about showing the supply connection explicitly, and instead use the following symbol:



This document will use the latter symbol from now on.

Op amps use a *differential* input. That is to say, neither of the two input connections is necessarily ground, and the amplifier amplifies the *difference* between the two input ports. That is, ideally,

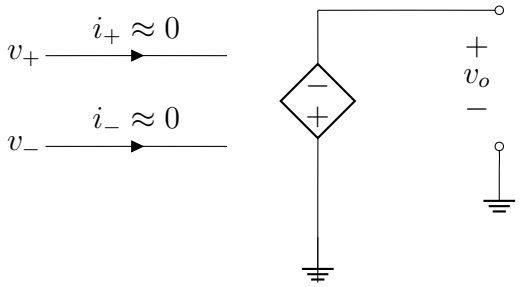
$$v_o = A(v_+ - v_-)$$

where A is the amplifier's voltage gain. In the ideal op amp model, $A \to \infty$. In real op amps, A is often on the order of 100,000 or more. Needless to say, most of the time, we don't want to amplify a signal that much - even a input voltage difference as tiny as 12μV would be enough to drive the output to the maximum 12 V ($12\mu V \cdot 10^6 = 12V$) for an amplifier with $A = 10^6$, which most certainly exist on the market.

So, to summarize what we've seen so far:

1. The op amp is a *differential* amplifier ...
2. ... that outputs a ground-referenced voltage.
3. Its gain is extremely high (ideally infinite), often on the order of 10000 to a million, or even higher. If we want less gain, we'll have to take care of that outside the abstract "amplifier box" that is the op amp.
4. Op amps can be powered by a single supply (a positive voltage and ground) or a dual supply (a positive and a negative voltage, with ground being in between the two). There are op amps made specifically for either of these two, though many works in either configuration.

To analyze circuits with ideal op amps, we model them as a VCVS, that is, a voltage-controlled voltage source. Again, no current flows into the input terminals, so we model them as open circuits.



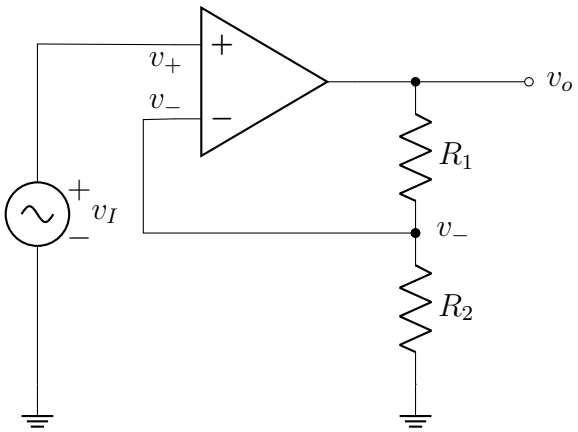
Since real-life op amps are powered by a power supply (rather than being ideal sources themselves), they can never output a voltage that is higher than that of the power supply. Thus, when the difference between the inputs is greater than $\frac{\text{Power supply voltage}}{\text{Gain}}$, the op amp *saturates*, and the output stays close to the power supply rail.
Here's a plot of an op amp's transfer function, with a gain of 10^6:

Note the extreme difference in scale between the axes: the v_i axis is in microvolts, while the v_o axis is in volts. The op amp behaves linearly as long as the gain multiplied by the input voltage difference is less than the power supply voltage (we call this the *active region*, or sometimes the linear region), but *saturates* whenever the voltage difference is greater.

Most op amps can't bring the output very close to either power supply rail; an op amp powered by a ± 12 V supply can usually swing between roughly -11 and +11 volts, or thereabouts. Therefore, an op amp will generally saturate before it the power supply voltage. Op amps which can output very close (usually within a few millivolts) of the power rails are called *rail-to-rail output* op amps. There are also *rail-to-rail input* and *rail-to-rail input/output* (RRIO) op amps.

7.2 Noninverting amplifier

Let's use an op amp to build a noninverting amplifier, one of the simpler (and perhaps more common) uses of an op amp. This configuration uses a resistive divider between the output and the v_- terminal to limit the gain to more usable values. This also has the huge advantage of making the op amp very stable, as opposed to its "raw" form, where the gain A can vary extremely much for pretty much no discernable reason.
It looks something like this:



As always, v_o is ground-referenced, though that will not be shown explicitly here or in future circuit diagrams.
Let's analyze this circuit. First, we know that the op amp input/output relation is, always,

$$v_o = A(v_+ - v_-)$$

Let's find the values of v_+ and v_- for this circuit.
v_+ is simply our input voltage v_I.
v_- is given by a simple voltage divider between v_o and ground, so

$$v_- = v_o \cdot \frac{R_2}{R_1 + R_2}$$

(Remember that v_- is an open circuit; the input impedance is infinite, and the connection there doesn't affect the voltage divider ratio.)

Let's substitute these values back into the first equation, and get:

$$v_o = A(v_I - v_o \cdot \frac{R_2}{R_1 + R_2})$$

$$v_o = Av_I - Av_o \cdot \frac{R_2}{R_1 + R_2}$$

We have v_o terms on both sides; let's collect all v_o terms on the left-hand side:

$$v_o + Av_o \cdot \frac{R_2}{R_1 + R_2} = Av_I$$

$$v_o(1 + A \cdot \frac{R_2}{R_1 + R_2}) = Av_I$$

Now we can simply divide both sides by the expression multiplying v_o and we're pretty much done:

$$v_o = \frac{Av_I}{1 + A \cdot \frac{R_2}{R_1+R_2}}$$

Note that if $A \cdot \dfrac{R_2}{R_1 + R_2} \gg 1$, we can ignore the 1 term without losing any noteworthy amount of precision. For example, if $A = 10^6$, $R_1 = 3000$ and $R_2 = 1000$, which gives a gain of 3.999984, the value we get using the approximation (ignoring the 1) is exactly 4. Such tiny errors don't really matter to us, so we can say that

$$v_o \approx \frac{Av_I}{A \cdot \frac{R_2}{R_1+R_2}}$$

Note that we can now cancel out the A terms!

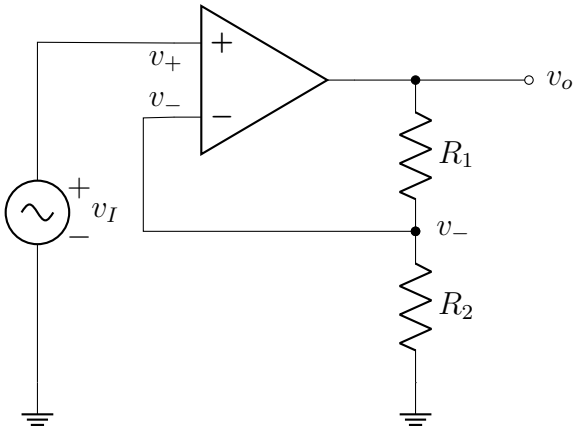
$$v_o \approx \frac{v_I}{\frac{R_2}{R_1+R_2}} = v_I \cdot \frac{R_1+R_2}{R_2}$$

Therefore, this configuration gives a voltage gain of almost exactly $\frac{R_1+R_2}{R_2} = 1 + \frac{R_1}{R_2}$, as long as we make sure that $A \cdot \frac{R_2}{R_1+R_2} \gg 1$ is true, by making sure we never make R_1 much, much bigger than R_2. However, even when R_1 is *1000 times* greater than R_2, the approximation works out quite well: for $A = 10^6$, $R_1 = 10^6\ \Omega$ and $R_2 = 1000\ \Omega$ the error is about 0.1%. Such a configuration would give a gain of roughly 1000.

So, this circuit works! It gives as a gain we can set via simple resistor values, and thus the op amp gain A is virtually irrelevant to this circuit, as long as it's very large.
The question is, though, *how* does it work? The answer is that the resistor divider connected to the inverting input creates a *negative feedback* loop.

7.3 Negative feedback

Negative feedback is a quite simple concept. This explanation may be a bit confusing at first, but think it through carefully and you should get it - and when you do, you'll agree that it's simple! I'll copy the schematic from above for convenience:



Let's think about how this circuit works. Say v_I is a static value of 5 volts, and $R_1 = R_2$ (in which case their actual values won't matter). The node voltage v_- will be exactly half of v_o. v_o, in turn, will be (according to the gain equation we found) $v_I \cdot \dfrac{R+R}{R} = 2 \cdot v_I$.
So, to begin with, we have

$$v_I = v_+ = 5V$$
$$v_o = 2 \cdot v_I = 10V$$
$$v_- = \frac{v_o}{2} = 5V$$

... Well, $v_+ \neq v_-$, but they're very, very close to equal. If they were exactly equal, v_o would be the gain times 0 ($A(v_+ - v_-) = 10^6(5-5) = 0$), which clearly isn't the case here. The difference depends on the op amp's gain A, but will be on the order of microvolts for a gain of 10^6. For now, we'll consider them equal, to drive home the point of negative feedback. In reality, v_o is just ever so slightly different than $2 \cdot v_I$.

So, right now, $v_+ = v_- = 5V$, and $v_o = 10V$. Now, let's think about what would happen if, for whatever reason (op amp instability, external noise, etc.) the voltage at node v_o suddenly increased to, say, 12 volts, without a change in the input.
Clearly, the voltage v_+ would remain unchanged; it will always stay at 5 volts due to the voltage source being directly connected there.
v_-, on the other hand, is half of v_o, and so that node would become 6 volts.
So now, after this output voltage bump, we have $v_+ = 5$, $v_- = 6$ and $v_o = 12$. What happens then?

Well, again, we have the op amp equation: $v_o = A(v_+ - v_-) = A(5-6) = -A$. In words, the op amp would try, as best it could, to drive the voltage v_o down to -10^6 volts (using the gain $A = 10^6$ we've been using so far). Needless to say it couldn't, due to the limited power supply, but it would try really hard to bring the voltage at the output node down.
Let's say it managed to drive the output down to 8 volts in a short time period. At that point, $v_+ = 5V$, $v_- = \dfrac{8}{2} = 4V$, which means the op amp would try to drive the output *up* really hard, since $v_+ > v_-$, and so on.
In practice, the output wouldn't fluctuate this much; rather, the reaction would be fast enough that the change virtually doesn't happen in the first place. This *can* certainly lead to unwanted oscillations, but

that topic is a bit too advanced for this section and chapter.

The important point here is that the negative feedback loop causes the op amp to do whatever it can on the output to keep the two inputs very close to each other. If the output rises too high, causing v_- to increase above v_+, it tries to reduce the output voltage to compensate. If the output drops too low, causing v_- to drop below v_+, it tries to increase the output voltage to compensate.
This is a *negative feedback loop*, a concept which can be applied to many situations, including many not in electronics.
In negative feedback, a gap between the output and expected output will be used to *reduce* this gap, and bring the output closer to the expected value. Therefore negative feedback helps the system become stable.

Positive feedback is, as one might expect, the opposite. With positive feedback, a small gap between the expected and actual values will be used to create an even bigger gap, which in turn will increase the amount of feedback, and make the gap larger yet, etc. A common example is when a microphone picks up sound from a nearby speaker, causing the sound from the microphone to be sent to the speaker, which leads to more sound being picked up by the microphone, increasing the volume in a loop until only an extremely loud noise remains.

