Electronics for Pros

Holger Müller, certified analog wizard

Spring 2020 notes

1 Introduction

2 Passive Components

Passive components are defined by a linear voltage and current relationship.

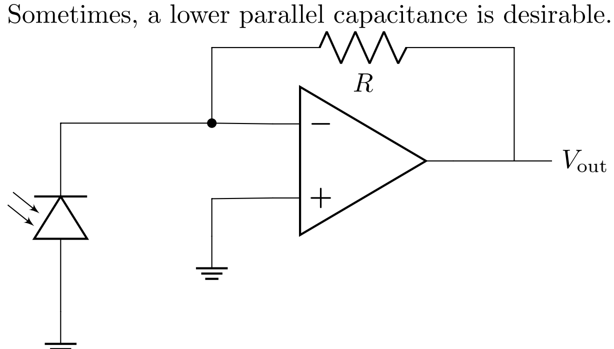
2.1 Fixed Resistors

Resistors are generally well behaved, meaning they approximate the ideal resistor very closely in practice. The key specifications are power (in watts), resistance (in ohms), the tolerance (usually 1%, but sometimes better). Resistors come in logarithmically spaced values. The E96 series is the most common, divides each decade into 96 logarithmically spaces values.

The most common resistors are metal film resistors. They are close to ideal resistors. They have some parallel capacitance, but this is usually as low as $C_p = 0.05\,\mathrm{pF}$. In order to actually get a capacitance this low, attention must be payed to layouts: the leads on the resistors should be straight.

2.1.1 Reducing Parallel Capacitance

Sometimes, a lower parallel capacitance is desirable. Consider the circuit



With parallel capacitance, the upper frequency limit of the circuit is

$$f_{\text{up}} = \frac{1}{2\pi R C_p}$$

For large R, this can become problematic if you want a fast circuit. Reducing R is difficult due to small resistors having larger current noise. C_p can be reduced by wrapping grounded copper tape around the resistor, which will break electric field lines between the two leads that are causing parasitic capacitance. This effectively makes many small capacitors to ground between each of the pieces of copper tape. To minimize the effect of the new capacitors, the tape should be fairly thin. Capacitors to ground hurt the circuit less because they load the output of the op-amp to ground, and at the output there is enough current to properly charge the capacitors. Note that this is especially effective in this op amp circuit, and may not be a good idea in other places.

2.1.2 Resistor Types

For a metal film resistor, the resistance can range from milliohms to gigaohms. The power is typically 1/4 W, but they are available at up to 100 W.

For higher powers, a wire wound resistor is used. This is literally a long, wound wire, which leads to large inductance, so be careful at high frequency.

Carbon film resistors are bad. They have substantial $1/f$ noise, in addition to the Johnson noise in all resistors. Bulk carbon resistors are good for pulsed power dissipation.

2.1.3 Form Factors

SMD resistors are fairly common. They are good for PCBs, and painful to solder by hand elsewhere. The most common size is 1206, which means the length is 0.12" and width is 0.06". This is the smallest size that can be hand soldered. SMD resistors have higher parasitic capacitances than through hole resistors, but allow more compact circuits to be constructed. SMD resistors do have a lower inductance than through hole resistors. For a PCB trace, stray inductances are approximately a few nanohenries per centimeter.

Through hole resistors have parasitic capacitance due to capacitance between the leads, capacitance between windings, and capacitance between internal caps. The capacitance between windings is distributed, which leads to the impedance decreasing as $1/\sqrt{\omega}$. However, you should assume that it will be simply a resistor and capacitor in parallel. If you need something better than that, you should consider redesigning the circuit.

When impedances are larger than $100\,\Omega$, stray capacitance is more of a problem. When impedances are smaller than $100\,\Omega$, stray inductance is more of a problem.

2.1.4 Potentiometers (Adjustable Resistors)

Potentiometers should be avoided. In theory, the issue is that the wiper can become dirty. In lab, the issue is that things that can be adjusted will be adjusted, which might mess up an experiment.

There are two main types of small potentiometers: the longer (20 turn) versions are better than the smaller 5 turn versions. Small potentiometers are not very stable. There are single turn and 10 turn front panel potentiometers. Single turn potentiometers are fine if precision is not needed. 10 turn potentiometers are expensive, but very stable. They also come with locking knobs. Note that 10 turn potentiometers are wire wound resistors, and thus have a large inductance. They should be avoided at high frequencies. 10 turn potentiometers might be usable to 1 MHz by using values between 1 kΩ and 10 kΩ.

We note that these should probably be avoided on drones. In general, trim potentiometers should have no more range than needed.

Only use pots for good reasnos. Bad reason to use a pot: I'm too lazy to calculate what resistor I actually need. Good reason: setting needs to be changed during normal operation. In either case, you don't want to give the pot more range than you actually need because more range means less stability.

2.1.5 NTC and PTC

NTC and PTC are temperature sensitive resistors. NTC: negative temperature coefficient. PTC: positive temperature coefficient. Standard component has 10 kiloohm resistance at room temperature. You can use these to temperature compensate a circuit with a known temperature dependence.

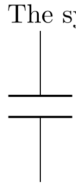
2.1.6 VDR

A VDR is a voltage dependent resistor. They are used as surge protectors. They have a high resistance at low voltage and are almost short circuits at high voltage.

They can also be used to absorb energy from large inductances. For example, a VDR can be placed in parallel with a MOSFET that is switching a magnetic coil. The current-voltage curve of the VDR is similar to that of a Zener diode. A VDR is preferable for this because they have a larger volume over which to dissipate energy. However, a Zener diode is more precise, so only use a VDR if you need to dump a lot of energy quickly.

2.2 Capacitors

The symbol for a capacitor is



2.2.1 Ceramic Capacitors

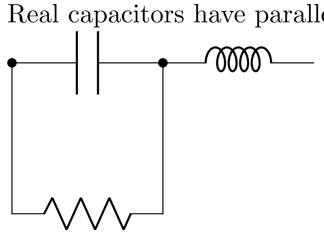
Ceramic capacitors are the most common. SMD capacitors are preferable, because parasitic capacitance is irrelevant. Otherwise, leads should be kept short.

Values from 100 pF to 1 nF are very good, especially NP0 capacitors, which have zero temperature coefficient.

Ceramic capacitors larger than 100 nF have many problems. Poor tolerance, high temperature coefficient, and they can be piezoelectric (if you flick them you can see voltage spike of a couple volts...). They do have low inductance and low leakage. They can be used as bypass capacitors (e.g. for connecting op amp power supplies to ground), since for this application you don't care about tolerance, temperature dependence, etc.

2.2.2

Real capacitors have parallel resistance and series inductance:



Ideally, the impedance of a capacitor is $Z = \frac{1}{\omega C}$. However, the series inductance leads to the impedance increasing at high frequencies. Use short leads to minimize parasitic inductance.

2.2.3 Other Types

Film capacitors are the more expensive versions of ceramic capacitors. They have good tolerance, low leakage, low temperature coefficients, and they are not piezoelectric. Film capacitors do have somewhat higher inductance. Film capacitors are more precise, and should be used in filters and integrators. Above 1 microFarad, film capacitors are very bulky.

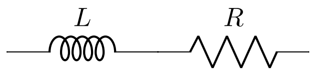
Electrolytic capacitors are used for larger capacitors. These are polarized, and will explode if reversed. There are two versions. Aluminum electrolytic (up to 1 F), which have poor tolerance, high leakage, high temperature coefficients, and a limited lifetime. Aluminum electrolytic capacitors are usually used in power supplies. They also have a relatively high series resistance, which will heat the capacitor and lead to a shorter lifetime (1000s of hours). Inside they have a roll of paper with aluminum on each side, which has terrible thermal conductance as a whole. Tantalum electrolytic capacitors have better tolerance, low leakage, and very low inductance, but are very expensive, so only use if you know why you are using it!

2.2.4 Variable Capacitors

Variable capacitors are also known as trimmers. They usually have capacitance less than 100 pF. Large bulky ones exist up to 1 μF. These should be avoided.

2.3 Inductors

Inductors are the most complicated passive component. They are rarely near ideal, almost always having a series resistance.



The Q of an inductor is given by

$$Q = \frac{i\omega L}{R}$$

We typically want to maximize this.

There is also a parallel capacitance, which leads to a self resonance, where the impedance increases rapidly. Above the self resonance frequency, the capacitance dominates and the impedance decreases. Parallel capacitance emerges from both capacitance between the leads and capacitance between windings. If one end of the inductor is grounded, the inductor can be wound as a cone. There is more inductance and more capacitance towards the bottom of the cone, but since the potential at the bottom is smaller, this is less important.

2.3.1 Air Core Inductor

For an air core, wire wound inductor of length l and diameter d with n turns

$$L = \frac{d^2 n^2}{18d + 40l}$$

(when $l > 0.4d$) when L is in microhenries and d and l are in inches. This equation replaces the standard physics equation for an inductor when the length becomes comparable to the diameter.

The Q of an inductor is $\omega L/R$. We want to maximize Q, which implies $l \approx 0.4d$. This is because $R \propto l$. At high frequencies, the resistance is increased due to the skin effect. Due to the skin effect, only a depth of $\frac{64\,\mu\mathrm{m}}{\sqrt{f/\mathrm{MHz}}}$ is used.

Inductors generate stray magnetic fields. These induce currents in other places, which will reduce Q as well as cause issues with the rest of the circuit. Thus, an air core inductor should be placed in a conducting box of at least $3-5l$ by $3-5l$ if you are concerned about coupling to other parts of your circuit.

Air core inductors are good at high frequencies for small inductances. They have some issues. Air core inductors pick up external magnetic fields. Air core inductors are limited in size.

The upper frequency limit depends on parallel capacitance C_{par}. Above the frequency $f_T = 1/(2\pi L C_{par})$, the inductor will begin to behave like a capacitor with impedance proportional to $1/f$. The capacitance to ground can be aleviated somewhat with cone shaped inductors. The idea is that capacitance to ground is a larger problem in the high voltage end of the inductor, so the high voltage end is at the tip of the inductor so the parasistic capacitances on this end are smaller. The larger end of the inductor will have larger parasitic capacitances, but the voltage difference to ground is smaller so they have less of an impact on speed.

2.3.2 Toroidal Inductor

To fix the issues with air core inductors, we can wrap our wire around a toroid of ferrite. The inductance is

$$L = A_l n^2$$

A_l is a parameter of the ferrite toroid (you can order toroids with this value specified). We want materials with a higher saturation field, B_{sat}. Iron has a higher μ_r and B_{sat}, but iron conducts, which leads to eddy currents and a decrease of Q. Ferrite is a worse conductor but has lower saturation field strength.

The geometry of toroids makes them less sensative to stray magnetic fields.

2.3.3

Inductors can also be bought, but these tend to be maximally sensative to stray fields.

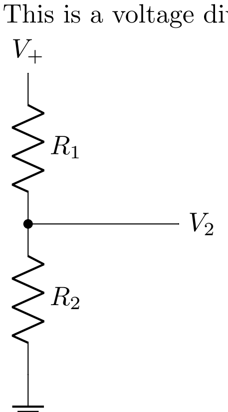
2.3.4 Mutual Inductance

For an inductor, the induced voltage is $V = L \cdot \dot{I}$. For two inductors sharing a magnetic field, there is a mutual inductance M. For a current I_1 in the first inductor, the voltage across the second inductor is $V_2 = M\dot{I}_1$.

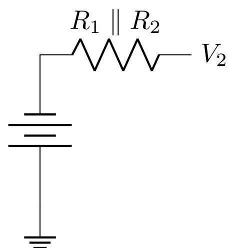
2.4 Simple RLC Networks

2.4.1 Voltage Divider

This is a voltage divider:



The smaller R_2, the small V_2. We could consider this to be a voltage source:

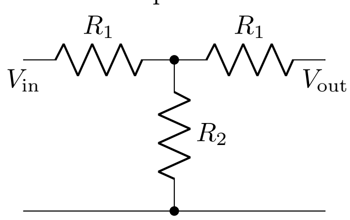


where

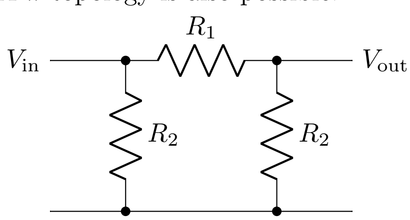
$$R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2}$$

2.4.2 Attenuator

An attenuator is similar to a voltage divider, but we add another resistor to match the impedance.

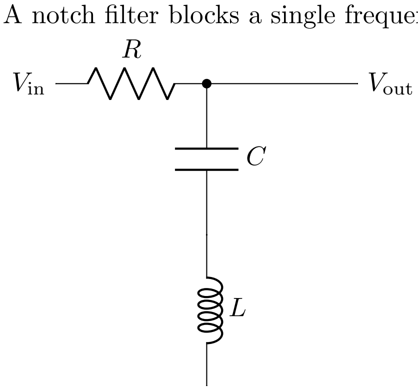


A π topology is also possible.