7.4 Virtual short/virtual ground

When we have an op amp with negative feedback (and *only* when we have negative feedback!), the feedback causes v_- to be almost equal to v_+. We can show this by rearranging the gain equation:

$$v_o = A(v_+ - v_-)$$
$$v_+ - v_- = \frac{v_o}{A}$$

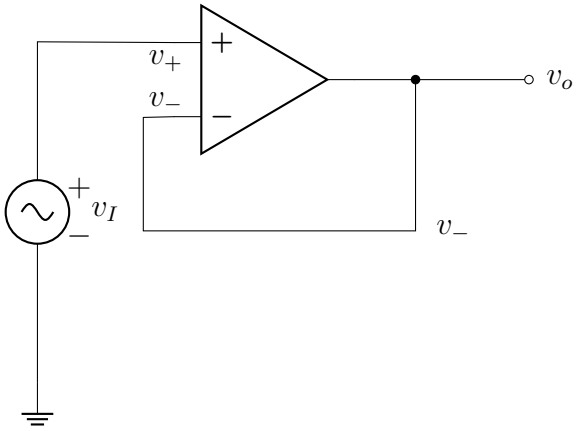
If we have a typical op amp with a very high gain, let's continue assuming 10^6, this means that for v_o in the typical range of $\pm 20V$ or less, the difference between v_+ and v_- will never be larger than on the order of $\frac{20V}{10^6} = 20\mu V$.
Since a short circuit means a zero (or, with wires that have resistances, just a very small) voltage difference, this is referred to as a *virtual short*, or (especially when v_+ is tied to ground) as a *virtual ground*.
This leads to us having three useful constraints when dealing with op amp circuits with negative feedback:

$$i_+ \approx 0$$
$$i_- \approx 0$$
$$v_+ \approx v_-$$

When combined, these constraints give rise to a new, very easy method of analysis for these circuits, as we'll see in the following sections.

7.5 Op amps as buffers

Let's see what happens if we take the noninverting amplifier in the section above, and instead of coupling v_o to v_- through a voltage divider, we simply connect them with a wire.
Since this connection provides negative feedback, we can analyze it extremely easily using the virtual short method outlined above.



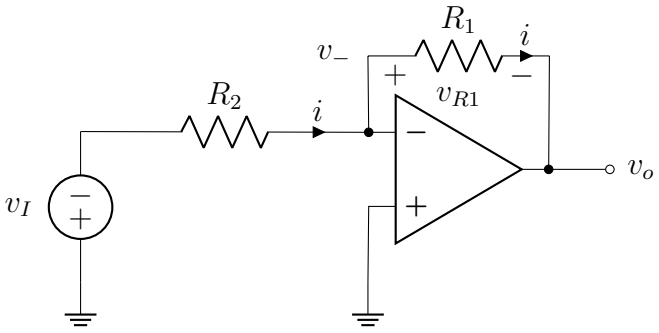
First, we note that $v_+ = v_I$. Via the virtual short, this means that $v_- \approx v_I$. Since $v_o = v_-$ via the wire they share, we draw the conclusion that $v_o \approx v_I$. Well, that was quick!

This circuit is known as a buffer, or source follower. It presents a very large (ideally infinite) impedance towards the input v_I, so that only a very small current is drawn. However, the op amp's output can deliver an ideally infinite current (which is very far from true in practice, as signal op amps, as opposed to power op amps, are often limited to less than 50 or so milliamps).
However, even if the op amp were to be limited to a 10 mA output current, the buffer could still be useful, as many voltage sources such as sensitive sensors may be unable to provide even 1 mA without negative effects. Thus, even a trivial task such as connecting the sensor to some reader, perhaps an ADC (analog-to-digital converter) in a microcontroller, may not work without buffering the voltage first.

We can also build inverting buffers; see the next section.

7.6 Inverting amplifier

Another amplifier topology that uses negative feedback is the *inverting* amplifier, which causes e.g. a 1 volt signal to become $-1 \cdot G$ at the output, where G is the circuit's gain, set by R_1 and R_2.



Since this circuit also uses negative feedback, we can use the virtual short method to analyze it quickly. First, we note that v_+ is tied to ground, i.e. 0 volts. Therefore, by the virtual short constraint, v_- will be approximately 0 volts. Since by our definitions no current flows into the inverting input ($i_- = 0$), and we take $v_- = 0$, the current i through R_2 must be simply $\frac{v_I - 0}{R_2}$. Since, again, no current flows into the inverting input, the same current i must flow through R_1 as they are then in series.
Therefore, the voltage drop across R_1 is that current times the resistance, or

$$v_{R1} = \frac{v_I}{R_2} \cdot R_1$$

The current flows left to right (because of v_I being the driver towards the inverting input at 0 volts), which means the *positive* side of the drop is to the left of R_1 (see the above circuit). Therefore, the node v_o is equal to the voltage at v_- minus that resistor drop:

$$v_o = 0 - \frac{v_I}{R_2} \cdot R_1 = -v_I \cdot \frac{R_1}{R_2}$$

That concludes our analysis! The circuit is indeed inverting, and the (negative) gain is simply $\frac{R_1}{R_2}$.
For the record, the full expression that governs the behavior of this circuit, without the simplifications, is

$$v_o = -\frac{A v_I R_1}{R_1 + R_2 + A R_2}$$

... from an analysis that was performed, but was not outlined in these notes.
As $A \to \infty$, the $R_1 + R_2$ term in the denominator becomes irrelevant, and expression reduces down to one where we can cancel out the As:

$$v_o \approx -\frac{\cancel{A} v_I R_1}{\cancel{A} R_2} = -v_I \cdot \frac{R_1}{R_2}$$

Note that if $R_1 = R_2$, this circuit simply inverts the input, and is therefore an *inverting buffer*, or *unity gain inverting amplifier*, which can certainly be useful.

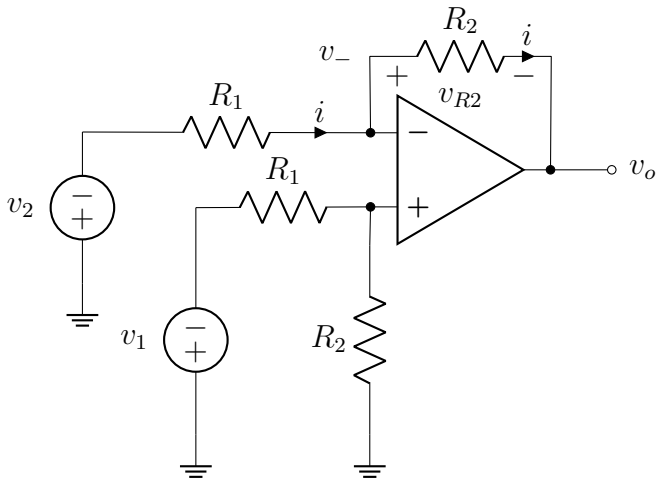
The input impedance of this circuit, that is, $\frac{v_I}{i_I}$, where i_I is the current flowing from the input voltage source, is

$$R_i = \frac{R_1}{A + 1} + R_2$$

Thus, as $A \to \infty$ (or just becomes "very large"), the first term is generally irrelevant, and we can say that $R_i \approx R_2$.

7.7 Op amp subtractor

Let's analyze this circuit:



As with the rest of the op amp circuits we've analyzed, this one uses negative feedback (see R_2), so we will again use the virtual short method to analyze it. However, this analysis is slightly more complex than the previous ones, due to more algebra at the end.
Anyway, let's get started.

Since the noninverting (+) input doesn't draw any current, the voltage at v_+ is given by the voltage divider formed by the v_1 - R_1 - R_2 series circuit:

$$v_+ = v_1 \cdot \frac{R_2}{R_1 + R_2}$$

By the virtual short method, $v_- = v_+$:

$$v_- = v_+$$

Therefore, since we know v_- and v_2, we can calculate the current through R_1, which we'll simply call i:

$$i = \frac{v_2 - v_-}{R_1}$$

Note that, since the inverting (−) input draws no current, all of this current must flow through R_2 as well (see the circuit diagram). Therefore, the we can easily calculate the voltage drop across R_2 as current times resistance, or

$$v_{R2} = \frac{v_2 - v_-}{R_1} \cdot R_2$$

By know, we know v_-, and we know the voltage drop across R_2. Since that voltage drop is positive "towards the left", v_o will be v_- *minus* the drop across R_2:

$$v_o = v_- - v_{R2}$$

$$v_o = v_- - \frac{v_2 - v_-}{R_1} \cdot R_2$$

Let's gather the v_- terms together:

$$v_o = v_-(1 + \frac{R_2}{R_1}) - \frac{v_2}{R_1} \cdot R_2$$

$$v_o = v_-(1 + \frac{R_2}{R_1}) - v_2 \cdot \frac{R_2}{R_1}$$

Let's substitute in v_-; remember, it's equal to the expression we have for v_+:

$$v_o = (v_1 \cdot \frac{R_2}{R_1 + R_2})(1 + \frac{R_2}{R_1}) - v_2 \cdot \frac{R_2}{R_1}$$

What a mess. Let's rewrite $1 + \frac{R_2}{R_1}$:

$$v_o = (v_1 \cdot \frac{R_2}{R_1 + R_2})(\frac{R_1 + R_2}{R_1}) - v_2 \cdot \frac{R_2}{R_1}$$

Aha! Now we can cancel out those $R_1 + R_2$ terms:

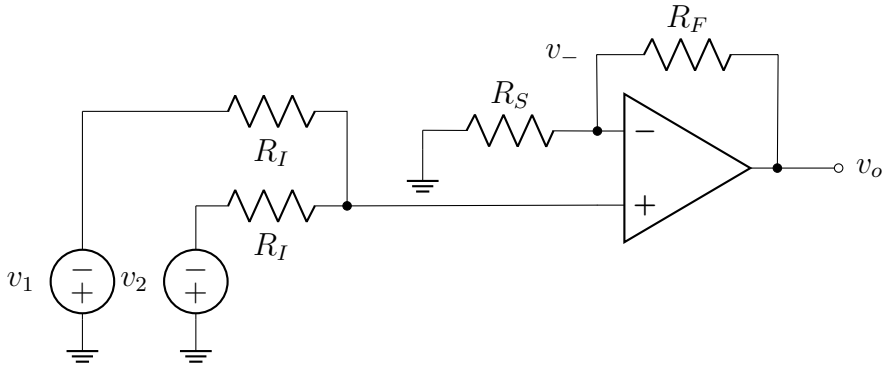
$$v_o = v_1 \cdot \frac{R_2}{R_1} - v_2 \cdot \frac{R_2}{R_1}$$

$$v_o = \frac{R_2}{R_1}(v_1 - v_2)$$

As promised in the section title, this is an op amp *subtractor* circuit. It calculates $v_1 - v_2$, scaled by the ratio of R_2 to R_1.

7.8 Summing amplifier

Since the previous section covered an op amp subtractor, let's cover an adder, usually called a *summing amplifier*:



Yet again, this circuit uses negative feedback, so we can use the virtual short method.
First, we can figure out v_+. Since no current passes through the noninverting (+) input, we can find v_+ by analysing the v_1 - R_I - R_I - v_2 subcircuit.
Since it's linear, we can use superposition, and figure out the voltage due to v_1 acting alone (v_{+1}) and v_2 acting alone (v_{+2}). To use superposition, we "set one source to zero" at a time, which for voltage sources means shorting it out. For v_1:

$$v_{+1} = v_1 \cdot \frac{R_I}{R_I + R_I} = \frac{v_1}{2}$$

(since the resistors have the same value.)

Similarly, for v_2:

$$v_{+2} = v_2 \cdot \frac{R_I}{R_I + R_I} = \frac{v_2}{2}$$

So, the sum of the responses will be v_+:

$$v_+ = \frac{v_1 + v_2}{2}$$

By the way, note that if the resistors had unequal values, the above would change. Not only would they not be divided by exactly two, but we would get a *weighted average*, since the top resistor in this schematic is in the numerator for the v_2 divider, while the bottom resistor is in the numerator for v_1.
If we relabel the top resistor R_1 and the bottom resistor R_2, and set $R_1 = 1000$, $R_2 = 3000$:

$$v_+ = v_1 \cdot \frac{R_2}{R_1 + R_2} + v_2 \cdot \frac{R_1}{R_1 + R_2} = \frac{3v_1}{4} + \frac{v_2}{4}$$

After that aside, let's get back on track. We found v_+, and via the virtual short method, $v_- = v_+$.

There are several ways to find v_o from here, but the node method gives a solution with only minor algebra. We set up a node equation for v_-:

$$\frac{v_-}{R_S} + \frac{v_- - v_o}{R_F} = 0$$

Move v_o to the right-hand side and multiply both sides by R_F:

$$\frac{R_F v_-}{R_S} + v_- = v_o$$

Switch the sides for clarity and factor out v_-:

$$v_o = \frac{R_F v_-}{R_S} + v_-$$

$$v_o = v_-(\frac{R_F}{R_S} + 1)$$

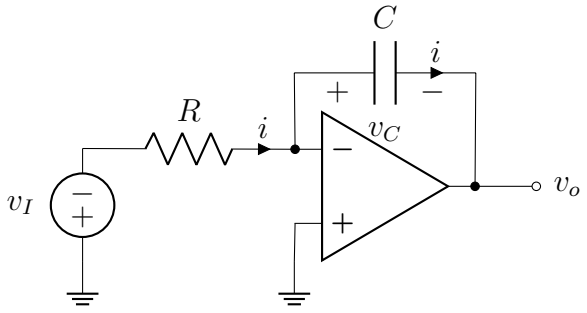
And, finally, get the actual value of v_- in there:

$$v_o = \frac{v_1 + v_2}{2}(\frac{R_F}{R_S} + 1)$$

We're done! We can also easily see that if we set $R_F = R_S$, the output is simply $v_o = v_1 + v_2$.

7.9 Op amp integrator

Due to some magic of capacitors, we can use op amps to create circuits that integrate and differentiate signals (with respect to time). Here's an integrator circuit:



How does this circuit work? Well, remember the formula for the voltage over a capacitor:

$$v_C = \frac{1}{C}\int_{-\infty}^{t} i(t)dt$$

v_+ is grounded, and thus at 0 volts. v_- is, via the virtual short technique, also 0 volts. Therefore, the current through R is simply $\frac{v_I}{R}$.

That same current must go through the capacitor, since there's no other way (the noninverting input has an infinite impedance), so the capacitor voltage drop is

$$v_C = \frac{1}{C}\int_{-\infty}^{t} \frac{v_I}{R} dt = \frac{1}{RC}\int_{-\infty}^{t} v_I dt$$

As with the previous op amp circuits where the noninverting input is grounded, we can find v_o as

$$v_o = v_- - v_C = 0 - \frac{1}{RC}\int_{-\infty}^{t} v_I dt$$

$$v_o = -\frac{1}{RC}\int_{-\infty}^{t} v_I dt$$

Thus this is an inverting integrator. If needed, we could feed the output into an inverting buffer (aka inverting unity-gain amplifier) to get it positive.

If we do a more complete analysis, we get the differential equation

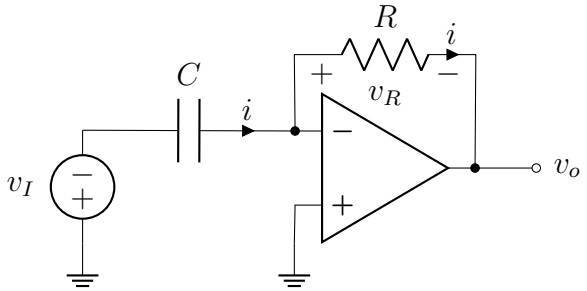
$$RC(1+A)\frac{dv_-}{dt} + v_- = v_I$$

indicating that the time constant for this circuit is multiplied by the op amp gain A. Therefore, the equation that describes v_o for a step input is of the same form as we've seen in regular series RC circuits:

$$v_o = -AV(1 - e^{-\frac{t}{(1+A)RC}})$$

... where V is the voltage of the step input.
Thus, for short time scales (short compared to $(1+A)RC$), this circuit is a very good approximation of an ideal integrator.

7.10 Op amp differentiator



This circuit is identical to the integrator except for *one* change: the capacitor and resistor places have been swapped. So, how does this work, then?

v_+ is grounded, and so via the virtual short method, v_- must also be 0 volts. Therefore, the voltage drop across the capacitor equals v_I. The current through the capacitor is given by

$$i = C\frac{dv_I}{dt}$$

That same current i is then forced through the resistor R, which causes a voltage drop of $R \cdot i$ across the resistor:

$$v_R = RC\frac{dv_I}{dt}$$

And, as with several of the previous circuits, v_o equals v_- minus the drop across the resistor:

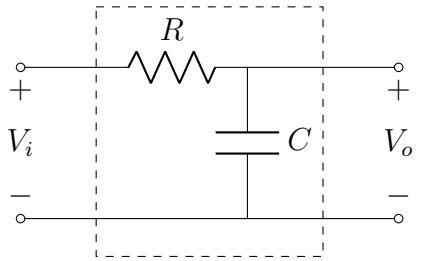
$$v_o = v_- - v_R = 0 - RC\frac{dv_I}{dt}$$

$$v_o = -RC\frac{dv_I}{dt}$$

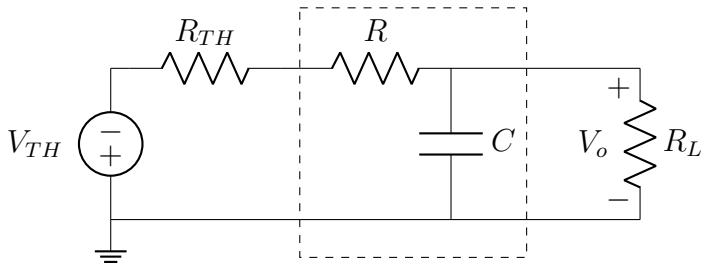
So, as with the previous circuit, this one also inverts. And, again, we could pass it through an inverting buffer op amp to get rid of the inversion.

7.11 Buffered filters

Consider a basic passive RC filter:



The dashed lines represent the filter components (the reason for them will be clear soon).
We've analyzed this circuit in a previous chapter. However, let's look at how a more complete circuit might look, when we attach a Thevenin equivalent source and a resistive load (perhaps a speaker):



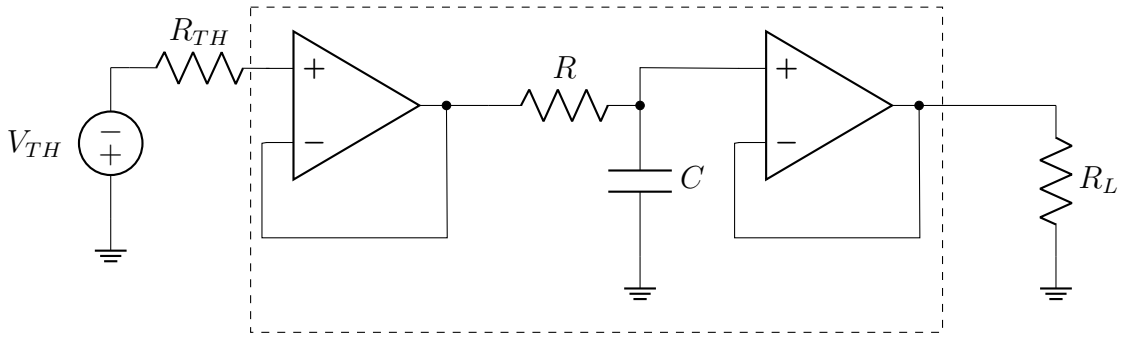
If we analyze this circuit, we'll soon find that the transfer function is not at all the one we would expect from just R and C. Rather, the transfer function would be

$$H(s) = \frac{\frac{1}{sC}||R_L}{(\frac{1}{sC}||R_L) + R_{TH} + R}$$

Instead of what we actually wanted:

$$H(s) = \frac{\frac{1}{sC}}{\frac{1}{sC} + R}$$

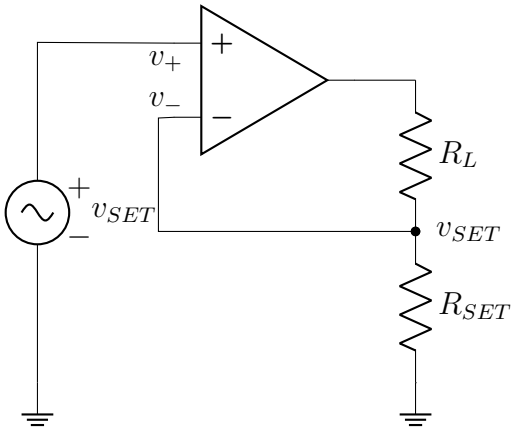
What can we do about this? Well, one solution - that works just fine, but is rather wasteful - is to buffer both the input and output of the filter:



That's a fair bit of extra complexity! However, the input op amp will make sure that virtually 0 current is drawn from the source, and decouple the input from the filter.
Likewise, the output op amp will make sure virtually no current is drawn from the RC network, so that the load resistance doesn't affect the transfer function.
Assuming ideal op amps, this filter should behave as a passive filter does when there is no load (infinite load resistance) and 0 source resistance. This filter also has (with ideal op amps) infinite input impedance and 0 output impedance!