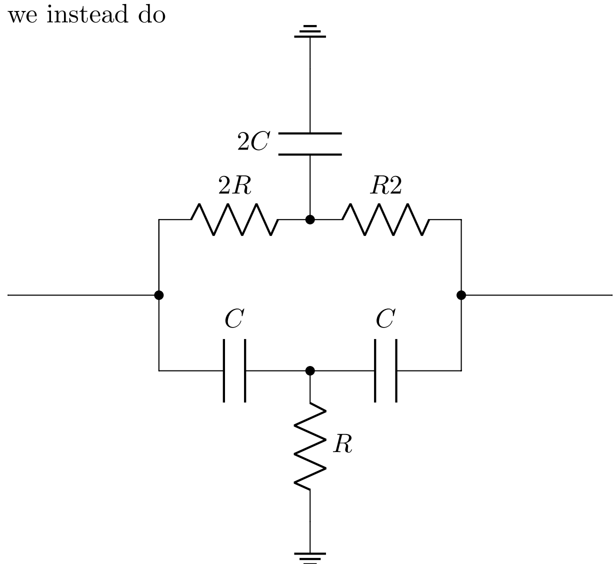


2.4.3 Notch Filter

A notch filter blocks a single frequency. The naive way to build this is



However, inductors are not good, especially at frequencies below 1 MHz. Thus, we instead do



The notched frequency is

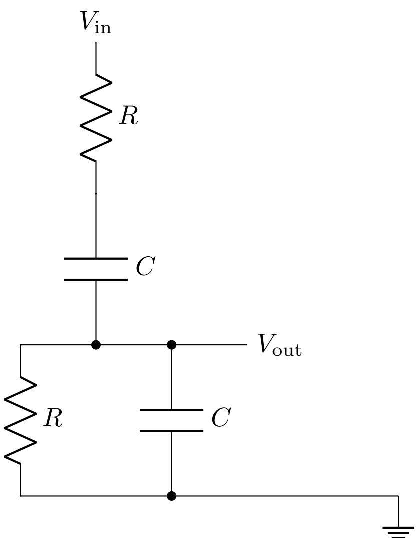
$$f_N = \frac{1}{4\pi RC}$$

The resistors and capacitors should be well matched (with a multimeter). This means you need stable capacitors, so avoid high capacitance ceramic and instead opt for film capacitors. This filter can achieve over $30 - 40\,\mathrm{dB}$ of attenuation if components are matched well.

Notch filters can be used to compensate for resonances, but this is not a silver bullet. Typical application would be to filter 60 Hz noise, or filter out mechanical resonances in a feedback loop. This circuit as drawn cannot be adjusted, but there exists a design for an adjustable notch filter using a single pot.

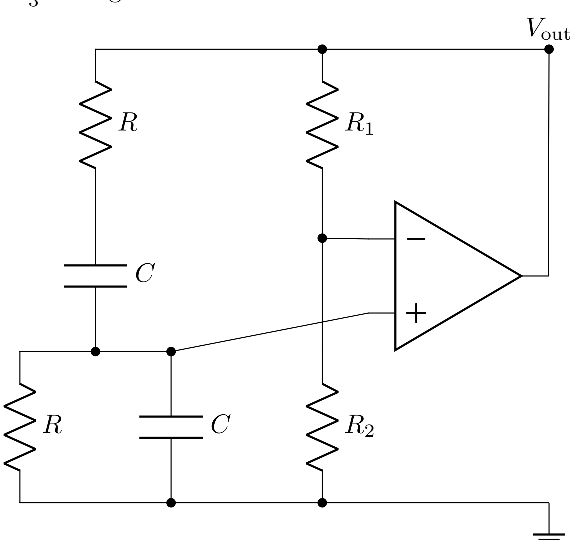
2.4.4 Notch Amplifier?

To produce a gain at a single frequency, a notch filter can be placed in a feedback loop. The Wien bridge is an example of this.



The voltage at V_{out} is $\frac{1}{3}$ of the input voltage.

With an op amp, we can produce oscillations at frequency $\frac{1}{2\pi RC}$ by adding a $\frac{1}{3}$ voltage divider.



However, the oscillation amplitude will grow without bound, until the op-amp is saturated and the output becomes a square wave. By substituting the lower resistor with a lightbulb, the oscillation amplitude can be constrained (as it will oscillate less effectively as the amplitude changes).

By making the lower resistor adjustable, we can make an adjustable notch amplifier.

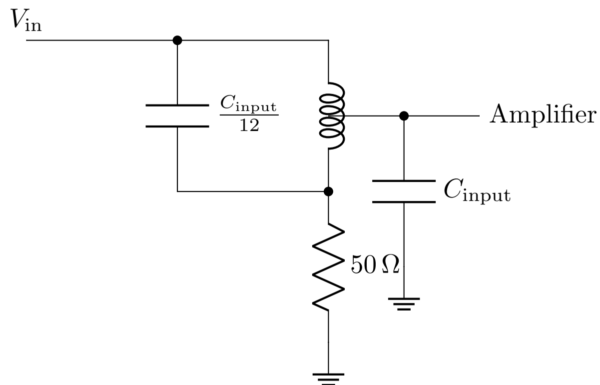
2.5 Passive Component Networks

At high frequencies, there tends to be a "hidden circuit." For example, there is always a capacitor in parallel with a resistor.

2.5.1 T-Coil Circuit

Suppose we have a $50\,\Omega$ radio frequency source connected to an amplifier. There will be a capacitor between the amplifier input and ground in parallel with the $50\,\Omega$ input impedance of the amplifier. At a frequency $f = \frac{1}{2\pi RC}$, there will be a $3\,\mathrm{dB}$ attenuation of the signal.

Consider a center-tapped inductor, with a mutual inductance M between the two sides.



With this circuit, the input impedance is always $50\,\Omega$, and there is a 2.83 speed up in the charging time of C_{input} (which is the voltage the amplifier sees) for current sources.

⟨Need to add explanation of how this works...⟩

Let each half of the coil have inductance L, mutual inductance M, and R be the general resistance ($50\,\Omega$ above). Then we want

$$L = \frac{1}{2}R^2C \quad M = \frac{1}{2}R^2C$$

and this results in

$$f_{3\,\mathrm{dB}} = \frac{2.83}{2\pi RC} \quad t_r = 0.79RC$$

where t_r is the charging time.

If there is parasitic inductance in series with C_{input}, we can increase the mutual inductance of the two half inductors to compensate.

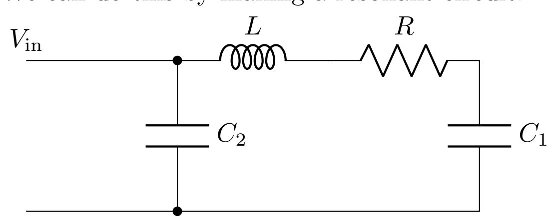
To build a large amplifier, we need large components, which have large C_{input}. Instead, the T-coil network can be used to build a **distributed amplifier**, where we chain together many t-coils and MOSFETs in order to amplify in parallel and reduce the load on each component.

We can add a T-coil network at the input of our amplifier (gate of a MOSFET). At input of the T-coil network looks like a resistor, so we can put a second T-coil network at the input of the first. ⟨Need to add schematic here...⟩ This allows us to drive several MOSFETs with the same input. There will only be a delay between the different MOSFETs. We can then add T-coil networks at the drain of the MOSFETs, which allows us to add the signals. The source of the MOSFETs is grounded.

A MOSFET has a gate-source, drain-source, and gate-drain capacitors. The gate-source capacitor is denoted C_{iss} and drain-source capacitor is denoted C_{oss} in datasheets. In the distributed amplifier, the gate-drain capacitor must be negligible. The T-coil networks cancel C_{iss} and C_{oss}.

2.5.2 Tank Circuit or Pi Network

We often have a $50\,\Omega$ source and a capacitive load C. Suppose we want to operate at one frequency and maximize the voltage across the capacitor. For example, this would be the case for driving an EOM from a function generator. We can do this by making a resonant circuit.



Note that this is an inductor in parallel with a capacitor $C_{\text{eff}}^{-1} = C_1^{-1} + C_2^{-1}$. The resonant frequency is

$$f = \frac{1}{2\pi\sqrt{LC_{\text{eff}}}}$$

The voltage across C_1 is $V_{\text{in}}\frac{C_1}{C_2}$. We want to maximize the power disappated in R. Thus, we want to match the impedance to that of the input. The maximum amplification is $\sqrt{Q}$, for Q of the inductor.

We usually do not know the values exactly, and thus can not calculate the values for components. We start with $C_2 = 10C_1$, then optimize to increase the voltage across C_1. However, measuring that voltage is difficult for small capacitors (since oscilloscopes have input capacitance). Thus, it is better to have the oscilloscope probes close, but not connected. Due to capacitive coupling to the oscilloscope probes, the scope will see some small fraction of the actual voltage across the load capacitor, but this signal can still be used to find the maximum voltage on the load when swapping out the C_2 resistor.

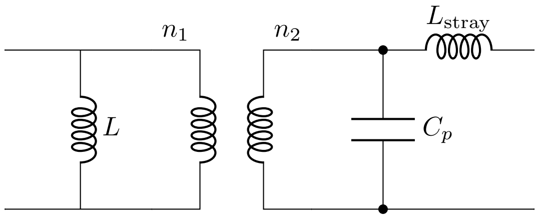
A power splitter can also be used. A power splitter can be thought of as a beam splitter that passes, e.g. 90% of the input to the output, and reflects 10% of the input to the coupling port. We can use one in reverse. When the impedance is matched, there will be no reflected power, and thus no power will come from the coupling port. So, we can connect a power splitter before the input of this circuit which we use backwards so that the coupling port picks up signal reflected from the capacitive load. Then you tune the circuit to minimize the amplitude of the signal at the coupling port.

2.6 Transformers

A transformer is two coupled coils. Let the winding on the left have n_1 coils and the winding on the right have n_2 coils. Then

$$V_2 = \frac{n_2}{n_1} V_1 \quad I_2 = \frac{n_1}{n_2} I_1$$

That is for an ideal transformer. In reality, we tend to have



In theory, the lower frequency limit would be $f_{\text{low}} = \frac{1}{2\pi Z_i L}$, where Z_i is the input impedance. In practice, the magnetic field in the core will saturate when the frequency is too low (the inductor can be considered to integrate the incoming voltage), at which point the inductance drops to zero, leading to a short. This results in a voltage drop each time the current peaks (near the node of the voltage wave), causing crossover distortion.

There is also a high frequency limit. For iron core transformers, laminated transformers are usually limited to 10 kHz. Toroidal transformers have a range of 60 Hz to 100 kHz.

2.6.1 Transformers for Noise Reduction: Ground Loops

Consider an measuring instrument in a lab. The measurement instrument will be grounded. Suppose we have a signal generating instrument, which is also grounded. We then connect a signal from the signal generating instrument to the measurement instrument. We usually also connect a ground between the instruments. The lab is full of changing magntic fields, so by having this ground connected between the instruments, which are also connected to ground, we have a (very large) loop, with a changing magnetic field. This is a ground loop. The changing magnetic field will generate AC currents in this loop. Since the signal connection has some resistance, we get an (unpredictable) AC voltage on the signal ground line. For a small signal (mV), this is quite problematic.

There are some strategies for mitigating ground loops

- We can reduce the impact of ground loops by reducing the resistance of our signal cables. Connecting both instruments to an optical table might provide a very low resistance path.

- We can also cut one of the other grounds. However, it is dangerous to cut the grounds on instruments. We can use an isolation transformer in order to power the instruments, which has a similar effect.

- We can float the power supplies of each device. Most power supplies have transformers, so the ground can be separated.
- Batteries will break ground loops. This works well for photodiodes.
- Amplifying signals earlier will reduce the impact of ground loops.
- Finally, we can use a transformer to break a ground loop. With no connection in the signal path, there is no longer a continuous loop. This will only work if the signal does not have (interesting) DC components. Some transformers have distortions, but you can buy transformers that will work over some range. This is usually done in audio, XLR cables are ungrounded and there are signal transformers in most devices.
- A differential amplifier can be used in place of a transformer when DC is needed. Differential amplifiers tend to be less good at high frequencies, in particular, the common mode rejection decreases. There also tends to be a limit between the input and output voltages in a differential amplifier. Differential amplifiers might offer higher signal integrity, minicircuits has good options for this.

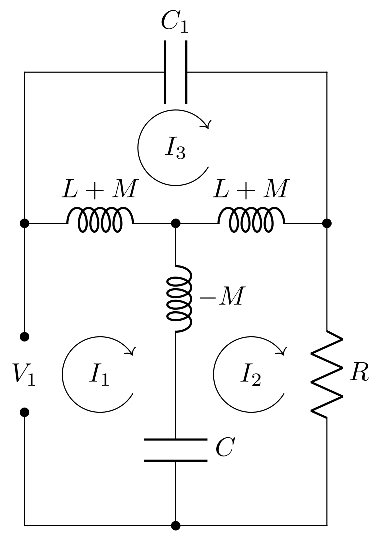
The skin depth of a conductor is less than $64\,\mu$m: it scales as $\delta = 64\,\mu\mathrm{m}/\sqrt{f/(1\,\mathrm{MHz})}$. At frequencies above $1\,$MHz the noise is no longer able to penetrate electronics boxes or BNC cables. Thus, we do not need to worry about ground loops for high frequency magnetic fields. This also implies that we should never use insulated BNC connectors for RF signals ($> 1\,$MHz), as this brings the surface currents inside our nice shielded box. For low frequency signals, insulated BNC connectors may be useful in order to make it easier to connect a transformer inside the box.

2.7 Network Analysis

There are two main methods: nodal analysis and mesh analysis.