7.12 Op amp current source

We can use op amps to make very precise current sources. Here's one of several ways to do so:

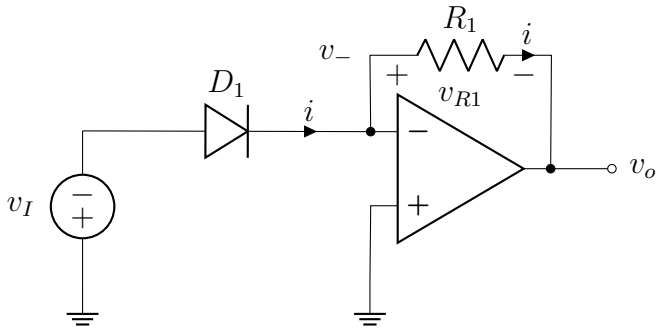


Note that the topology above is identical to the noninverting amplifier! I only changed the resistor and voltage source names, to make things more clear.
How does this circuit work? Well, we have negative feedback, as usual. Therefore, we can apply the virtual short method, and see that the voltage across R_{SET} is v_{SET}. Therefore, the current through it is $\frac{v_{SET}}{R_{SET}}$.
Since the current into the noninverting input is negligible (even taken as 0 for our model), R_L and R_{SET} are in series, and must share the same current! Therefore, the current through R_L is the same as the current through R_{SET}, which the op amp tries to keep equal to $\frac{v_{SET}}{R_{SET}}$ at all times.

Of course, there are some limitations. Most op amps have very low current limits, so this will only work for small currents (a few mA, almost certainly less than 40 or so) or when using power op amps rated for high currents. Second, the current can clearly never be larger than the supply voltage divided by $R_{SET} + R_L$, since the supply voltage is the maximum the op amp can have on its output - in theory, in practice the actual maximum output voltage will always be slightly lower than the supply voltage.

7.13 Exponential amplifier

The inverting amplifier topology is extremely versatile, as we've already seen with the integrator and differentiator configurations. We can do more, however! Lets see what happens with a diode in there:



The analysis is very similar to that of the inverting amplifier and the integrator/differentiator.

v_+ is grounded, so $v_- \approx 0$ by the virtual short method (since we have negative feedback). Therefore, the voltage across the diode is simply v_I. The current through a diode is given by the Shockley diode equation:

$$i_D = I_S(e^{v_D/V_T} - 1)$$

I_S and V_T are diode parameters. I_S is the "reverse bias saturation current", usually a very small value on the order of 10^{-14} A, while V_T is the thermal voltage, roughly 25 mV at room temperature.
Now, given that we know the diode voltage drop, we can calculate that the current through it is:

$$i = I_S(e^{v_I/V_T} - 1)$$

That same current will, as with the previous configurations, be forced to flow through R_1 and create a voltage drop across it.
v_o can be seen to be v_- minus the drop across R_1, and since v_- is 0:

$$v_o = -i \cdot R_1 = -R_1 I_S(e^{v_I/V_T} - 1)$$

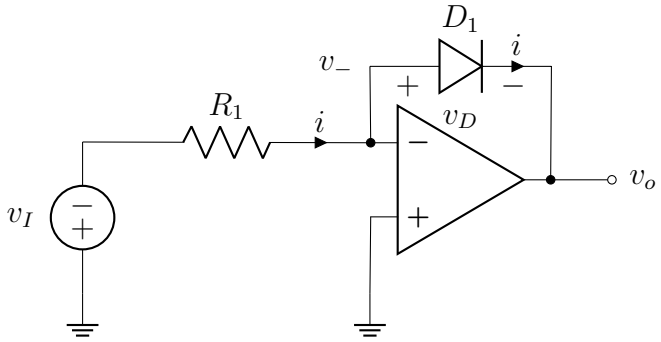
The above equation makes it clear that there's an exponential relationship between v_I and v_o. For a typical silicon diode at room temperature, with $R_1 = 1000\Omega$, we can estimate(!) the relationship as

$$v_o = -10^{-11} \cdot e^{v_I/0.025}$$

Note that the last "1" was dropped, as its value was 10^{-11} A, a current which is insignificantly small to us.
Though the 10^{-11} term looks like it might dominate, it does not; the exponential increases extremely fast, and the output voltage goes below $-100V$ (remember that it's an inverting amplifier) when the input is still as low as 0.75 volts. Of course, the op amp will saturate if the power supply voltage doesn't reach that low.

7.14 Logarithmic amplifier

By swapping the position of the diode and resistor in the exponential amplifier, we get this schematic:



This is still easy to analyze, but there's a bit more math. We still use the diode equation, but this time, we need to solve the equation for v_D before we can use it to calculate a voltage drop given a current (the form we have only calculates current from a known voltage).

Again, the Shockley diode equation is

$$i_D = I_S(e^{v_D/V_T} - 1)$$

The current i_D through the diode will be given by the current through R_1, which is $\frac{v_I - v_-}{R_1}$. As v_- is 0 via the virtual short (this should be second nature at this point!), the current is simply given by $\frac{v_I}{R}$, so let's make that substitution. Then, we will need to solve the equation for v_D. Let's start by distributing the I_S, and getting the exponential on its own:

$$\frac{v_I}{R} = I_S(e^{v_D/V_T} - 1)$$
$$\frac{v_I}{R} = I_S \cdot e^{v_D/V_T} - I_S$$
$$\frac{v_I}{R} + I_S = I_S \cdot e^{v_D/V_T}$$

Divide both sides by I_S:

$$\frac{v_I}{RI_S} + 1 = e^{v_D/V_T}$$

Take the natural log of both sides:

$$\ln\left(\frac{v_I}{RI_S} + 1\right) = \frac{v_D}{V_T}$$

And, finally, multiply both sides by V_T, and just switch the sides:

$$v_D = V_T \cdot \ln\left(\frac{v_I}{RI_S} + 1\right)$$

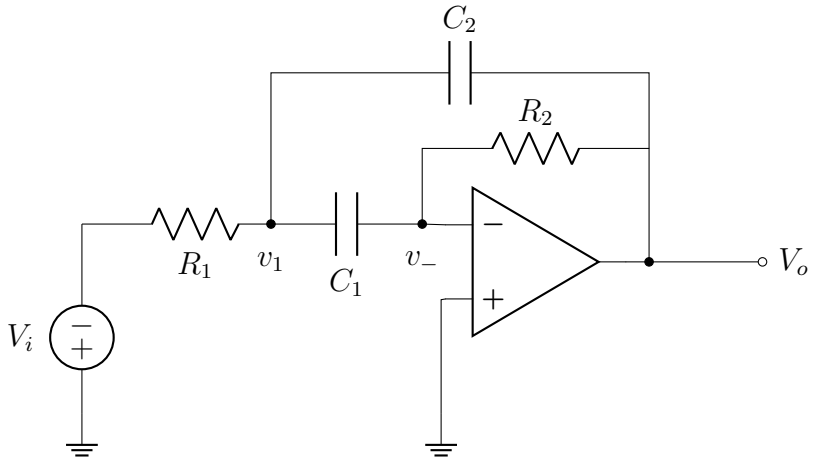
Now we know how to calculate the diode's voltage drop, which makes us very close to finding v_o. We know that $v_o = v_- - v_D$, and that v_- is 0, so

$$v_o = -v_D$$
$$v_o = -V_T \cdot \ln\left(\frac{v_I}{RI_S} + 1\right)$$

And we are done! This is indeed a logarithmic amplifier.

7.15 Active RC filters

We can also construct filters that use op amps as part of the actual circuitry, rather than just having them as input/output buffer. Such filters are called *active filters*, and they come in various topologies, where the op amp's function differs.
Here's an example for an active RC filter that we'll analyze in this section:



Let's get started. We know that $V_o = A(v_+ - v_-) = -Av_-$.
Let's write node equations for v_1 and v_- using the impedance method:

$$\frac{v_1 - V_i}{R_1} + \frac{v_1 - V_o}{Z_{C2}} + \frac{v_1}{Z_{C1}} = 0$$

(since $v_- \approx 0$ via the virtual short technique.) and

$$\frac{-v_1}{Z_{C1}} + \frac{-V_o}{R2} = 0$$

(Again, since $v_- \approx 0$ that term disappears, to simplify our calculations.)
After substituting $Z_{C1} = \frac{1}{sC_1}$ and $Z_{C2} = \frac{1}{sC_2}$, we solve the equations for V_o (I used Mathematica), simplify and get

$$V_o = -\frac{sC_1R_2V_i}{s^2 + \frac{C_1+C_2}{C_1C_2R_2}s + \frac{1}{C_1C_2R_1R_2}}$$

Comparing the denominator to the canonical form $s^2 + 2\alpha s + {\omega_0}^2$ we get

$$2\alpha = \Delta\omega = \frac{C_1 + C_2}{C_1C_2R_2}$$

$$\omega_0 = \frac{1}{\sqrt{C_1C_2R_1R_2}}$$

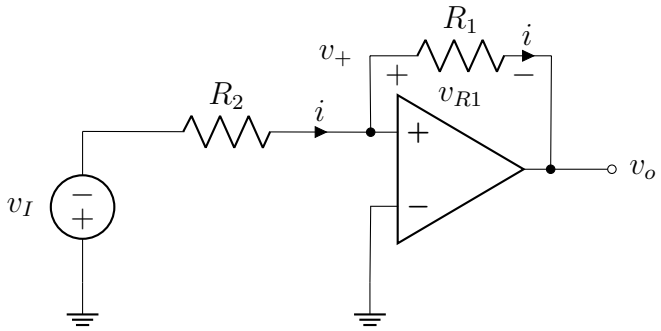
$$Q = \frac{\omega_0}{2\alpha} = \frac{\sqrt{C_1C_2R_1R_2}}{R_1C_1 + R_1C_2}$$

If we substitute $s = j\omega$ and divide by V_i, we get the circuit's transfer function $H(j\omega)$, also simply $\frac{V_o}{V_i}$:

$$H(j\omega) = -\frac{j\omega C_1R_2}{-\omega^2 + \frac{j\omega(C_1+C_2)}{C_1C_2R_2} + \frac{1}{C_1C_2R_1R_2}}$$

7.16 Positive feedback

Positive feedback was briefly touched upon in the section on *negative* feedback, but deserves more study. Let's see what happens if we make a connection similar to the inverting amplifier, but connect the feedback and input to the *positive* terminal, and ground the *negative* terminal.

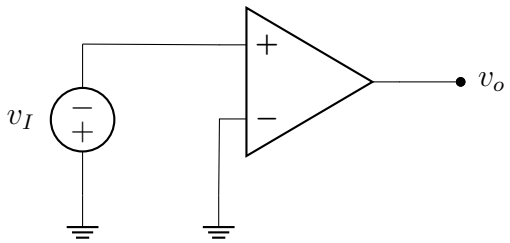


Let's analyze this circuit intuitively first. Note that we're feeding a portion of the output back to the *positive* terminal of the op amp, which means we now have *positive* feedback.
Assume that the circuit is in equilibrium, to begin with. What would happen in this circuit if, for whatever reason (noise, external excitation, etc.) the output node v_o would increase in voltage?
Well, we have a voltage divider back to the positive terminal, so a portion of that voltage increase would appear at the positive terminal. Via the op amp equation $v_o = A(v_+ - v_-)$, the op amp would hugely amplify that voltage, and vastly increase the output. That, in turn, would make an even higher voltage appear at the + terminal, which would get amplified, etc. Very quickly, the op amp will spiral out of control, saturate and hit the positive supply rail.

Let's then think about what would happen if we start out at $v_+ = v_- = v_o = 0$ volts. Say v_o drops down to -1 volt. A portion of that would be fed back to the positive input; for this example, let's say $R_1 = R_2$, so that exactly half gets fed back.
-0.5 volts would appear at v_+, while v_- would remain at 0 volts as always (it's grounded!). Via the good old op amp equation, we have $v_o = A(-0.5 - 0)$, so the op amp will try to force the output down as low as it can (as the value of A is very, very large, the op amp wants v_o to be much less than the supply voltage that might be on the order of -12 volts).