2.7.1 Example: The T-Coil Circuit

We will consider the T-coil circuit (this is an equivalent circuit, because analyzing mutual inductance is complicated).



with I_1, I_2, and I_3 being clockwise currents. Each current loop is known as a mesh. Let $p = i\omega$. Using Kirckoff's laws, we can setup an equation for the potential in terms of the currents.

$$V_1 = I_1 \left[p\,(L+M) - pM + \frac{1}{pC} \right] - I_2 \left[\frac{1}{pC} - pM \right] - I_3 \left[p\,(L+M) \right]$$

We can then construct two more equations and put them into a matrix:

$$\begin{bmatrix} p\,(L+M) - pM + \frac{1}{pC} & -\frac{1}{pC} + pM & -p\,(L+M) \\ -\frac{1}{pC} + pM & p\,(L+M) + R - pM + \frac{1}{pC} & -p\,(L+M) \\ -p\,(L+M) & -p\,(L+M)\,\frac{1}{pC_1} + 2\,(L+M) & \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} V_1 \\ 0 \\ 0 \end{bmatrix}$$

The diagonal elements are the sums of impedances in a mesh. The off diagonal elements are negative the impedance common to two meshes, and this means that the matrix is symmetric. We can solve this with a computer, for example, in Mathematica. ⟨I am not going to try to write this down.⟩

This will give us solutions as a function of frequency. Bode plots are very useful for visualization of the circuit behavior in frequency space, where the magnitude and phase of the response are plotted next to each other. The poles p of the response function give the complex frequency of resonances in the circuit, with the imaginary part of p corresponding to the frequency of the resonance (since $p = i\omega$) and the real part gives the damping rate. Using a Laplace transform, we can determine the time domain behavior as well. To obtain the impulse response in the time domain, multiply the voltage response by $1/p$ (the Laplace transform of a unit step impulse). Taking the Laplace transform of this gives a series with terms e^{pt} weighted by the residues of the poles p. Note that the real part of p is negative for passive components, otherwise one would have unbounded oscillations.

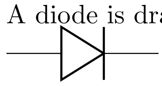
Circuit simulation software will do this automatically, so this rarely needs to be done in practice. SPICE is the most common simulation tool.

3 Diodes

While we can do interesting things with passive components, the real power in electronics is from active components. Usually, this only includes transistors, but we will consider diodes to be active components as well. This is justified due to the ability of diodes to change the spectrum of a signal. For example, one can create a square wave from an input sine wave, therefore genearating new frequency components at $3f$, $5f$, etc. A diode can be configured to act as a frequency doubler. Diodes can function well at 10s of GHz.

3.1 Diode Properties

A diode is drawn



The anode is on the left, the cathode is on the right. The cathode is marked with a ring for through hole components and a stripe/dot for a SMD diode. The nomenclature is based on vacuum tube diodes. A diode consists of a pn junction, p at the anode and n at the cathode.

For a positive voltage applied from anode to cathode, a current can begin to flow at about 0.6 V for a silicon diode or 0.3 V for a germanium diode (due to the lower band gap in germanium). We usually use silicon diodes: germanium diodes have larger reverse current, they are more temperature dependent, but they have a lower threshold voltage.

A plain diode in a glass package will be light sensitive. The smaller band gap in germanium makes it more suitable for infrared photodiodes, whereas silicon diodes are used for visible and UV.

The current in a diode is

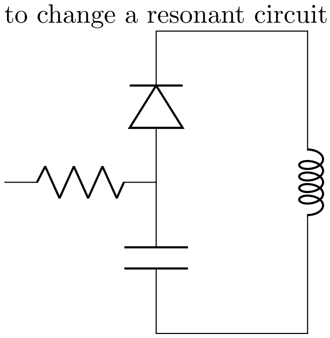
$$I = I_0 \left(e^{V/V_t} - 1 \right)$$

where V_t is the thermal voltage,

$$V_t = \frac{k_B T}{e} \approx 25\,\mathrm{mV}$$

The material properties are in I_0. I_0 is much larger for germanium. Then the current becomes too large, you get deviation from this exponential behavior because of serier resistance.

The normal specifications are the maximum DC current $I_{\max}$, which ranges from a few milliamps to thousands of amps, and the maximum voltage, which ranges from 5 V to thousands of volts. There is a maximum reverse voltage, the breakdown voltage. The reverse capacitance C_r ranges from 0.1 pF (for small RF components) to a few nanofarads for large diodes. In RF mixers, this capacitance is parasitic and is typically the limiting factor in performance. On the other hand, varactor or varicap diodes exploit this effect as a voltage tuneable capacitance. C_{AC} is maximal at zero reverse voltage, and decreases as the voltage increases. This can be used in a voltage controlled ocillator (VCO) to change a resonant circuit:

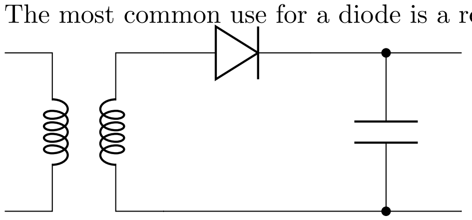


where the resistor is large.

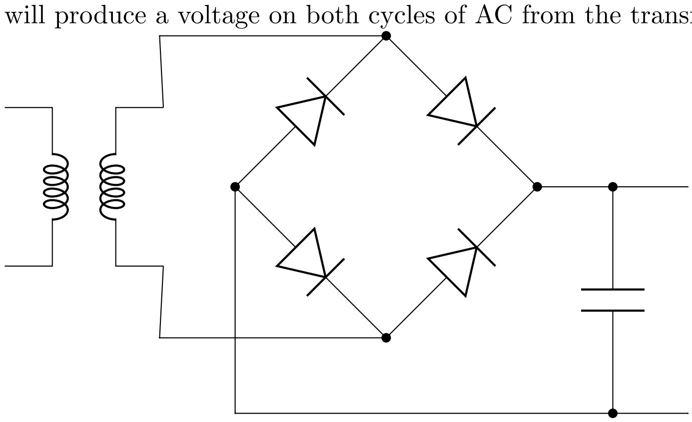
3.2 Diode Circuits

3.2.1 Rectifier

The most common use for a diode is a rectifier.



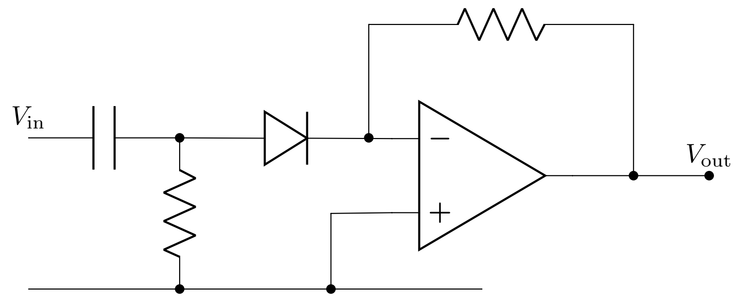
This is not a very good rectifier. It is better to build a full bridge rectifier, which will produce a voltage on both cycles of AC from the transformer.



3.2.2 RMS

We can determine RMS voltage with a diode. The RMS voltage is

$$V_{\mathrm{RMS}} = \sqrt{\langle V^2 \rangle}$$



The op amp functions as a current to voltage converter. The average current through the diode is

$$I_{\text{av}} = \frac{1}{2\pi}\int_0^{2\pi} I_0\left(e^{V\sin t/V_t} - 1\right)\,dt = I_0\left(\mathcal{I}_0\left(\frac{V}{V_t}\right) - 1\right) \approx I_0\frac{V^2}{4V_t^2}$$

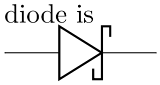
where $\mathcal{I}_0$ is the modified Bessel function. This holds only for a small voltage where the Taylor expansion is valid. The op amp stage serves as a current to voltage converter.

3.3 Diode Types

Germanium diodes have higher reverse current, which makes them suboptimal.

3.3.1 Schottky Diode

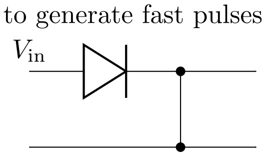
A Schottky diode is good for when speed is needed. The symbol for a Schottky diode is



The threshold voltage of a Schottky diode can be as low as 0.2 V. They use metal bound to a semiconductor, rather than the pn junction. Thus, there is no carrier storage, so they can function at several gigahertz. In a normal diode, some carriers are kept in the pn junction. Thus, for the first bit of a time that a current is reversed, the diode will keep conducting. Diodes have a switch off time.

3.3.2 Step-Recovery Diode

A step-recovery diode has a deliberately longer switch off time; it turns off faster after a delay. It can store for 100 ns, then switch off in 10 ps. This can be used to generate fast pulses with this circuit.



V_{in} is fed with a sine wave.

3.3.3 PIN Diode

Consider the derivative of the diode equation:

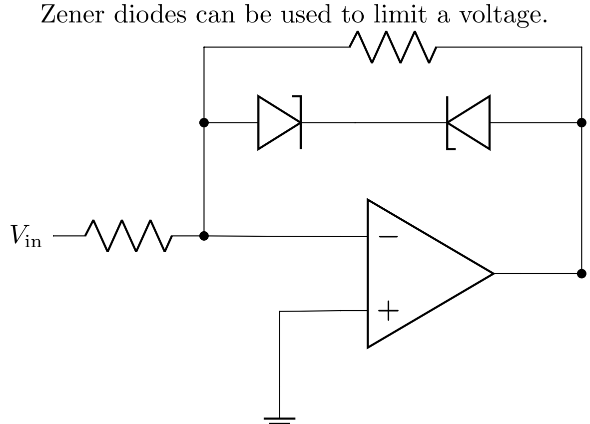
$$\frac{\partial I}{\partial V} = \frac{I_0}{V_t} e^{V/V_t} = \frac{I}{V_t}$$

This suggests that a diode acts as a resistor, with $R_{\text{eff}} = V_t/I$. Note that this only holds for small voltage changes (nearly useless). However, at RF frequencies, the PIN diode does this. In a PIN diode, the pn junction is made artificially thicker by adding an "intrinsic" layer between the p and n. This leads to charge storage and makes the diode very slow. The PIN diode acts like a variable resistor at RF frequencies because there is no rectification of RF signals, but the current will vary with a DC voltage. This will be temperature dependent, but only slightly.

3.3.4 Zener Diode

The Zener diode is also known as the Z-diode (because Zener did not want a diode named after him). Zener diodes are used in reverse. In the forward direction, they conduct at about 0.6 V. Zener diodes are built so that when used in reverse above the breakdown voltage, the diode is not destroyed. When the Zener voltage $V_z < 6\,\text{V}$, they have a negative temperature coefficient, the Zener voltage decreases as temperature increases. When the Zener voltage $V_z > 6\,\text{V}$, they have a positive temperature coefficient. These are not actually using the Zener effect. Zener diodes at 5.6 V and 6.3 V are relatively temperature independent.

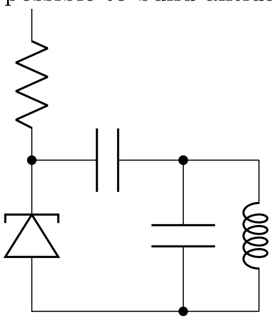
Zener diodes can be used to limit a voltage.



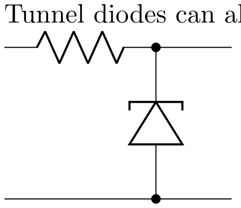
This circuit caps an input signal at the Zener voltage plus 0.6 V. A key specification is the maximum current.

3.3.5 Tunnel Diode

The tunnel diode (also known as the Esaki diode) starts conducting at 0 V, increases to some maximum, begins decreasing again as the voltage increases, then finally starts increasing again. They are specified by the maximum current. They have a negative effective resistance in some region, which makes it possible to build antidamping circuits.



Inductor and capacitor in parallel. Tunnel diodes can also be used as an amplifier.
⟨capacitor in series with TD across output of transformer in parallel with load resistor⟩
Tunnel diodes can also be used as fast triggers.



⟨Should add resistor and tunnel diode load lines⟩ As the voltage at the input increases, the diode goes over its current peak, then rapidly begins conducting at a higher voltage. The switching time is several picoseconds, though there will be substantial histerysis. The switching voltage is around 0.2V for Ge and 0.5V for Si. It is typically hard to procure tunnel diodes, but they are faster than Schmidt triggers.

The backward diode is a tunnel diode with a shallow current peak. Tunnel diodes conduct backward, and the backward diode does as well. This allows the backward diode to be used as a rectifier without the threshold.

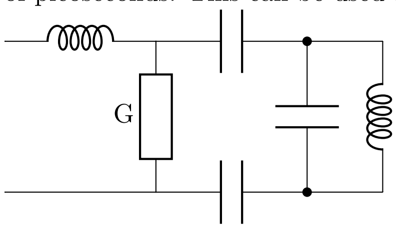
3.3.6 Gunn Element

The Gunn element is not really a diode. It does not rectify. It is made of gallium-arsenide, and has a strange band structure. The effective mass of a carrier in a semiconductor might be

$$m \propto \left(\frac{\partial E}{\partial k}\right)^{-1}$$

By applying a voltage to the Gunn element, the effective mass changes. Most of the voltage will tend to be across a short region of the element. The carriers in that region are heavy and slow. The badly conducting region drifts across the element. When it reaches the end, there will be a pulse from the element.

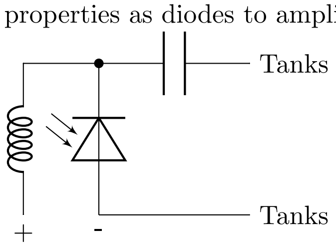
The cycle then restarts. This leads to a chain of pulses, with a period of tens of picoseconds. This can be used as a microwave generator with a tank circuit:



where the element labelled G is the Gunn element. This is used in radar guns.

3.3.7 Photodiode

A photodiode is a diode with a deliberately large cathode. This will be light sensitive, with incoming photons producing carriers. We can also use their properties as diodes to amplify signals.



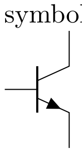
Three tank circuits in series are then connected in parallel with the diode. One tank circuit is at the signal frequency f_s, one is at $2f_s$, and is driven through a transformer at the inductor, and one is at $3f_s$. The diode then acts as a parametric amplifier at f_s, producing a signal large enough to be detected.

4 Transistors

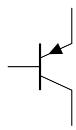
Diodes are not easy to use as amplifiers. The important difference between diodes and transistors is that a transistor has three terminals.

4.1 BJT

There are NPN and PNP Bipolar Junction Transistors. An NPN transistor has symbol



An PNP transistor has symbol



The arrow is at the emitter, the base is on the flat side, and the remaining terminal is the collector.

An NPN transistor turns on if the voltage from base to emitter, $V_{\mathrm{BE}} > 0.6\,\mathrm{V}$. A PNP transistor turns on if the voltage from base to emitter, $V_{\mathrm{BE}} < -0.6\,\mathrm{V}$.

4.1.1 BJT Parameters

An important parameter is β_{DC}, the DC current gain, defined by

$$\beta_{\mathrm{DC}} = \frac{I_{\mathrm{CE}}}{I_{\mathrm{BE}}}$$

where I_{CE} is the collector-emitter current and I_{BE} is the base-emitter current. It can be from 5 (for high power or microwave transistors) to 5000 (for "superbeta" transistors). It is usually around 100. Superbeta transistors are usually used at the inputs of op amps.

The transistor is often said to be a current control device. For an NPN transistor, I_{CE} rises with V_{CE} at low V_{CE}, but will saturate at some value set by the base current I_{BE}. The region where I_{CE} depends on V_{CE} is called the saturation region. The datasheet will contain a saturation voltage. A small saturation voltage is better to use the transistor as a switch. It turns out that even at high V_{CE}, I_{CE} has a dependence on V_{CE}. If you extrapolate the current-voltage lines backwards, they will intersect at some negative voltage, known as the Early voltage. This will be several hundred volts.

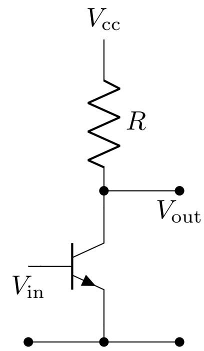
β is not constant. It will decrease at very low I_{CE} and high I_{CE}. Only the general shape is common to all transistors, the change is more or less pronounced in different transistors. This is quite important for using transistors as photodiode pre-amplifiers, since I_{CE} will be very small (and β is usually not specified in this region on the datasheet). β is also a function of frequency.

$$\beta = \frac{\beta_{\mathrm{DC}}}{1 + i\frac{f\beta_{\mathrm{DC}}}{f_T}}$$

f_T is the transistion frequency, where β goes to one. In addition, at high frequencies, there will be a phase shift of 90°. f_T can be as low as 10 kHz for high power transistors and as large as 100 GHz for RF front ends (small transistors). It is typically about 100 MHz. There is an inverse relationship between the maximum voltage of a transistor and f_T.

4.1.2 Common Emitter Amplifier

A typical BJT amplifier, the common emitter amplifier, looks like



We will assume we have small signals at about 0.6 V.

The gain is

$$g = -R\frac{\partial I_{\text{CE}}}{\partial V_{\text{BE}}}$$

$$I_{\text{CE}} = \beta I_{\text{BE}} = \beta I_0 e^{V/V_t}$$

where the second equality uses the diode equation based on the diode in the transistor (so V_t is the diode thermal voltage). The transconductivity is

$$g_m = \frac{\partial I_{\text{CE}}}{\partial V_{\text{BE}}} = \beta\frac{I_{\text{CE}}}{V_t}e^{V/V_t} = \frac{I_{\text{CE}}}{V_t}$$

which is "a great equation", as it is the same expression for any bipolar junction transistor, independent of the transistor construction. For $I_{\text{CE}} = 1\,\text{mA}$.

$$g_m = \frac{1\,\text{mA}}{25\,\text{mV}} = 40\,\text{mA/V}$$

The thermal voltage will be relatively constant. Using our expression for the transconductivity, the gain is

$$g = -R\frac{I_{\text{CE}}}{V_t}$$

The collector-emitter current can be determined from the resistor.

$$I_{\text{CE}} = \frac{V_{\text{cc}} - V_{\text{out}}}{R} \approx \frac{1}{2}\frac{V_{\text{cc}}}{R}$$

Thus, the gain is

$$g = -R\frac{I_{\text{CE}}}{V_t} = -\frac{1}{2}\frac{V_{\text{cc}}}{V_t}$$

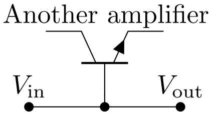
As V_{cc} increases further, we will run into issues with the Early voltage. Eventually, the gain will approach $g = -\frac{V_{\text{Early}}}{2V_t}$.

The input impedance is also important. It is

$$R_{\text{in}} = \left(\frac{\partial I_{\text{BE}}}{\partial V_{\text{in}}}\right)^{-1} = \frac{\beta}{g_m}$$

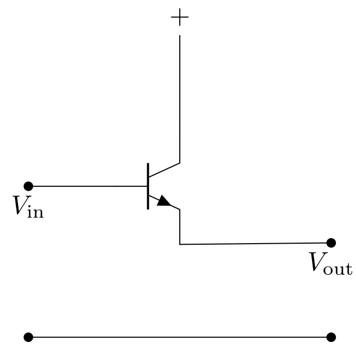
4.1.3 Other Amplifiers

Another amplifier is the common base amplifier



⟨Add explanation of what this does⟩.

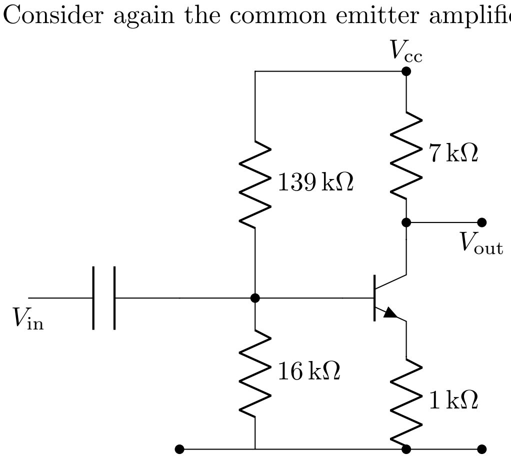
Another amplifier is the common collector amplifier, which is also known as the emitter follower.



This has no voltage gain, but provides a current gain, with a high input impedance of $R_{\text{in}} = \beta R_{\text{out}}$. In this circuit, $V_{\text{in}} = V_{\text{out}} + 0.6\,\text{V}$. The current out is $I_{\text{out}} = V_{\text{out}}/R_{\text{load}}$ (for some load attached at V_{out}). The current in is only I_{out}/β, which will be fairly small since β is large.

4.1.4 Improved Common Emitter Amplifier

Consider again the common emitter amplifier



⟨Show how resistor values are calculated⟩. The goal here is to shift the signal voltage so the output fluctuates about zero, without needing to elevate the voltage from ground. The solution is to use a voltage divider from $+V_{cc}$ to ground, and AC couple the input to avoid adding an additional offset. The resistor values for the voltage divider are chosen so that the signal is centered around $1.6\,V$. We also need to make the current through the voltage divider much larger than the base current I_B to maintain integrity of the voltage divider. The base current will be around $5\,\mu A$, so the resistors are chosen accordingly.

We would choose a transistor with $\beta \approx 200$. The gain is approximately the ratio of the collector and emitter resistances:

$$g = -\frac{7\,\mathrm{k}\Omega}{1\,\mathrm{k}\Omega}$$

This is due to "emitter degeneration". More exactly,

$$g = -\frac{R_2}{R_e + 1/g_m} \approx -\frac{R_2}{R_e}$$

Without degeneration,

$$g = -g_m R_2 = -40\mathrm{mA/V} times 7\,\mathrm{k}\Omega = -280$$

However, g_m depends on the current, so this is nonlinear. Thus, with emitter degeneration we have a lower gain but linear amplifier. To get high gain, we add a capacitor in parallel with R_e, which effectively reduces R_e at high frequencies.

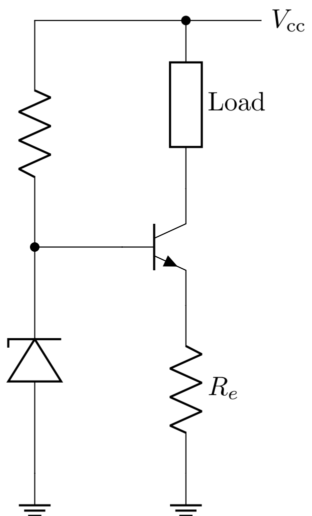
The input impedance with degeneration is

$$R_{\text{in}} = \beta\left(\frac{1}{g_m} + R_e\right) = 200\,(25 + 1000)\,\Omega = 200\,\mathrm{k}\Omega$$

⟨Explain⟩

4.1.5 Transistor Current Source

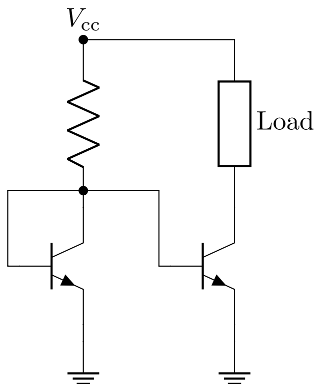
We can make a transistor current source



The voltage across R_e will be $V_z - 0.6\,\mathrm{V}$, so the current through the load, which is the same as the current through R_e will be constant. To make this temperature independent, a Zener diode with $V_z \approx 10\,\mathrm{V}$ should be selected. This will lead to the temperature effects of the transistor and temperature effects of the Zener diode cancelling out.

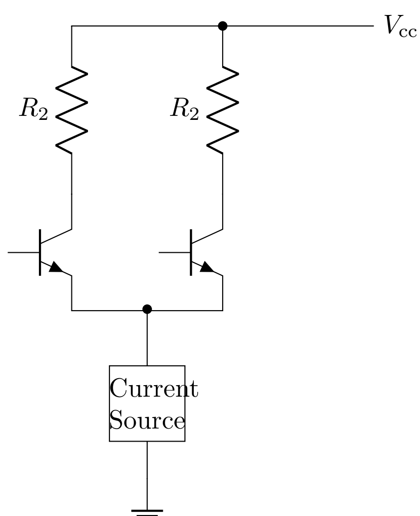
4.1.6 Current Mirror

To make this circuit, we need the I_{CE} and V_{BE} of two transistors to be the same. This requires a matched pair.



The current through the load will be the same as the current through the resistor. ⟨Add explanation⟩

4.1.7 Differential Pair



⟨add explanation (this diagram was hard)⟩ The input voltage is between the two bases, the output voltage is between the two collectors.

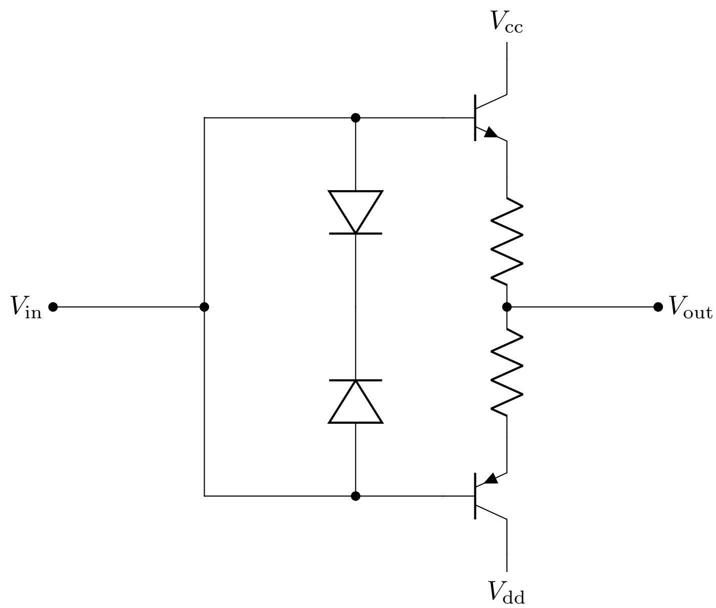
4.1.8 Op Amp

We rarely want a floating output, so we change the circuit a bit. This is similar to what is in an op amp.
⟨Diagram⟩
The current mirror resistor is chosen such that we have $\frac{30\,\mathrm{V}}{100\,\mu\mathrm{A}} = 300\,\mathrm{k}\Omega$. We replace our upper resistors in the differential pair with more current mirrors, set such that we have $50\,\mu$A. This makes that resistor $600\,\mathrm{k}\Omega$. Note that many pairs of transistors in this circuit must be matched.

We connect the output to a PNP transistor based differential amplifier with one of the outputs short circuited. The other output is connected to a NPN-PNP complementary pair before the output.