7.17 Op amp comparator

Let's try to build a circuit that exploits the positive feedback phenomenon, where the output voltage is virtually always either $+V_S$ or $-V_S$. We can use such a circuit as a comparator:



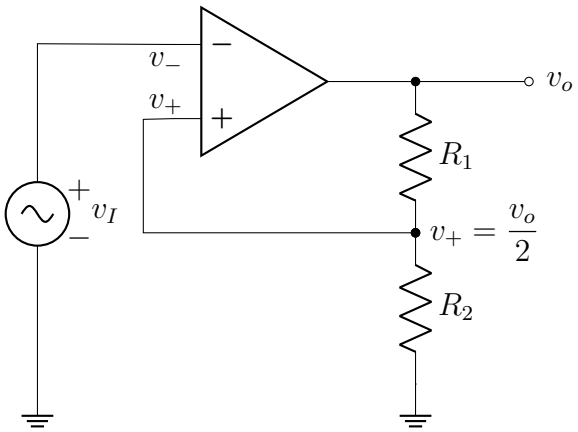
This circuit will have $v_o = V_S$ whenever v_I is positive ($v_I > v_-$, and v_- is grounded) and $v_o = -V_S$ when v_I is negative ($v_I < v_-$), as long as $Av_+ > V_S$ or $Av_+ < -V_S$ is true, so that the op amp is saturated.
Of course, the v_- terminal doesn't have to be grounded - it could be set to some reference voltage that we want to compare to v_+, via a voltage divider, perhaps. It could even be a second signal.
Here's the transfer function for a comparator, normalized so that v_o ranges between 1 and -1:

(The filled areas are the output; the unconnected lines were hard to spot without the filling.)
Looks great! However, what whould happen if there were noise on the input voltage? The distorted input could cross zero multiple times due to noise, so the output is not a clean square wave:

Because the noise causes the input to cross zero multiple times extra per cycle, there are glitches in the square wave output. This example is obviously exaggerated, but we still want a better solution. If we could move the threshold from 0 to some positive/negative voltage pair, so that the noise doesn't cross those thresholds very easily, this could be avoided.
Let's see (in the next section) what would happen to this transfer function if we add positive feedback.

7.18 Schmitt trigger

If we change the circuit slightly from the previous comparator to add positive feedback, and at the same time provide the input to the inverting input, we get this circuit:



This circuit is a *Schmitt trigger*; more specifically, an inverting Schmitt trigger. It has *hysteresis* (it "remembers" the previous state of the circuit), and will solve the problem we had. How? Well let's analyze the circuit and find out!

Let's do a manual "simulation" again, like we did with the previous comparator. Also, let's use some numerical values to make it a bit more intuitive. We'll use $R_1 = R_2$ (their values don't really matter, only the ratio) and $V_S = 15$ V (and therefore $-V_S = -15$ V). Also, we'll start off our simulation with $v_I = 0$ V and $v_o = 15$ V, a state that we'll soon see is valid for the normal operation of this circuit.

With the values we've chosen, v_+ will be half of v_o, so $v_+ = 7.5$ volts. The op amp equation always applies, so with $v_- = v_I = 0$ volts, the op amp will saturate the output, as $v_o = A(7.5 - v_I) = A(7.5 - 0)$ far exceeds $V_S = 15$ volts. Let's look at that equation a bit more carefully. As long as A is very big (again, it's usually over 10^4, and not rarely over 10^6), v_o will stick to V_S as long as the input is less than 7.5 volts. For example, say the input increases all the way to 7.4 volts; via $v_o = A(7.5 - 7.4)$, the output will still be $v_o = 0.1 \cdot A$ which still far exceeds V_S.
Interesting things happen when $v_- > v_+$, however. For example, when $v_I = v_- = 7.6$ volts, $v_+ - v_-$ becomes negative, and the op amp equation evaluates to a hugely negative number, and so the op amp will saturate at the *negative* rail $-V_S$, since $v_o = A(7.5 - 7.6) = -0.1 \cdot A \ll -V_S$.

Now that the input has gone above the old "switching threshold" of 7.5 volts, let's see what happens when we go back below again. Note that now that $v_o = -15$ volts, $v_+ = -7.5$ volts, thanks to the voltage divider at the output.
We have $v_o = A(v_+ - v_-) = A(-7.5 - 7.6)$ at the moment, which evaluates to a very negative number, so the op amp should indeed be in negative saturation. What happens if we lower the input voltage to 7.4 volts, below the old switching threshold? It still evaluates to a negative number! Even if we go to $v_I = 0$, it will *still* evaluate to $v_o = A(-7.5)$. In order to get the output to switch, we need to go all the way down to $v_I \approx -7.5$ volts (ever-so-slightly lower). Let's use $v_I = -7.6$ volts. Via the good old op amp equation, we have $v_o = A(-7.5 - (-7.6)) = A(0.1)$, which means the output will now switch to positive saturation again, and v_+ will therefore become +7.5 volts.

Note that now that we've switched again, the only way to switch the output back to the negative voltage is to go above $v_I = 7.5$ volts, at which point $v_O = A(7.5 - v_I)$ will turn negative.

So, as can be seen in the rather messy analysis above, a signal like our sine wave will cause a clean square wave output, as the noise we had comes nowhere near $V_S/2$ or $-V_S/2$, and so the output won't switch more

than once per intentional zero crossing. This trigger's inverting nature is also clear, as it only switches to a negative output when the input gets *positive* enough.

Let's go back to the noisy sine wave, and see how the same input looks through a Schmitt trigger, instead of a comparator:

(Again, the filled areas are the output.)
As in the textual example above, we start out with the input at 0 volts (plus a whole ton of noise, which makes it start out closer to 5 than 0), and start the world with $v_o = 15$ V, and take it from there.
Despite the extreme, exaggerated noise, the output is a perfectly clean square wave! Note how the output switches whenever the input crosses the ± 7.5 V threshold set by the voltage divider, but only from "one direction" - for example, when the sine wave is starting to go up from its bottom position around the graph's middle, it crosses the -7.5 volt threshold without anything happening. In just this graph, the sine wave crosses the thresholds 6 times (6 intended times - not counting the extra ones due to the noise), but only switches 3 times. This is due to hysteresis - the past matters; which "direction" it comes from decides whether the output will switch or not.
Also, of course, note how the output is an inversion of the input. It is however perfectly possible to build non-inverting Schmitt triggers.

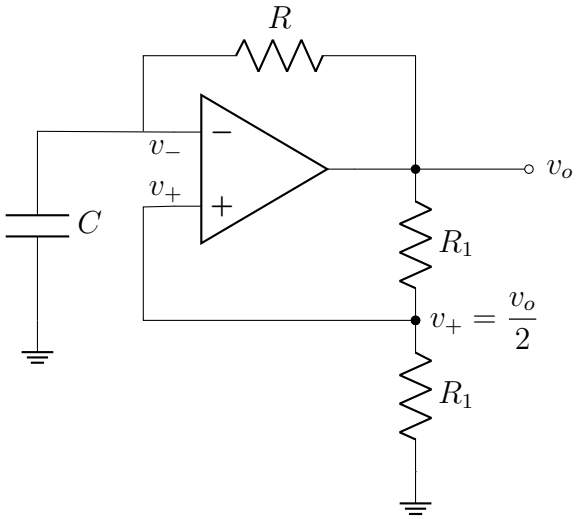
We can also vary the switching thresholds. The equation is

$$V_{thres} = \pm v_I \cdot \frac{R_2}{R_1 + R_2}$$

... where R_2 is the resistor between v_+ and ground.

7.19 Relaxation oscillator

We can use the inverting Schmitt trigger to make an op amp oscillator, by allowing a capacitor to control the v_- voltage. The capacitor is then charged/discharged through a resistor connected to v_o, like this:



Notice that the bottom resistors are the same value, so that the thresholds are again set at $\pm V_S$.
The "top" resistor R is the only one that matters for the rise/fall times, while the "bottom" ones are the only ones that matter for setting the switching thresholds.

Let's take the usual intuitive look at how this would work. Imagine it starts out with the capacitor voltage (which is the same as v_-) at 0 volts, and v_o at the positive rail, 15 volts (so -V_S is then -15 volts).
In this state, a current will flow from v_o to v_-, due to the voltage difference (15 - 0 volts). This current flows through R_1, and has nowhere to go but through the capacitor, so the capacitor is charged by this current. It keeps charging along happily with no other change in the circuit, up until it passes just slightly above 7.5 volts ($V_S/2$). At that point, according to the op amp equation, $v_o = A(7.5 - v_-)$ will turn negative, and force v_o down to the negative rail.
When that happens, the capacitor voltage will be almost exactly 7.5 volts, while v_o will be at -15. Therefore the capacitor will discharge through R, until it's down to $V_S/2 = -7.5$ volts, at which point $v_o = A(-7.5 - v_-)$ will turn positive, and force v_o to the positive rail. This cycle will repeat over and over, and produce - at least in theory - a perfect square wave, with a 50% duty cycle (i.e. the "on time" is exactly equal to the "off time").

Now, let's turn to calculating the oscillation frequency of this circuit. The way we do this is to calculate the rise time and the fall time for the RC circuit at the inverting input, and use that information to find the frequency.

7.19.1 Rise time

To calculate the rise time, let us define $t = 0$ such that the capacitor voltage is as low as it ever gets, i.e. $-V_S/2$. We can then use our old trusty formula $v_C = V_S + (V_0 - V_S)e^{-t/RC}$, and set that equal to the target voltage, $V_S/2$:

$$\frac{V_S}{2} = V_S + (-\frac{V_S}{2} - V_S)e^{-t_r/RC}$$
$$\frac{V_S}{2} = V_S - \frac{3V_S}{2}e^{-t_r/RC}$$

To start with, we subtract V_S from both sides, and then multiply by -1:

$$-\frac{V_S}{2} = -\frac{3V_S}{2}e^{-t_r/RC}$$

$$\frac{V_S}{2} = \frac{3V_S}{2}e^{-t_r/RC}$$

We can now divide both sides by $\frac{3V_S}{2}$, which is the same as multiplying by $\frac{2}{3V_S}$:

$$\frac{V_S}{2} \cdot \frac{2}{3V_S} = \frac{2}{3V_S} \cdot \frac{3V_S}{2}e^{-t_r/RC}$$

$$\frac{\cancel{V_S}}{\cancel{2}} \cdot \frac{\cancel{2}}{3\cancel{V_S}} = e^{-t_r/RC}$$

This remains:

$$\frac{1}{3} = e^{-t_r/RC}$$

We take the natural log of both sides, and multiply both sides by -RC:

$$\ln\left(\frac{1}{3}\right) = -\frac{t_r}{RC}$$

$$-RC\ln\left(\frac{1}{3}\right) = t_r$$

Via log rules, $\ln(1/3) = \ln(1) - \ln(3) = 0 - \ln(3)$:

$$t_r = RC\ln(3)$$

Ah, finally: we know the rise time. Next up: fall time.

7.19.2 Fall time

For the fall time, we use the same equation $v_C = V_S + (V_0 - V_S)e^{-t/RC}$, but plug in other values for the variables, and then set it equal to the target value of $-V_S/2$. Remember that the "supply" that is "charging" the capacitor is now $-V_S$, and the initial voltage now $V_S/2$:

$$-\frac{V_S}{2} = -V_S + (\frac{V_S}{2} - (-V_S))e^{-t_f/RC}$$

$$-\frac{V_S}{2} = -V_S + \frac{3V_S}{2}e^{-t_f/RC}$$

Compare this form to the rise time:

$$\frac{V_S}{2} = V_S - \frac{3V_S}{2}e^{-t_r/RC}$$

If we multiply both sides of the fall time equation by -1, we can easily see that these equations are exactly the same, which means the answer must also be the same!

$$t_f = t_r = RC\ln(3)$$

7.19.3 Frequency

Frequency is the inverse of cycle time, and $t_r + t_f$ make up one full cycle. Therefore,

$$f = \frac{1}{t_r + t_f} = \frac{1}{2RC\ln(3)}$$

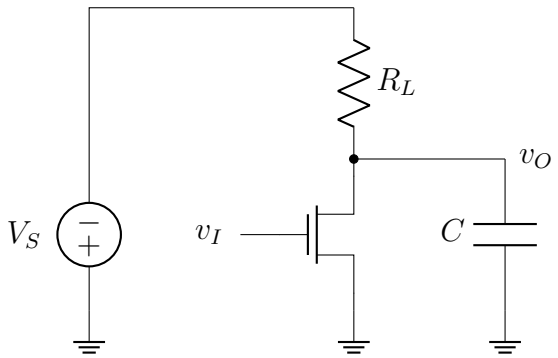
If the negative supply is not equal to exactly $-V_S$, the formula above (and its derivation) becomes quite a bit more complex. We will not worry about that case, as having equal supplies to within a small margin of error is common.

Chapter 8

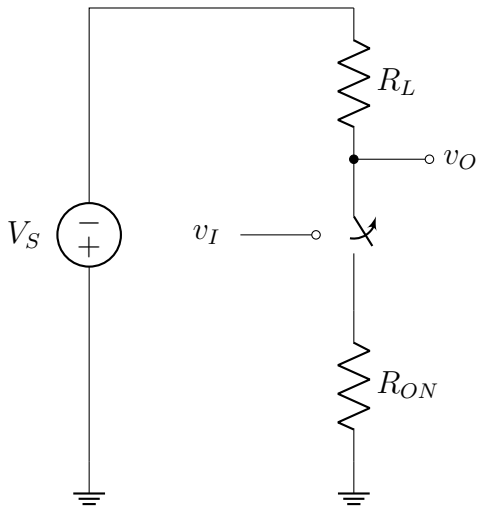
Energy and power in digital circuits

8.1 Power usage for an NMOS inverter

Let's look at the power usage of our good old resistor/MOSFET inverter. However, we'll add a capacitance between v_O and ground, to model the capacitance being driven (the input of other gates, perhaps, and the parasitic capacitances in the wires).



Using the SR model for the MOSFET, and considering the case for a static v_I, which means the capacitor will act as a long-term open circuit, this circuit is equivalent to the previous one:



We will use this circuit to calculate this inverter's *static power*, i.e. the power used when the output is not switching, but either constantly low or constantly high.
How do we calculate the power used by this circuit? Doing so is of course extremely simple. There are two cases, switch open and switch closed.

Switch open: no current flows, and so the power usage is zero.

Switch closed: A current $\frac{V_S}{R_L + R_{ON}}$ flows (we assume the switch is ideal and has 0 resistance). The power usage is that current squared times the resistance, or (much simpler) via $P = \frac{V^2}{R}$, just

$$P = \frac{{V_S}^2}{R_L + R_{ON}}$$

Since there are two cases with vastly different power draws, we can use the average of the two: if we assume a 50% probability for the gate to be on, and 50% for it to be off, the average power is roughly

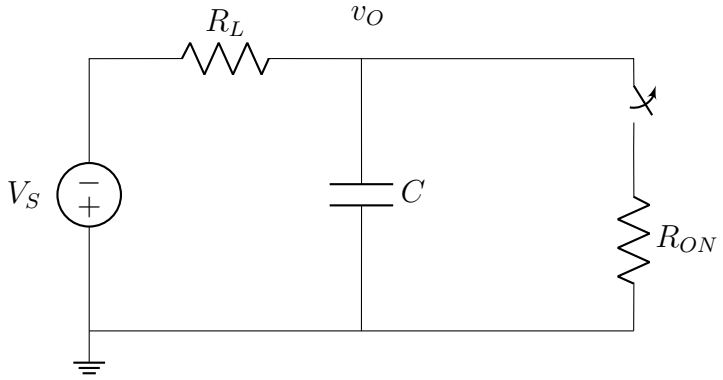
$$\overline{P_{static}} = \frac{\frac{{V_S}^2}{R_L + R_{ON}} + 0}{2}$$

$$\overline{P_{static}} = \frac{{V_S}^2}{2(R_L + R_{ON})}$$

So, we now know the static power of this circuit.
The reason we average the two cases is not that any single inverter is going to be at each state an equal amount of time, but rather that on a chip with millions or even billions of such gates, we consider the probability of any given one to be either on or off to be 50%.

As you might expect, there's a counterpart to static power called *dynamic power*, which is the power used when switching the inverter at a certain frequency.
Unfortunately, the derivation of that expression is rather messy, though not strictly *hard*. Only techniques previously discussed are required, so it will be left as an exercise to the reader, if you are interested. If not, read on.

This circuit is equivalent to the inverter circuit, if we use the SR model for the MOSFET once again:



If we drive the input with a square wave with a 50% duty cycle - that is, the on time is exactly equal to the off time, it turns out (this is the skipped derivation, of course) that the dynamic power is

$$\overline{P_{dynamic}} = C{V_S}^2 f \frac{{R_L}^2}{(R_L + R_{ON})^2}$$

Let's add the static power to that:

$$\overline{P} = \overline{P_{static}} + \overline{P_{dynamic}}$$

$$\overline{P} = \frac{{V_S}^2}{2(R_L + R_{ON})} + C{V_S}^2 f \frac{{R_L}^2}{(R_L + R_{ON})^2}$$

That's a slightly complex expression, but we can simplify it quite nicely. R_L is usually much greater than R_{ON}, to the point where we can simply neglect R_{ON}. Doing so, the equation reduces to

$$\overline{P} = \frac{{V_S}^2}{2R_L} + C{V_S}^2 f \frac{{R_L}^2}{{R_L}^2}$$

Ah! Now we can get rid of the entire fraction to the very right:

$$\overline{P} = \frac{{V_S}^2}{2R_L} + C{V_S}^2 f$$

Not bad! Note how the dynamic power depends on three things: the capacitance C (a linear dependence), the switching frequency f (also a linear dependence) and the supply voltage V_S - a quadratic dependence! Halving the supply voltage cuts the dynamic power usage in four!

Now, let's run some example numbers for a hypothetical chip we might want to build. Say we have a chip with *5 million* of these inverter gates, with $R_L = 10\text{k}\Omega$, $C = 1$ fF (10^-15 F) and $V_S = 5$ volts, and we will run it at 3 GHz. Let's put these numbers into the power equation:

$$\overline{P} = \frac{{V_S}^2}{2R_L} + C{V_S}^2 f$$

$$\begin{aligned} P &= 5 \cdot 10^6 \left[\frac{5^2}{2 \cdot 10000} + 10^{-15} \cdot 5^2 \cdot 3 \cdot 10^9 \right] \\ &= 5 \cdot 10^6 \left[1.25 \text{ mW} + 75 \ \mu\text{W} \right] \\ &= 6250 + 375 = 6625 \text{ W} \end{aligned}$$

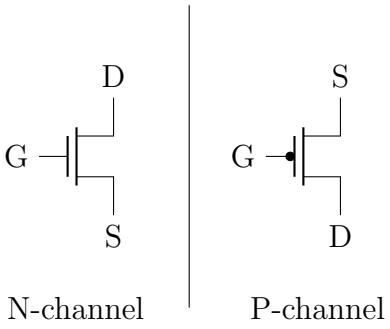
... so this chip apparently consumes, on average, *6.625 kilowatts* of power!!! That's enough to power 50-100 modern quad-core CPUs, and our relatively simple chip only has 5 million gates, which means we are probably a bit over a decade behind Intel's and AMD's current offerings! This is clearly untenable. The next section will discuss a means to eliminate *all* of the static power, and also list various ways to bring the dynamic power down to more reasonable levels. 375 watts is possible to cool, but not in a way that is both quiet and cheap.
As of mid-2012, Intel's top-end mainstream CPUs (not their most extreme ones, but the kind most enthusiasts would buy) have a Thermal Design Power (TDP) of 77 watts, down from 95 watts in the previous generations. Actual consumption values can be (much) lower due to power saving when not at full load, however.

8.2 The P-channel MOSFET

Note to 6.002x students (as I expect virtually all readers to be): this section is NOT covered on the final exam.

Remember that the static power for the inverter was divided into two cases: low input (MOSFET switch open) where static power was zero - excellent! - and high input (MOSFET switch closed), where the static power was very high - at least if we consider millions of gates. The reason that there was power dissipation when the MOSFET is closed is that there is a path for the current to flow from V_S to ground. If we could remove this path, for example by placing a switch in place of R_L, that was closed whenever the pulldown MOSFET was open, and vice versa, could solve our problem.

Fortunately, such a switch exist, and it's a different kind of MOSFET. The MOSFETs we've used so far are N-channel MOSFETs, named as such because the channel that forms between the source and drain terminals is made up of electrons - negative charge carriers.
The complementary MOSFET is known as the P-channel MOSFET. Here, as expected, the channel that forms is made up of positive charge carriers - "holes", which are the absense of electrons.



Before we discuss how the P-channel MOSFET works, let's review the N-channel.
As we remember, the N-channel MOSFET (in the SR model) is off/open when $v_{GS} < V_T$, and on/closed when $v_{GS} \geq V_T$, where v_{GS} is the voltage from gate to source, which is always positive. (If it's negative, we switch the drain/source terminals with each other, so that v_{GS} becomes positive.)

For the P-channel MOSFET to be complementary, or in fact for it to be useful here, it needs to have the opposite behavior, and it does. We still measure v_{GS}, but note in the schematic symbol above that the source is now "on top", and will connect to the power rail. Therefore, the source voltage is usually higher than the gate voltage (when both are taken with respect to ground), as opposed to the pulldown N-channel we've used in previous inverters, where we always connect the source to ground.
The P-channel MOSFET is on/closed when $v_{GS} \leq V_T$ and off/open when $v_{GS} >$V$_T$.