4.1.9 Complementary Pair

The NPN-PNP complementary pair is

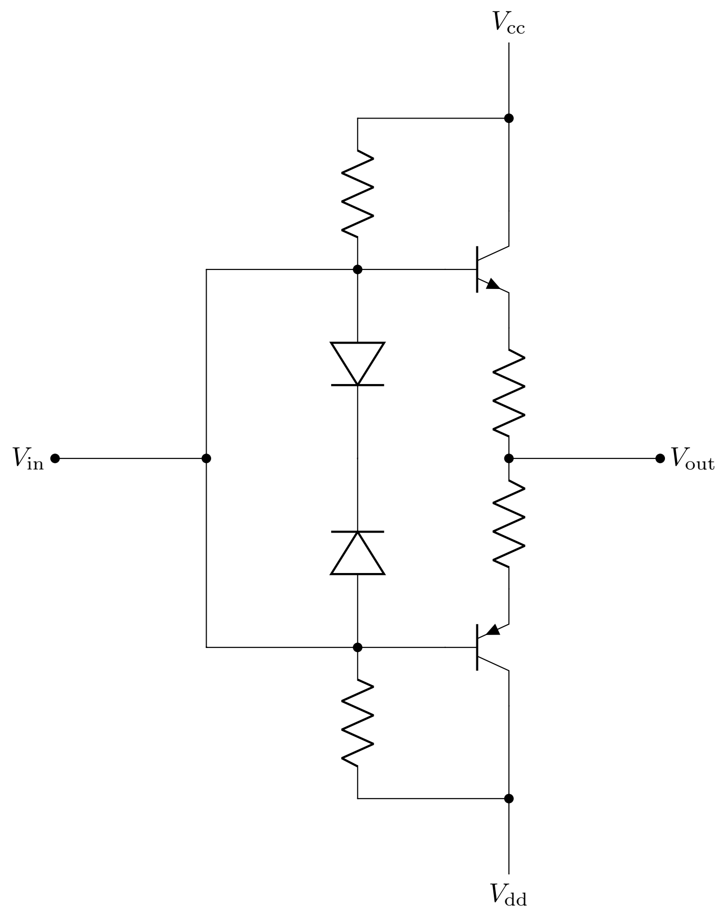


This can be used to drive a signal to positive and negative voltages with more current. The diodes are used to pull the voltage such that there is no region where neither transistor is driven.

4.1.10 Power Amps

The purpose of a power amp is to increase current at the output. A power amp should not change the voltage, but should produce the same voltage across a larger load. The emitter follower can amplify current, but only positive currents and voltages.

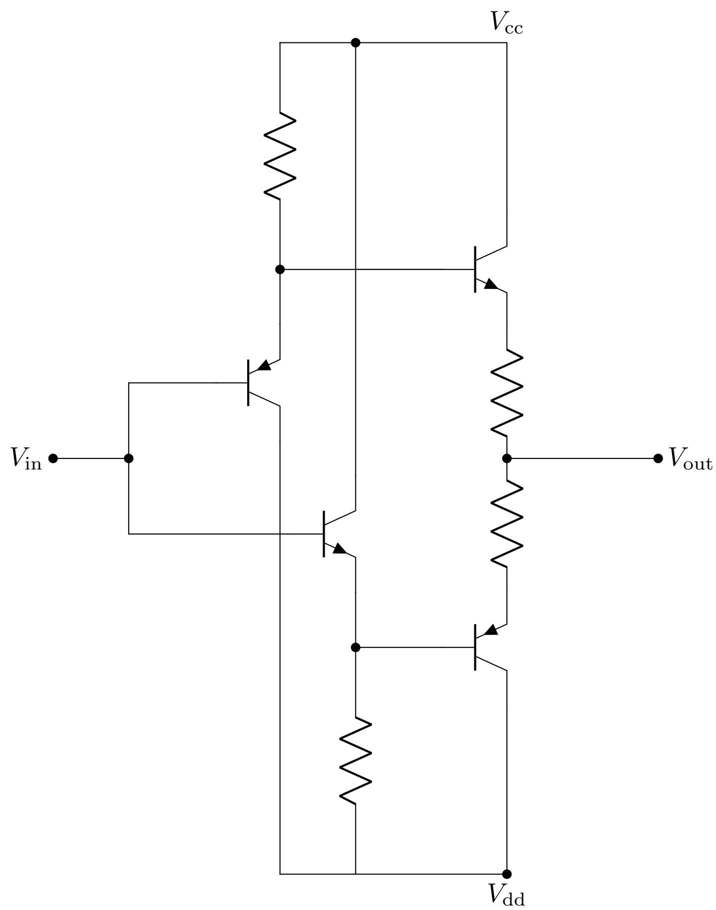
The NPN-PNP complementary pair can drive both polarities.



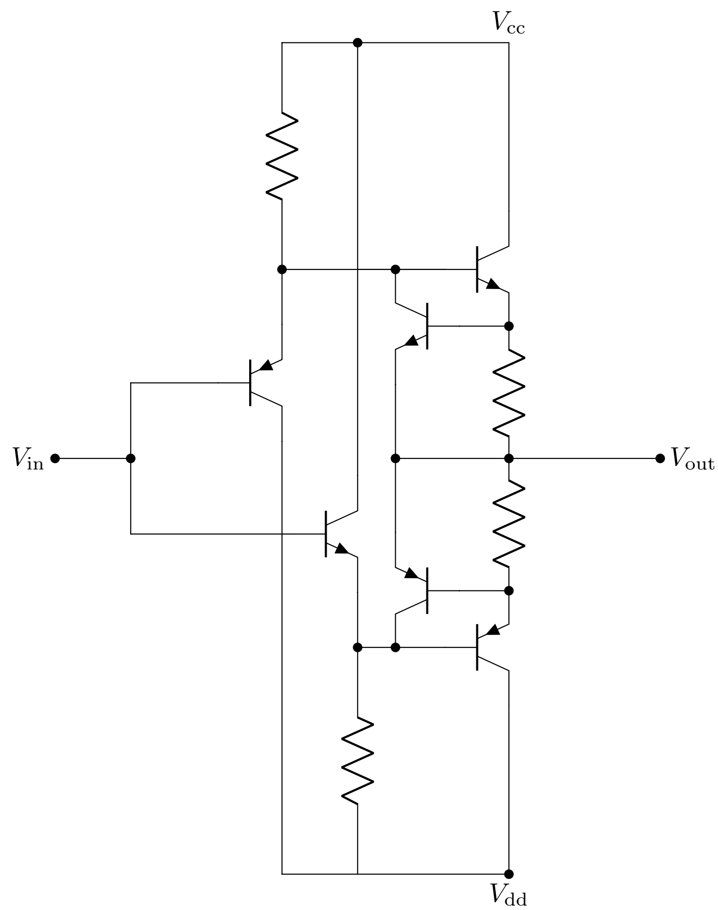
This can be used to drive a signal to positive and negative voltages with more current. The diodes are used to pull the voltage such that there is no region where neither transistor is driven. Without these, V_{out} will be zero for $|V_{\mathrm{in}}| < 0.6\,\mathrm{V}$. With the diodes, there is a $0.6\,\mathrm{V}$ drop across the diodes, which is enough to turn on the transistors. However, even with the diodes, there will still be a slight zero region. This is known as crossover distortion.

There are three classes of amplifiers. Class A has a heavy bias current, $I_{\mathrm{bias}} \geq I_{\mathrm{signal}}$. These have very low crossover distortion, but also low efficiency. The Mini-Circuits RF amps are Class A. Class B have zero bias current, which leads to high distortion and high efficiency. A Class AB has low bias, lowish distortion, and highish efficiency.

We can replace the diodes with transistors to improve the circuit.



This adds additional gain and fixes the biasing issue. We also often want to add current limiting, especially when driving large loads.



The extra transistors will turn off the two transistors driving V_{out} when the current across the resistors is too large, thus serving to limit the current through the amplifier.

4.1.11 Maximum Ratings

p_D is the maximum power dissipation, ranging from a few milliwatts to hundreds of watts. This is due to the transistor having some internal resistance, which leads to heating. There is a thermal resistance from the silicon junction to the case and a maximum temperature that the junction can reach. Thus, p_D is limited by the ambient temperature and heat dissipation. A heat sink can help improve heat dissipation. It is not good to bolt transistors to heat sinks, this can bend the case and reduce thermal contact. Thermal grease will also help. It is not a good heat conductor, merely better than air. Thus, thermal grease layers should be as thin as possible. This can be done by scraping most of it off with a razor blade.

A transistor has a safe operating area, set by a maximum collector current, maximum collector-emitter voltage, and maximum power dissipation. This would seem to set a trapezoidal region in which the transistor can be operated. However, there is "secondary breakdown", making the maximum voltage decay faster than $1/V_{\mathrm{CE}}$ at the power limit. This is due to the transistor being internally non-uniform. Current flows more easily in warm regions of the transistor, so the transistor will tend to develop hot spots. Secondary breakdown tends to occur above $10\,\mathrm{V}$ to $20\,\mathrm{V}$. Some transistors do not have secondary breakdown.

Power transformers use several transistors in parallel. This can not be done directly, or the warmer transistor will tend to take more current, heat up more, and run away. Thus, we need emitter degeneration: resistors from the transistor emitter to ground.

For short times, the maximum current can be exceeded. This effectively stores heat in the transistor.

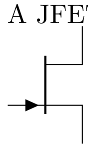
There is a leakage current from the collector to the base. As V_{CE} increases, this increases, which increases the base voltage and turns the transistor on. This will rapidly lead to increasing current through the transistor, until it breaks down. The breakdown voltage is denoted $V_{\mathrm{CE},\,0}$. If the base is connected to the emitter by a resistor, the leakage current can flow to ground. This will increase the maximum collector emitter voltage higher. This is denoted $V_{\mathrm{CE},\,\mathrm{R}}$ for a resistor connected or $V_{\mathrm{CE},\,\mathrm{S}}$ for a short from base to ground. It is best to choose a transistor such that the smallest, $V_{\mathrm{CE},\,0}$, is larger than the voltage you need. However, if you need something very fast, a SiGe transistor tends to have $V_{\mathrm{CE},\,\mathrm{S}}$ much larger than $V_{\mathrm{CE},\,0}$. Thus, it might be possible to use such a transistor, carefully.

Transistors do no live forever. They are limited by electro-migration. Electro-migration is the movement of atoms in the transistor due to electric fields. In particular, the atoms in the metals connected to the semiconductor are pulled into the semiconductor, which disturbs the doping. Electro-migration is a function of current and temperature. For maximum current and temperature, the lifetime might be limited to $10,000$ hours. In general, try to run transistors are lower currents and temperatures. Very high frequency circuits tend to have more issues with this. Faster transistors have to be thinner (to decrease drift time) and smaller (to decrease capacitance), which tends to cause these issues.

4.2 More Transistor Types

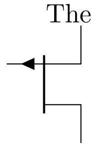
4.2.1 JFET

A JFET is



The left terminal is the gate (like the base), the lower terminal is the source, and the upper terminal is the drain. This is an N channel JFET, which is like an NPN BJT.

The P channel JFET is



This is like a PNP BJT.

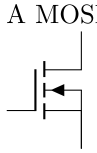
An NPN conducts at $V_B \geq 0.6\,\mathrm{V}$. The N channel JFET conducts until V_{GS} is negative. For an NPN BJT, the base current is $I_B = I_C/\beta$. For a JFET, the base gate current is essentially zero.

⟨JFET diagram⟩

A JFET stops conducting when V_{GS} is below some negative value, typically less than $-1\,\mathrm{V}$. V_{GS} can be a bit positive, but must be small enough that the diode from gate to source does not start conducting. JFETs are fairly sensative to static electricity on their gate pin.

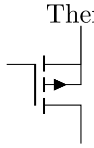
4.2.2 MOSFET

A MOSFET symbol is



The left terminal is the gate, the lower terminal is the source, and the upper terminal is the drain. This is an N channel enhancement mode MOSFET (NPN type). The schematic does relate to some of the internal connections.

There is also a P-channel MOSFET:



Enhancement mode means there is no current at $V_{GS} = 0$. MOSFET means metal-oxide semiconductor field effect transistor. Metal-oxide refers to the insulation from gate to source. The metal-oxide is an insulator, so $I_G = 0$, always.

The advantage of JFETs is lower noise at lower frequencies. Otherwise, MOSFETs should be used.

All the BJT circuits can be made with MOSFETs.

4.2.3 IGBT

Insulated gate bipolar transistor.

The IGBT is like a BJT, except the base current is zero.

4.2.4 Comparison of Transistors

	BJT	JFET	MOSFET	IGBT
β	10-1000	∞	∞	∞
V_{CE}	1 kV	10 V	2 kV	2 kV
I_{CE}	100 A	10 mA	100 A	1000 A
p_D	100 W	< 1 W	100 W	kW
f_T	100 GHZ	Limited by capacitances		
Switching speed	saturation	Limited by capacitances		saturation
g_m	I_C/V_t	less than I_S/V_t		
Noise	Good for low impedance	Good at high impedance	$1/f$ noise	terrible
Capacitances	Lowest	Higher	Higher	???

In general, larger transistors are slower.

Switching speed for BJTs is limited by how long the junction is saturated. Outside of the linear region, the speed of a BJT decreases.

For a BJT, $g_m = I_C/V_t$ is the same for all transistors. This is the highest possible (by thermodynamics). Related to this, the BJT has the lowest voltage noise. Voltage noise is noise when the base is short circuited. Current noise is noise produced when there is a resistor from the base to the ground. BJTs have higher current noise.

BJTs are the only transistors for which we have a good equation for the base current.

$$I_B = I_0 e^{V_{\text{CE}}/V_t}$$

JFETs are only useful for lower power applications, they do not conduct more than 10s of milliamps.