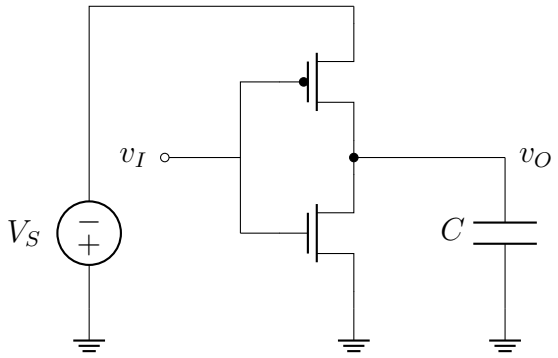
For example, let's say V_S is 5 volts, and we connect the source to V_S and the drain directly to ground (just for this experiment). If the gate is also connected to ground, v_{GS} will be $0 - 5 = -5$ volts. If V_T is, say, -1 volts, $v_{GS} \leq V_T$, so the MOSFET is on, and current flows through it. Clearly, for this configuration and these values, the MOSFET will be on as long as the gate voltage (with respect to ground) is less than or equal to 4 volts. If we make it 5 volts, however, v_{GS} will no longer be less than or equal to V_T, since we will have $v_{GS} = 5 - 5 = 0$.

8.3 CMOS logic

Note to 6.002x students (as I expect virtually all readers to be): this section is NOT covered on the final exam.

We can now begin using the P-channel MOSFET in our gates. Gates designed using a combination of P-channel and N-channel MOSFETs are known as CMOS gates, for Complementary MOS.
The introduction of CMOS was a revolution for computing and electronics in general, as it virtually eliminates static power entirely. In modern microprocessors, leakage current is one of the bigger problems, which only increases as transistors become smaller and smaller with increasingly advanced technology nodes. However, leakage current will not be discussed here.

Now, let's get on with it. Our new, nice CMOS inverter looks just about you'd expect: we replace R_L with a P-channel MOSFET:

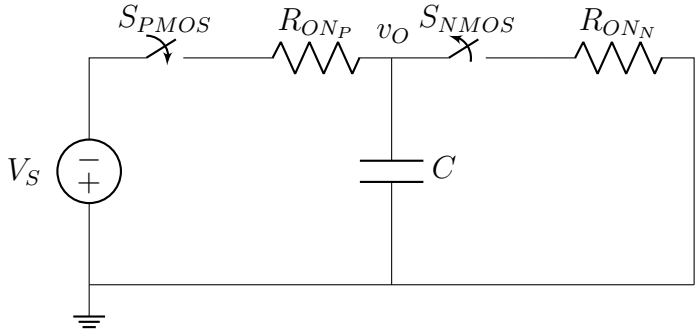


(Note that the capacitor meant to model the load capacitances etc. remains, as we will soon analyze the power usage for this circuit with those capacitances included.)

Since the one input goes to the gates of both MOSFETs, as long as we don't have v_I in the circuit's forbidden region (that is, we avoid $V_{IL} < v_I < V_{IH}$), the MOSFETs should never be on at the same time. Of course, if they were to be on at the same time, a very large current would flow through them (as long as R_{ON} is small for both MOSFETs), which is highly undesirable.

Now, let's analyze the power usage of this inverter! Due to the nature of the CMOS logic, there will never be a direct path from V_S to ground through a resistance (R_{ON}), only through the capacitor, which acts like an long-term open circuit. Therefore the resistance between V_S and ground is always infinite for the static power case (constant/non-switching input), and so the static power is zero! As mentioned earlier, due to leakage currents and other such effect (quantum effects and so on), this is not quite true. However, compared to the previous NMOS inverter with a > 6 kW power draw per 5 million gates, it might as well be zero.
So, if the static power is zero, let's calculate the dynamic power.

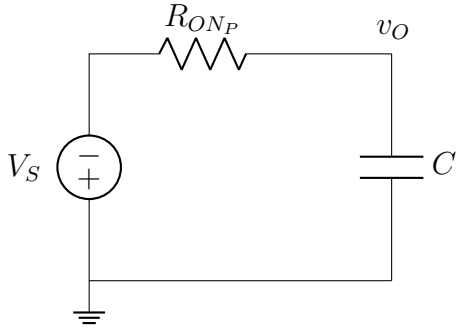
If we use the SR model for both MOSFETs, this circuit is fully equivalent:



Compare the two until you're convinced that they are equal!
Now, remember that the switches are never closed, and never open, at the same time. If the PMOS switch is closed, that means the NMOS switch is open, and vice versa.
Now, we will consider the case when this circuit is driven by a square wave, yet again with a 50% duty cycle. Let's call the time when the PMOS switch is closed T_1, and the time when the NMOS switch is closed T_2. $T = T_1 + T_2$ is then the cycle time, and $f = \frac{1}{T}$ is the switching frequency. Our goal is to find the average dynamic power. We'll start by considering the case where the PMOS switch is closed and the NMOS switch open, i.e. when the input is low and output is high (S_{PMOS} connects v_O to V_S, while S_{NMOS} is an open circuit), i.e. we'll calculate the energy supplied by the voltage source during T_1.

8.3.1 Energy supplied during T_1

Okay. So when the PMOS switch is closed, we can replace it with a short circuit, while the NMOS switch can be replaced with an open, which makes the entire NMOS fall out of the circuit:



Well, this looks simple enough. We know that for a series RC circuit with a step input, the waveform is either a rising exponential or a decaying exponential. If we assume that the initial capacitor voltage is zero, which we will (we'll assume that the switch is closed long enough for it to charge fully, and in the next step of analysis, the NMOS switch is closed long enough for it to *dis*charge fully), we get the two common forms for the capacitor voltage and current, respectively:

$$v_C(t) = V_S(1 - e^{-\frac{T_1}{R_{ON_P}C}})$$

$$i_C(t) = \frac{V_S}{R_{ON_P}} e^{-\frac{T_1}{R_{ON_P}C}}$$

We want to calculate the total energy supplied by the source, however, not the energy stored in the capacitor. We know that $P = VI$, but the V in question will be the supply voltage, not the capacitor voltage. We can however use the capacitor current, since this is a series circuit, and that current must be equal to the supply current.
Since the current is changing, we can't simply multiply the two. Instead, we need to integrate over the interval $[0, T_1]$:

$$E = \int_0^{T_1} V_S \cdot i_C(t) dt$$

$$E = \int_0^{T_1} V_S \frac{V_S}{R_{ON_P}} e^{-\frac{T_1}{R_{ON_P}C}} dt$$

We can move the fraction outside, as it's made up of constants only:

$$E = \frac{{V_S}^2}{R_{ON_P}} \int_0^{T_1} e^{-\frac{T_1}{R_{ON_P}C}} dt$$

To speed things up, we'll use the remembered result that

$$\int e^{ax}dx = \frac{e^{ax}}{a} + C$$

so:

$$E = \frac{{V_S}^2}{R_{ON_P}} \left[-R_{ON_P} C e^{-\frac{T_1}{R_{ON_P} C}} \right]_0^{T_1}$$

$$E = {V_S}^2 \left[-C e^{-\frac{T_1}{R_{ON_P} C}} \right]_0^{T_1}$$

If we evaluate this expression at 0, the exponential goes to 1, so we end up with

$$E = {V_S}^2 \left[-C e^{-\frac{T_1}{R_{ON_P} C}} + C \right]$$

$$E = {V_S}^2 \left[C \left(1 - e^{-\frac{T_1}{R_{ON_P} C}} \right) \right]$$

$$E = {V_S}^2 C - {V_S}^2 C e^{-\frac{T_1}{R_{ON_P} C}}$$

Phew. We can simplify the above, however. Remember that we said that we would let the capacitors change and discharge fully[1]. Therefore we can assume that $T_1 \gg R_{ON_P} C$ (since several time constants must pass during T_1 for it to discharge) which means the second term goes to 0, and all that remains is

$$E_{T1} \approx {V_S}^2 C$$

Aha! We now know the energy supplied by the source during T_1 (when the PMOS switch is the one that's closed). Now for some bad news.

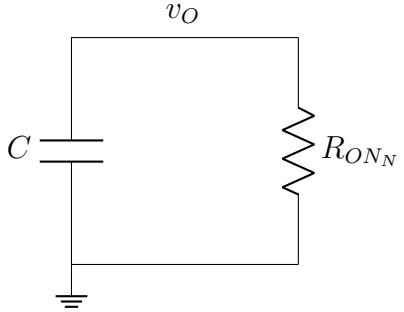
Recall that we let the capacitor charge fully. Also recall that the formula to calculate energy stored on a capacitor is $\frac{1}{2}CV^2$. Here, we charge it all the way to V_S, so the energy stored on the capacitor is $\frac{1}{2}C{V_S}^2$. However, if we compare this to E_{T1} above, this is only (exactly) *half* of the energy supplied! We could integrate the power dissipated by the resistor, or we could get the same result by using conservation of energy: the rest of the energy, i.e. half of it, must have dissipated through the resistor (i.e. the MOSFETs R_{ON_P})!

It appears that no matter what we do, half the energy will be lost as heat, while the other half remains usable in the capacitor. Bummer. Well, at least now we know why dynamic power even exists.

Next, let's find the energy dissipated during T_2.

8.3.2 Energy dissipated during T_2 and total energy used

Thankfully, this subsection will be much shorter than the previous one. With the PMOS switch now open, and the NMOS switch closed, the equivalent circuit is simply this:



[1] Mathematically, this would take infinite time. In practice, 5 time constants or so is usually close enough to consider the process pretty much done.

Again, remember that we said we would give the capacitor time to discharge fully. Therefore, all of the energy stored on it ($\frac{1}{2}CV_S{}^2$) must be dissipated through the resistor, and that's that.
Also, note that the source is not connected to this circuit, so it clearly doesn't draw any additional power. Therefore, the total energy supplied by the source is given by $E_{T1} + 0$, so $E = E_{T1}$.

Now that we know the total energy, we can calculate the average power. Power is just energy over time, so to find the average, we can simply divide the two:

$$\overline{P} = \frac{E}{T}$$

where $T = T_1 + T_2$. Since frequency is the reciprocal of T, the period, we can equivalently write

$$\overline{P} = Ef = CV_S{}^2 f$$

Ah! So we've finally found the dynamic power usage by this CMOS inverter, and therefore also the total power usage (since the static power usage is zero)!
Note that this equation for the CMOS inverter's power usage is the same as the equation for the old NMOS/resistor inverter's dynamic power usage. Therefore, we already know the power usage for our example scenario, but let's duplicate the results here. Again, we use $V_S = 5$ volts, $C = 1$ fF and $f = 3 \cdot 10^9$ Hz (3 GHz), and multiply that times 5 million gates:

$$\overline{P} = CV_S{}^2 f$$

$$\begin{aligned}\overline{P} &= 5 \cdot 10^6 \left[10^{-15} \cdot 5^2 \cdot 3 \cdot 10^9\right] \\ &= 5 \cdot 10^6 \left[75\ \mu\text{W}\right] \\ &= 375\ \text{W}\end{aligned}$$

Way, way better than the old inverter's 6.6 kilowatts, but still not acceptable. What can we do next? Well, there are several ways to reduce dynamic power, three of which are obvious from the above equation: we can reduce the load capacitance, reduce the supply voltage, and reduce the frequency.
Reducing the frequency might not be desirable, though, since that will reduce chip performance.