The speed of field effect transistors are limited by capacitances between terminals, which can be quite large. Large MOSFETs have especially large capacitances.

JFETs and MOSFETs do not follow an accurate current law, but have

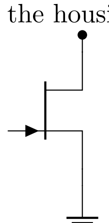
$$I_{DS} \approx V_{GS}^{3/2}$$

IGBTs are very high power, and should only be used when a high power MOSFET or BJT is not available.

4.3 JFET Circuits

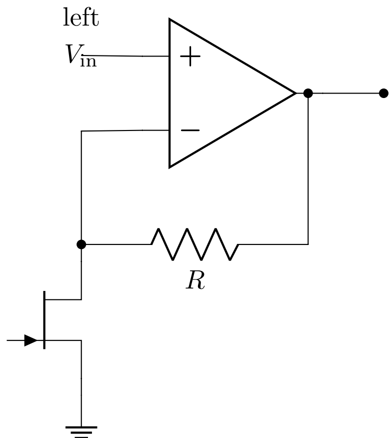
4.3.1 JFET as a Variable Resistor

Note that symbols with and without circles are equivalent, it is meant to evoke the housing of discrete transistors.

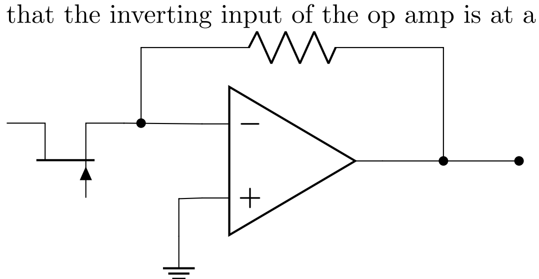


For a low V_{DS}, we have $I_D \propto V_{DS}$, so the JFET looks like a resistor. The slope of the proportionality varies with V_{GS} (decreasing as V_{GS} becomes more negative). However, for larger V_{DS}, the JFET saturates.

We can use this to make a variable gain amplifier

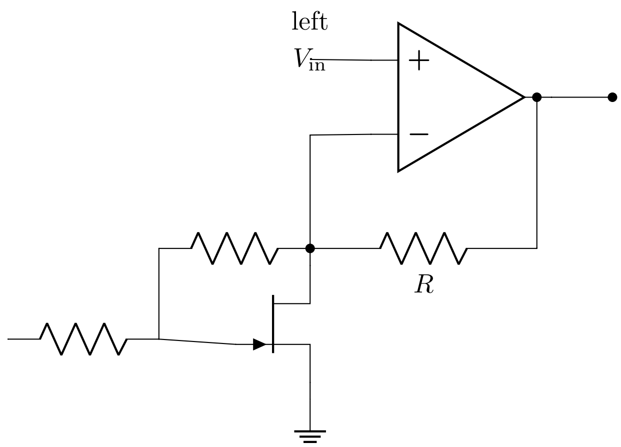


We can also make this work with a non-inverting amplifier. This uses the fact that the inverting input of the op amp is at a virtual ground.



To make this more linear, we can build an op amp amplifier with a LDR (light dependent resistor) coupled to an LED. The problem with LDRs is the resistance only increases slowly. It can decrease rapidly.

We can also improve the transistor version by adding feedback.



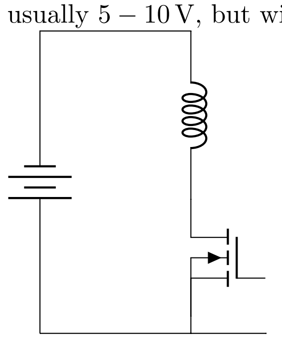
This helps linearize the transistor through negative feedback on V_{GS} of the transistor.

4.4 Transistors as Switches

We might want to switch high voltage and high current.

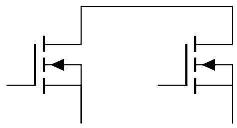
4.4.1 High Current

For example, switching a MOT (magneto-optical trap) magnet. The voltage is usually $5-10\,\mathrm{V}$, but with $100\,\mathrm{A}$.



We want to make sure that we stay in the safe operating area of the transistor. The MOSFET will have a resistance from drain to source specified, R_{DS}. We want this to be as small as possible. In general, R_{DS} decreases as V_{GS} increases. The MOSFET will start turning on around $V_{GS} = 2\,\mathrm{V}$, but will not fully turn on until around $V_{GS} = 10\,\mathrm{V}$. Operating between those voltages will lead to a larger R_{DS}. The voltage drop across the transistor is $R_{DS}I$.

We might want to connect MOSFETs in parallel. BJTs require emitter degeneration to be connected in parallel, since they tend to conduct better as they get warmer.



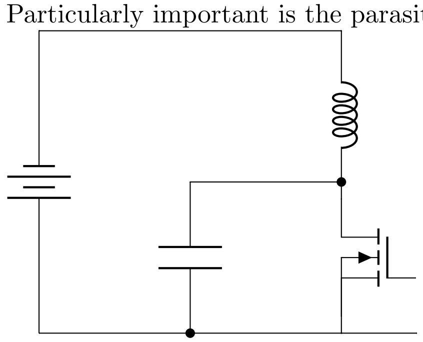
It is claimed that MOSFETs do not have this problem, and can be simply connected in parallel. However, the I_D versus V_{GS} curve does change depending on the temperature. The slope of the curve decreases at higher V_{GS}, which leads to hotter transistors conducting worse for a sufficiently high gate voltage. This means that we want the transistors to be not too oversized.

It is important to put small resistors before the gates of the MOSFETs to eliminate a resonant circuit between wire inductance and gate capacitance.

The switching on speed of the circuit will be given by

$$\dot{I} = V/L$$

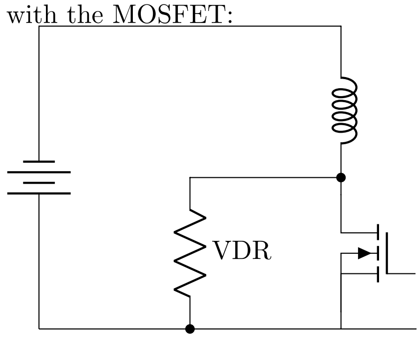
The switching off speed depends on the parasitic capacitances of the MOSFET. Particularly important is the parasitic drain source capacitance.



The capacitance can be as large as 5 nF. This leads to an LC resonant circuit. We thus get a high frequency ($f = \frac{1}{2\pi\sqrt{LC}}$) sinusoidal oscillation. The peak voltage will be given by

$$\frac{1}{2}LI^2 = \frac{1}{2}CV_p^2$$

V_p will be very large, since I is large. The resistance of MOT coils is low, so the Q of the resonator will be quite high. To resolve this, we put a VDR in parallel with the MOSFET:

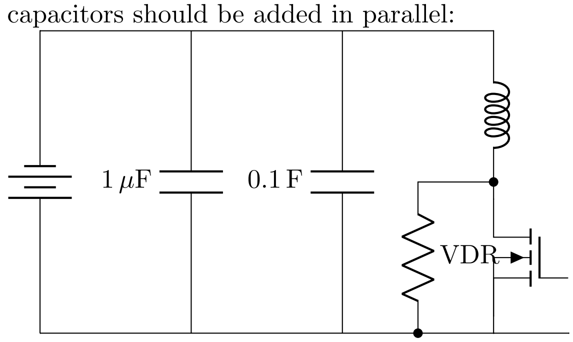


⟨Proper VDR symbol⟩ Recall that a VDR is a resistor that only begins conducting at some voltage. Choosing that voltage to be some value, we can dissipate the current and prevent oscillations. The current is now

$$I_0 \left(1 - t\frac{V_{VDR}}{L}\right) \implies t_0 = \frac{L}{V_{VDR}}$$

where t_0 is the time to finish switching off. Note that this implies we get a faster switch off with a higher voltage limit, but this requires the rest of the system to be able to withstand higher voltages.

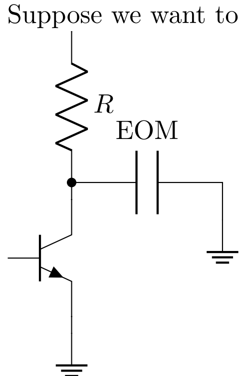
The power supply will not be pleased about the large transients, so large capacitors should be added in parallel:



When turning off the MOSFET, the gate drain capacitor will be charged to a large value. Discharging this capacitor rapidly thus requires fairly large currents. A MOSFET driver should thus be low impedance.

4.4.2 High Voltage

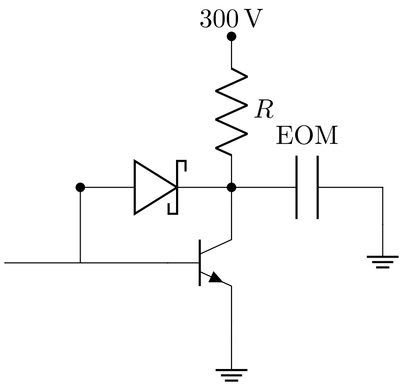
Suppose we want to switch an EO modulator.



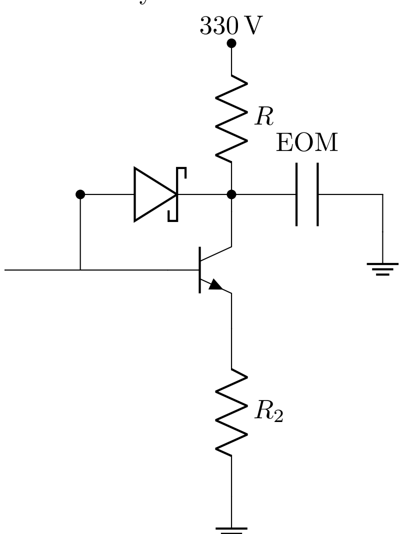
The capacitance of the EOM is about 20 pF. Note that RG-58 cable has 101 pF per meter, which could add a lot of capacitance. RG-62 cable is better at 45 pF/m, but avoiding cables is best.

Suppose we have 10 pF stray capacitance between cables and the transistor. Suppose $R = 10\,\mathrm{k\Omega}$. Then we can switch the EOM with time constant $RC = 10^4\,\Omega \times 10^{-11}\,\mathrm{F} = 1\,\mu\mathrm{s}$. So, we can charge the capacitor to 90% in $2.2\,\mu$s, if the transistor is not saturated.

We avoid saturating the transistor by adding a Schottky diode. We get saturation when V_{CE} is too low. With the Schottky diode, the collector is held above 0.3 V.



Another way to do this is



We choose R_2 such that we get a 10 V drop across it at the desired current, then drive the transistor with a 10.6 V pulse. This leaves us with 30 V at the collector, which will keep it quite far above saturation.

We are trying to drive a square wave. To have an edge of 100 ns, the Fourier component we care about is at

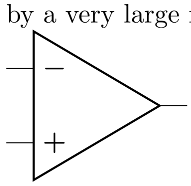
$$\frac{0.34}{100\,\mathrm{ns}} = 3.5\,\mathrm{MHz}$$

$$\beta\,(3.5\,\mathrm{MHz}) = \frac{f_T}{3.5\,\mathrm{MHz}} \implies I_B = \frac{I_C}{\beta\,(3.5\,\mathrm{MHz})} = \frac{30\,\mathrm{pF} \times 300\,\mathrm{V}}{100\,\mathrm{ns}}\frac{1}{30} = \frac{90\,\mathrm{mA}}{30} = 3\,\mathrm{mA}$$

assuming we have a transistor with $\beta = 100$ and $f_T = 100\,\mathrm{MHz}$. Note that the decreasing β implies the base voltage must be increased. ⟨Explain better⟩.

5 Op Amps

An op amp essentially takes the difference between two inputs and amplifies it by a very large factor. The symbol is



The inverted input is marked −. The noninverted input is marked +. The output is on the right. The power supply inputs are not shown.

Op amps are designed to be used with feedback to constrain the gain. With networks of resistors and capacitors, we can design complex linear circuits. With other elements, we can built logarithmic and exponential multipliers, multipliers, and other non-linear circuits.

Negative feedback in circuits was patented by Black at Bell Labs. However, human in the loop negative feedback has occured forever.

Issues with negative feedback occur when there is delay. This leads to oscillations in the feedback loop. The output of an amplifier will lag the input, leading to a phase shift. Once the phase shift is 180°, we have positive feedback instead of negative feedback.

In summary, op amps take the difference of two inputs, set the output to that times a large value, and work well with negative feedback. In some transistor amplifiers, there were level shifts. We want an op amp not to have a level shift relative to the input.

5.1 Specifications

Our specifications emerge from the requirements above.

5.1.1 Maximum Ratings

Maximum ratings should never be exceeded, but are insufficient to guarantee good operation.

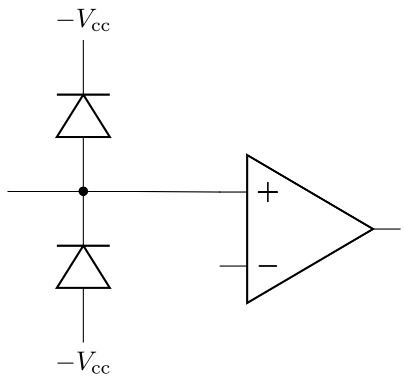
The difference between the positive supply and negative supply is limited. For an op amp designed to be used with ±15 V, this is typically 36 V. Some op amps have this limited to 12 V (designed to use ±5 V). If possible, choose one with the larger supply voltage range.

There is a maximum differential input voltage (the difference between the + input and the − input). This is typically ±10 V. Exceeding this may not entirely break the op amp, but will damage the input stage leading to additional noise or input drift. Such issues will be quite difficult to diagnose.

There can be a limit on the short circuit duration. Most op amps allow for the output to be shorted to ground indefinitely.

There is a maximum junction temperature, usually 150°C. Note that the internal temperature is higher than the surrounding temperature.

There are specifications on electrostatic discharge (ESD). This will tell you if you need addition protections against ESD. It is usually good to have protection diodes to the supply before the input of the circuit.



However, diodes will conduct a little bit of current, reducing the input impedance. To reduce this, the diode can be replaced with a JFET. Replace the anode of the diode with the gate and the cathode of the diode with the source.

5.1.2 DC Specifications

Ideally, the op amp outputs the positive input minus the negative times a large factor. However, it also adds an input offset voltage, denoted V_{OS}. A very good V_{OS} is 10 nV. This requires a chopper stabilized op amp, which internally converts the input to AC. They tend to add noise. A good value of V_{OS} is a few microvolts. Fast op amps tends not to have good V_{OS}. BJT op amps tend to be better, FET op amps tend to be worse.

The offset voltage drift is the change in V_{OS} with temperature. A good op amp (chopper stabilized or BJT) tends to be $0.1\,\mu\mathrm{V}/^\circ\mathrm{C}$. FET op amps tend to be higher, $1\,\mathrm{mV}/^\circ\mathrm{C}$. Solder joints tend to act as thermocouples, and can produce an offset voltage drift of a few $\mu\mathrm{V}/^\circ\mathrm{C}$.

The input bias current, I_B, is the flows into the inputs. For FET op amps, it can be as low as 10^{-15} A. For BJT op amps, it is around 10^{-4} A. Faster op amps tend to be BJT, and tend to have larger input bias.

The offset current is the difference between the input bias current at each input. This tends to be $I_B/10$.

The input resistance is $R_{\mathrm{in}} = \frac{\partial I_B}{\partial V_{\mathrm{in}}}$. This is 10s of megaohms for BJT op amps and teraohms for FET op amps.

The output resistance is R_O. For a good op amp, it is as low as $10\,\Omega$. It can be as high as $10\,\mathrm{k}\Omega$. This is usually less important, as feedback will compensate.

The common mode rejection ratio is CMRR. A good value is 120 dB, a bad value is 20 dB. It tends to decrease at high frequencies. The actual op amp output is

$$g\left(V_+ - V_-\right) + \frac{g}{\mathrm{CMRR}}\left(V_+ + V_-\right)$$

where g is the gain, V_+ is the positive input, and V_- is the negative input.

Similarly, there is the power supply rejection ratio, PSRR. A good value is 100 dB, a bad value is 20 dB. PSRR is usually good for small changes and bad for large changes.

The open loop voltage gain is denoted A_V. A good value is 120 dB, a bad value is 60 dB. This is at DC, and usually decreases at high frequencies.

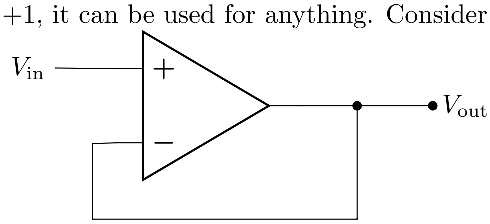
There will be a maximum output current. This can be as high as 100 A (but this will cost \$100). 10 mA is typical for a low power op amp.

5.1.3 AC Specifications

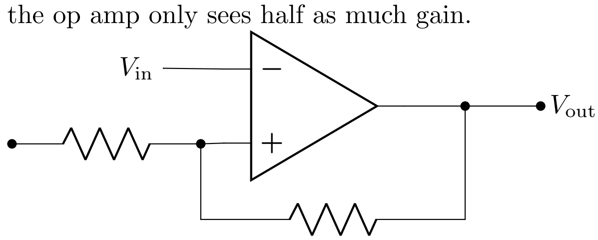
The slew rate, denoted SR, is the speed at which the output voltage can change. Imagine applying a step function to the input. The op amp output will not change infinitely fast. The maximum rate of change at the output $\frac{\partial V_{\text{out}}}{\partial t}$ is the slew rate. A good value is several kilovolts per microsecond, a bad value is a few volts per microsecond. You should not use a faster op amp than necessary though. The slew rate is when the input is overdriven.

For smaller signals, we consider the unity gain bandwidth f_T. A fast op amp would have hundreds of megahertz, a slow op amp would be a few megahertz.

The unity gain bandwidth needs to be combined with the minimum gain for stability. A good specification is +1. When an op amp is stable at a gain of +1, it can be used for anything. Consider a follower.



The gain of the op amp usually decays as $1/f$, which leads to a 90° phase shift. Note that this implies to a minimum phase network (no additional delays). Thus, it will be stable with no gain. When an op amp is made faster, there will be a point where the gain decays as $1/f^2$, in which case the phase shift increases, which can lead to positive feedback and oscillations. An inverting amplifier can be stable when the non-inverting amplifier is unstable. In an inverting amplifier, the op amp only sees half as much gain.



Very fast op amps tend to have a higher minimum gain.

5.1.4 Input Noise

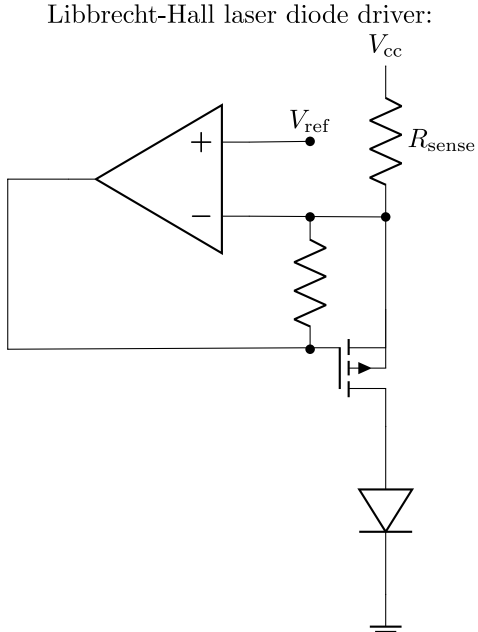
The input noise voltage is denoted e_n. A good specification is $1\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. We use $\sqrt{\mathrm{Hz}}$ since noise is incoherent, so power is added instead of voltage. It can be specified in terms of an equivalent resistor. For a resistor, $e_n = \sqrt{4k_B TR}$, where k_B is the Boltzmann constant and T is the temperature. For a $50\,\Omega$ resistor at room temperature, $e_n = 0.9\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. BJT op amps tend to have a better specification, FET op amps can be as high as $50\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$.

There is also an input noise current, I_n. A good value, for FET op amps, is a few femptoamps per root Hertz (this is usually not specified, because it is so small), a bad value is tens of picoamps per root Hertz. For a signal source with a $50\,\Omega$ resistor, our noise due to voltage noise is

$$g\sqrt{\left(\frac{0.9\,\mathrm{V}}{\sqrt{\mathrm{Hz}}}\right)^2 + e_n^2 + (I_n 50\,\Omega)^2}$$

For the LM7171, $e_n = 14\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ and $I_n = 1.5\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$. Thus, the $I_n 50\,\Omega$ term is only $75\,\mathrm{pV}/\sqrt{\mathrm{Hz}}$. This can be ignored. If we increase the resistor to $0.1\,\mathrm{M\Omega}$, the current noise results in $150\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. The lowest voltage noise op amp is probably the LT1028, which has $0.85\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. However, it has a current noise of $1\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$.

Libbrecht-Hall laser diode driver:

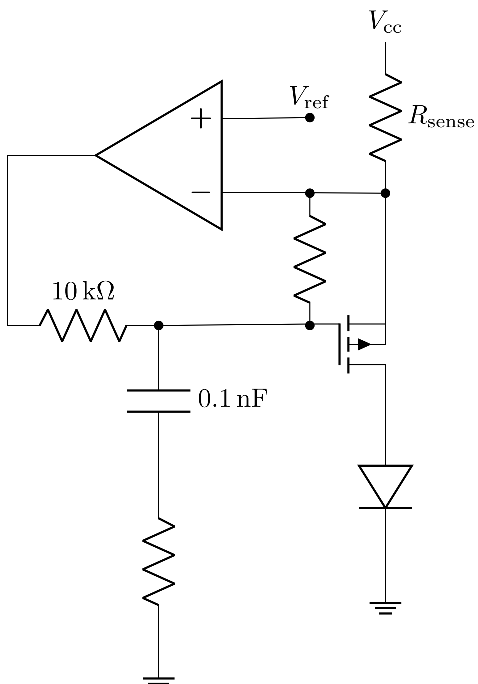


This acts as a current source. The short noise is $\sqrt{2eI}$, where I is the current. For the Johnson noise across the sense resistor to be below shot noise, we need $IR_{\mathrm{sense}} > 50\,\mathrm{mV}$. We choose $R_{\mathrm{sense}} = 50\,\Omega$. We set the reference voltage to $5\,\mathrm{V}$ to get $100\,\mathrm{mA}$.

The current noise is

$$\frac{V}{R_{\text{sense}}} = \frac{1}{R_{\text{sense}}}\sqrt{(4k_BTR_{\text{sense}})^2 + e_n^2 + (I_nR_{\text{sense}})^2}$$

The LT1028 is not unity gain stable, so there also needs to be a network at the output of the op amp. By placing the network at the output, the signal has been amplified, so the noise in the network is irrelevant.



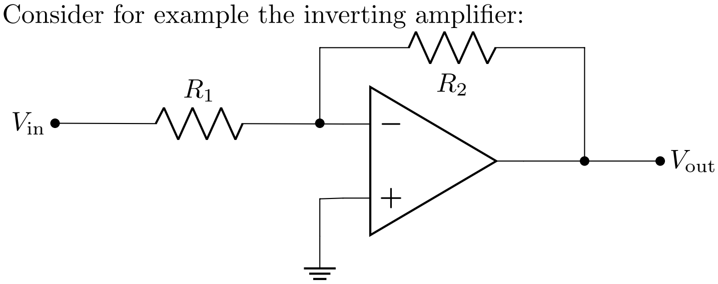
5.2 Op Amp Circuits

Consider an op amp. Most op amp circuits can be analyzed with the following assumptions:

- The input voltage difference is zero.
- There is no current into the op amp inputs.

5.2.1 Inverting Amplifier

Consider for example the inverting amplifier:

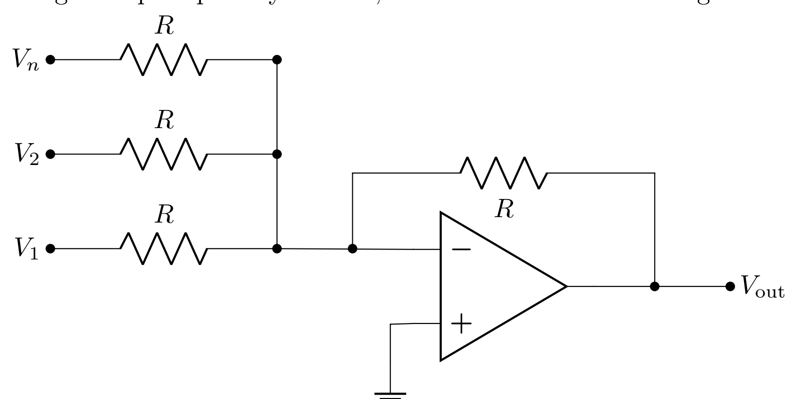


Since $V_- = V_+ = 0$, and no current flows into (or out of) the input, $V_{\text{in}}/R_1 = -V_{\text{out}}/R_2$. Therefore, $V_{\text{out}} = -\frac{R_2}{R_1}V_{\text{in}}$.

6 Analog Signal Processing (With Op Amps)

6.1 Sum

Using our op amp analysis rules, we can see that this is a signal adder:



The output is

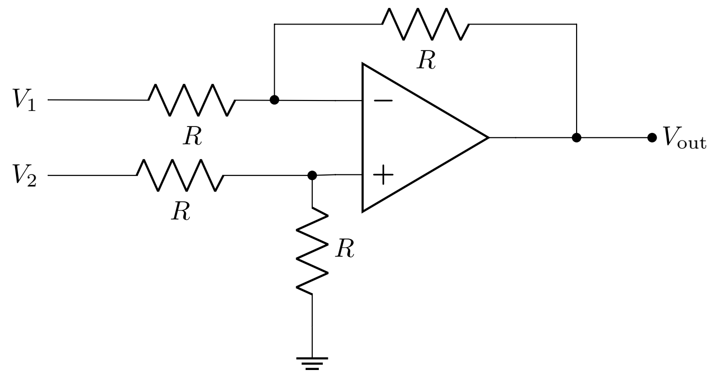
$$-\frac{V_1 + V_2 + \ldots + V_n}{R} R_2$$

Note that even if each individual signal does not exceed op amp limits, the combination might.

We can weight the inputs by changing the resistors in front of them.