Let's have a look at reducing the power usage of this chip.

8.3.3 Reducing CMOS power usage

Reducing V_S looks like a good idea to begin with. In fact, the 5 volts we've used places our technology a bit in the past; the last "PC microprocessor" (that is, for regular home use) to use a 5 volt core voltage was the Pentium, around 1996. Modern CMOS chips use supplies as low as 1 volt or even lower. If we were to run our chip above on $V_S = 1$ volt, the power would come down all the way to 15 watts!
Of course, we can't simply reduce the supply voltage until we're at virtually zero. Remember that lower voltages lead to smaller currents, which lead to longer capacitor changing times, and thus, in the end, slower chip operation. In addition, things such as a MOSFETs V_T can't go arbitrarily low. Still, 1 volt is enough for modern chips to run as fast as they do.

One of the main ways microprocessors go forward is by reducing the size of the chips. This is referred to as "shrinking". For example, perhaps the main difference between 2011's Sandy Bridge (the codename for Intel's Core i3/i5/i7 2xxx CPUs) and 2012's Ivy Bridge (Intel Core i3/i5/i7 3xxx) is a die shrink, from 32 nm to 22 nm. That size (22 nm) refers to the approximate size of the smallest transistors in the chip. The full definition is slighty complex.
When doing such die shrinks, the gate capacitances tend to decrease, which leads to a reduction in dynamic power. In general, die shrinks also allow for a decrease in the supply voltage.

What about the frequency? Well, as said before, we don't want to simply lower that too much, or performance will suffer. What we can do, and what is actually done in modern chips, is to vary it over time. For example, while the computer is idle, or doing something that is not very computationally intensive (web browsing, text editing), the processor can reduce its frequency, and only stay at the full frequency when really needed. Lower frequency also means the supply voltage can be reduced during those idle times, which further reduces power usage!

Yet another way to reduce power usage is *clock gating*. This technique can power down parts of a chip that aren't currently used. For example, the floating point-unit of a microprocessor might be currently unused, so we can disable the clock signal to it.
Doing so is simple: we can (for some designs) use an AND gate, and connect the clock to one input of the AND gate, connect an enable line to the other, and the gate's output to the floating point-unit's clock input.
When the enable line is low (0), the AND gate will produce a 0 on its output no matter whether the clock signal is high or low.
What the enable line is high (1), the AND gate will reproduce the clock signal on its other input (as $A \cdot 1 = A$, the clock signal is passed through unchanged in this case).

8.4 CMOS logic gate design

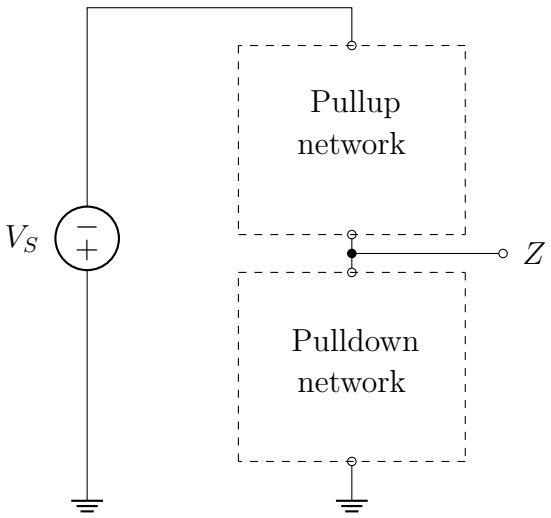
Note to 6.002x students (as I expect virtually all readers to be): this section is NOT covered on the final exam.

This section doesn't really belong in this chapter, but it'll have to stay here for the moment. If I add a proper chapter on the digital abstraction, I'll move it there at that point.
Using what we already know about digital gates and CMOS logic, we just need a small guideline to make designing CMOS gates easy.
In our old NMOS gates, we had a simple resistor as a pullup, and N-channel MOSFETs as pulldowns. All we had to do was to configure the pulldown network, and we'd be done.
For CMOS gates, we also need to configure the pullup network. However, doing so is rather simple.

If we want to implement the logical function F, perhaps NAND, so that $F = \overline{A \cdot B}$, we design it so that:

The pullup network is a short when F is true, but an open otherwise

The pulldown network is a short when $\overline{F}$ is true, but an open otherwise
We also only use P-channel MOSFETs for the pullup network, and only use N-channel MOSFETs for the pulldown network. That's all we need to do. We then insert those two networks into the "CMOS gate template", which looks like this:



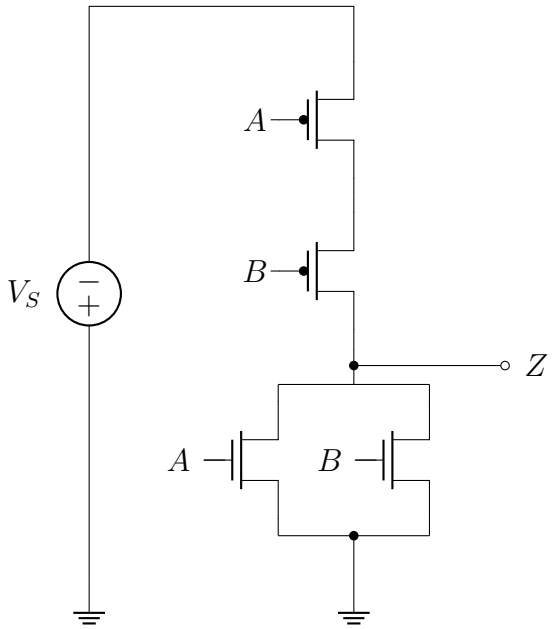
... where Z is the logic function's output. The inputs are connected to the MOSFETs in the pullup and pulldown networks, of course.
Remember that this topology is inverting by its very nature, so to make non-inverting functions such as AND/OR, we would make a NAND/NOR gate and then send that output to a NOT gate.

Let's use what we now know to design a NOR gate. For NOR, the function is $F = \overline{A + B}$, and the complement is $\overline{F} = \overline{\overline{A + B}} = A + B$.

We begin with the pulldown network, since we have designed NOR gates in the past. The pulldown remains the same as before, but let's go through the method anyway.
We need transistors such that it is a short circuit when $\overline{F} = A + B$ is true. That's just a regular OR function, which is true whenever either or both inputs is true, so we use two parallel N-channels here, so that it's a short when either input is true, but an open circuit when both are false.

What about the pullup? Here, we need a short circuit when $F = \overline{A + B}$ is true. This is (as the gate will be) NOR, which is true when both A and B are zero, and never otherwise. Therefore, this pullup network needs to be such that it's only a short circuit when $A = B = 0$, which we can do with two P-channel MOSFETs in series. (Remember that we only use P-channels for the pullup.)

That concludes our gate design - all that remains is to display the result:



And we are done! The same method can be used to design NAND and NOT gates, which we can then combine to form any boolean logic function. (In fact, having nothing but NOR gates still makes that possible!)

We can also make gates with more than two inputs this way. To turn the above into a 3-input NOR gate, we would simply add a "C" P-channel MOSFET in series with the other two pullups, and a "C" N-channel in parallel with the others in the pulldown network.

Chapter 9

Miscellaneous

9.1 Impulses and steps

9.1.1 A note

For a linear circuit - one with resistors, capacitors and inductors *ONLY* - no sources allowed - an interesting relation between input and output exists. If y(t) is the output for the input x(t), the output for the derivative of x(t) will be the derivative of y(t). The same goes for integration. That is,

$$x(t) \to y(t)$$
$$x'(t) \to y'(t)$$
$$\int x(t)dt \to \int y(t)dt$$

The usefulness for this will soon be clear. In a short preview, we can calculate the behaviour of such a circuit to for example an impulse, and then integrate the response we got; the integrated answer will then be the circuit response to a *step*, as a step is the result of integrating an impulse.

9.1.2 Impulses

Impulses are (theoretical) bursts of an infinite voltage/current over an infinitesimally small time. The Dirac Delta function, $\delta(t)$, is used to describe them. The delta function is defined as[1]

$$\delta(x) = \begin{cases} +\infty & x = 0 \\ 0 & x \neq 0 \end{cases}$$

In other words, it is zero everywhere except at $x = 0$, where it's infinite. There is an additional identity it is constrained to, however:

$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

In other words, the area under the curve is exactly 1. Mathematically, the delta function is not a proper function, but can be defined as a distribution.
In a parallel, current source-driven RC circuit, a current impulse at $t = 0$ delivers 100% of its charge to the capacitor, and essentially creates an initial condition for $t = 0^+$. Assuming the capacitor has no state, the capacitor voltage after the impulse will be

$$v_C(0^+) = \frac{Q}{C}$$

[1] Perhaps not rigorously, but good enough for our purposes.

where Q is the area of the impulse, i.e. the current source output is given by

$$Q\delta(t)$$

, so Q is in coulombs, the SI unit for electric charge. The unit is such that

$$\frac{1 \text{ coulomb}}{1 \text{ farad}} = 1 \text{ volt}$$

In such a circuit, the capacitor voltage (and thus the entire circuit's voltage, as there is only 1 node besides ground) for $t > 0^+$ will be given by

$$v_C(t) = \frac{Q}{C} e^{-\frac{t}{RC}}$$

The dual of this circuit is the series RL circuit, with a voltage source, a resistance and an inductance. In this case, we have a voltage impulse, that delivers a flux linkage Λ to the inductor. As above, this will essentially create an initial condition, this time for the inductor current. Assuming the inductor has no state, the current through it will be:

$$i_L(0^+) = \frac{\Lambda}{L}$$

where Λ is the area of the impulse, i.e. the voltage source output is given by

$$\Lambda\delta(t)$$

, so Λ is in webers, the SI unit for magnetic flux. The unit is such that

$$\frac{1 \text{ weber}}{1 \text{ henry}} = 1 \text{ ampere}$$

Again, an in the previous case (the circuits are duals, after all), the inductor current for $t > 0^+$ is given by

$$i_L(t) = \frac{\Lambda}{L} e^{-\frac{Rt}{L}}$$

9.1.3 Steps

Steps are the result of integrating an impulse. Steps are likely more intuitive than impulses. The unit step function (or the *Heaviside step function*) has the definition

$$u(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

In words, it "turns on" at $t = 0$, and so is a useful model for things such as a voltage source turning on at $t = 0$. Such a source could be written as

$$V_I u(t)$$

where V_I is the voltage of the source when turned on.

As noted in the introduction section, a step is the result of integrating an impulse. Thus, we can calculate the circuit response to a step by calculating the response to an impulse, and then integrate the answer. In a similar manner, we can calculate the response to a step, and differentiate the answer to find the response to an impulse.

9.1.4 Shifts

The above functions can be shifted to start at times other than $t = 0$ by shifting them.
For example, while the unit step function $u(t)$ "turns on" at $t = 0$, we can make it turn on at time $t = 5$ by shifting it, like so: $u(t - 5)$
The above function will have $t - 5$ be negative until $t = 5$ where it becomes positive, and the unit step function changes its output from 0 to 1.
The exact same principle applies to the unit impulse function $\delta(t)$.
In general, for the change to happen at time $t = T$, use $u(t - T)$ or $\delta(t - T)$.