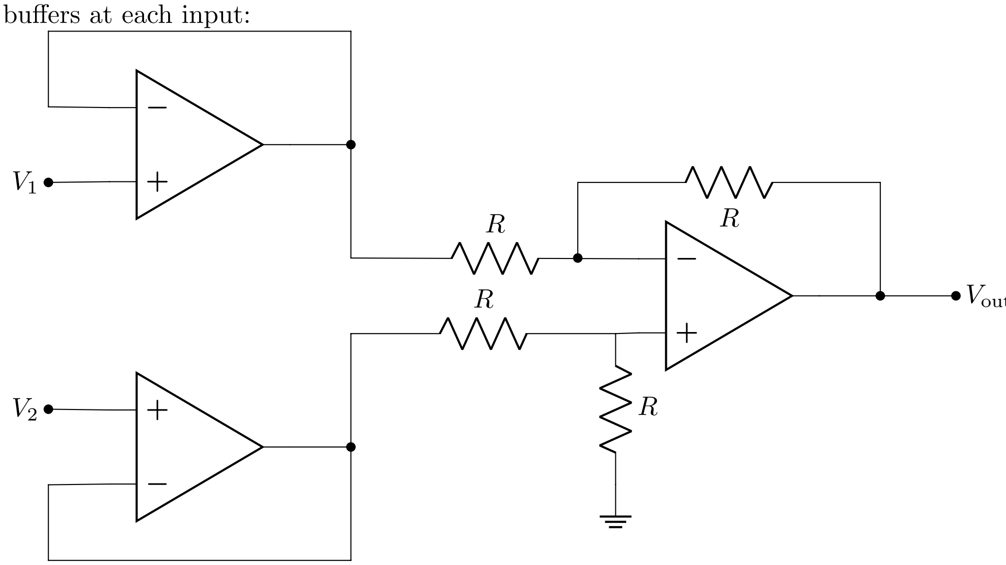
6.2 Instrumentation Amplifiers

To do differences, we need a more complicated circuit, which is found in instrumentation amplifiers.

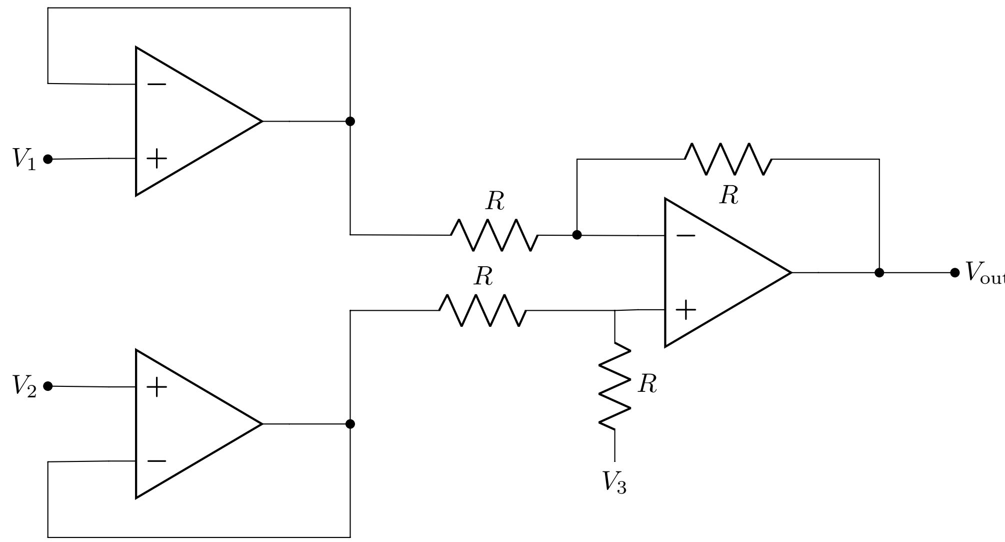


When $V_2 = 0$, we get $V_{\text{out}} = -V_1$. When $V_1 = 0$, $V_{\text{out}} = V_2$. Thus, in general, $V_{\text{out}} = V_2 - V_1$.

One issue with this circuit is that V_1 and V_2 are loaded by the resistors. We also want to keep the resistors small to reduce Johnson noise. Thus, we add buffers at each input:



Suppose we have a voltage V_3 at the ground.



When V_1 and V_2 are both zero, we get $V_3/2$ at the + input and $V_{\text{out}}/2$ at the − input. Thus, $V_{\text{out}} = V_3$. In general,

$$V_{\text{out}} - V_3 = V_2 - V_1$$

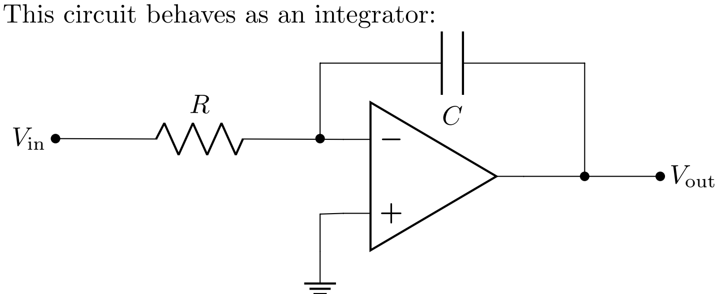
For this circuit to work well, the resistors must be matched very well, and the temperatures of the resistors should stay the same. These are often available in single chips, which works better than doing this manually.

This can be used to break a ground loop: V_2 and V_1 are treated as a differential signal, and V_3 defines the new ground.

6.3 Integration and Differentiation

6.3.1 Integration

This circuit behaves as an integrator:

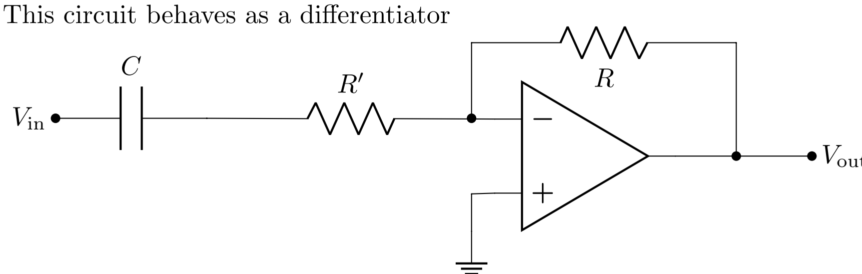


$$V_{\text{out}} = \int_0^t \frac{V_{\text{in}}}{RC}\,dt + C$$

This circuit is limited by leakage currents and non-linearity of the capacitors. Recall that high capacity ceramic capacity are quite non-linear. Some capacitors can hold a charge for days (polyster film or teflon film). However, the op amp offset voltage will be integrated, leading to the capacitor slowly charging until the op amp rails. Usually, an integrator should be used as part of a feedback loop to prevent this.

6.3.2 Differentiation

This circuit behaves as a differentiator



Ignoring R' for now,

$$V_{\text{out}} = -\frac{dV_{\text{in}}}{dt} RC$$

For op amps to be stable, the gain must be less than 1 by the time the op amp phase shift reaches 180°. The differentiator network adds a phase shift of 90°, so the op amp can oscillate at a lower frequency. Once we are at a high enough frequency that the impedance of C is negligible, the phase shift through the R-R' divider is zero, so there is less phase shift.

The other issue with differentiators is noise. The gain of a differentiator rises as $1/f$, so high frequency noise is emphasized at the output. In general, differentiators should be avoided.

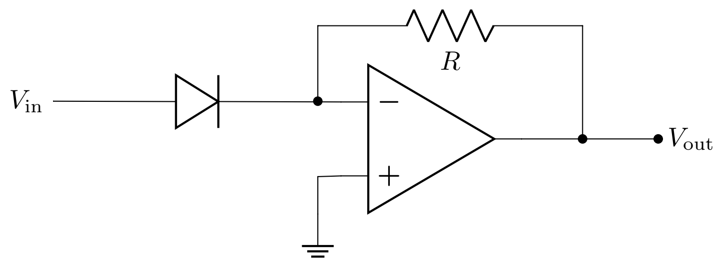
6.4 Logarithm and Exponentiation

To calculate logarithms and exponents, we use diodes. Recall thtat for a diode,

$$I_D = I_0 \left(e^{V/V_T} - 1 \right)$$

Typically, $V > 100\,\text{mV}$ and $V_T \approx 25\,\text{mV}$. Thus, we usually ignore the -1 term.

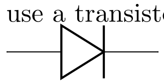
6.4.1 Exponent Calculation



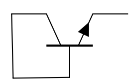
The op amp converts the current through the diode into a voltage. Thus,

$$V_{\text{out}} = -RI_0 e^{V_{\text{in}}/V_T}$$

One issue with this circuit is that the diode has a series resistance. This leads to the circuit becoming sub-exponential at large currents. To improve this, we use a transistor instead of a diode:



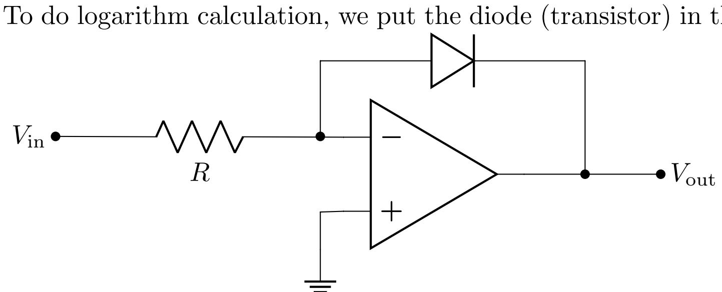
becomes



I will continue to use diodes in the diagrams, but diodes can be replaced by transistors as above throughout.

6.4.2 Logarithm Calculation

To do logarithm calculation, we put the diode (transistor) in the feedback loop.



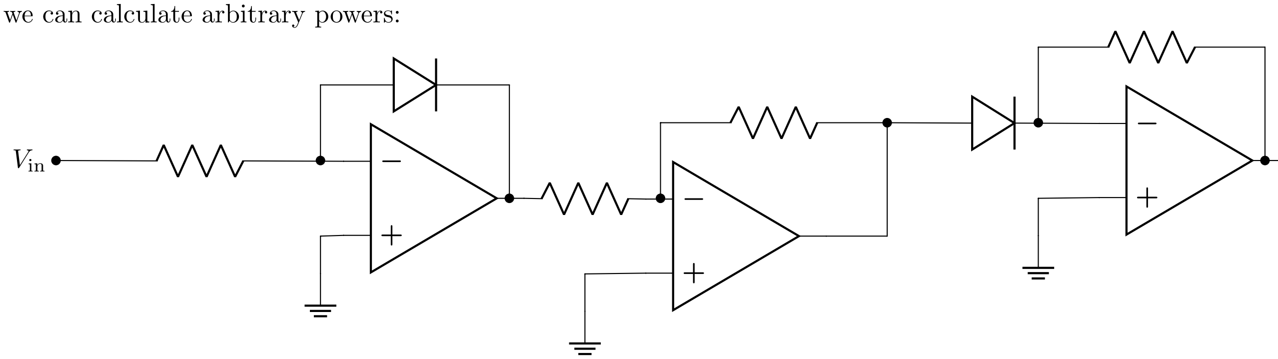
This can be used to reduce the range of a signal, which may be useful before digitizing.

6.4.3 Powers

Using

$$x^n = e^{n \ln x}$$

we can calculate arbitrary powers:



The center op amp supplies a gain of $-n$.

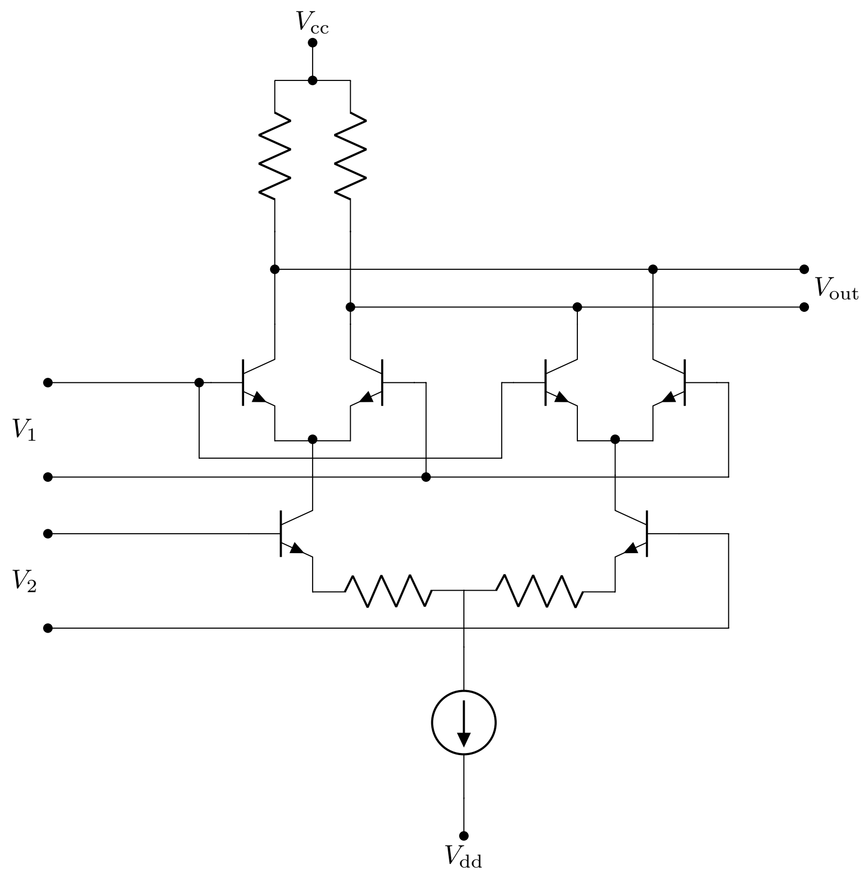
6.4.4 RMS Detector

We can combine op amps to make an RMS detector. ⟨diagram⟩ These are fine for audio frequencies, but usually insufficient in general.

6.5 Multipliers

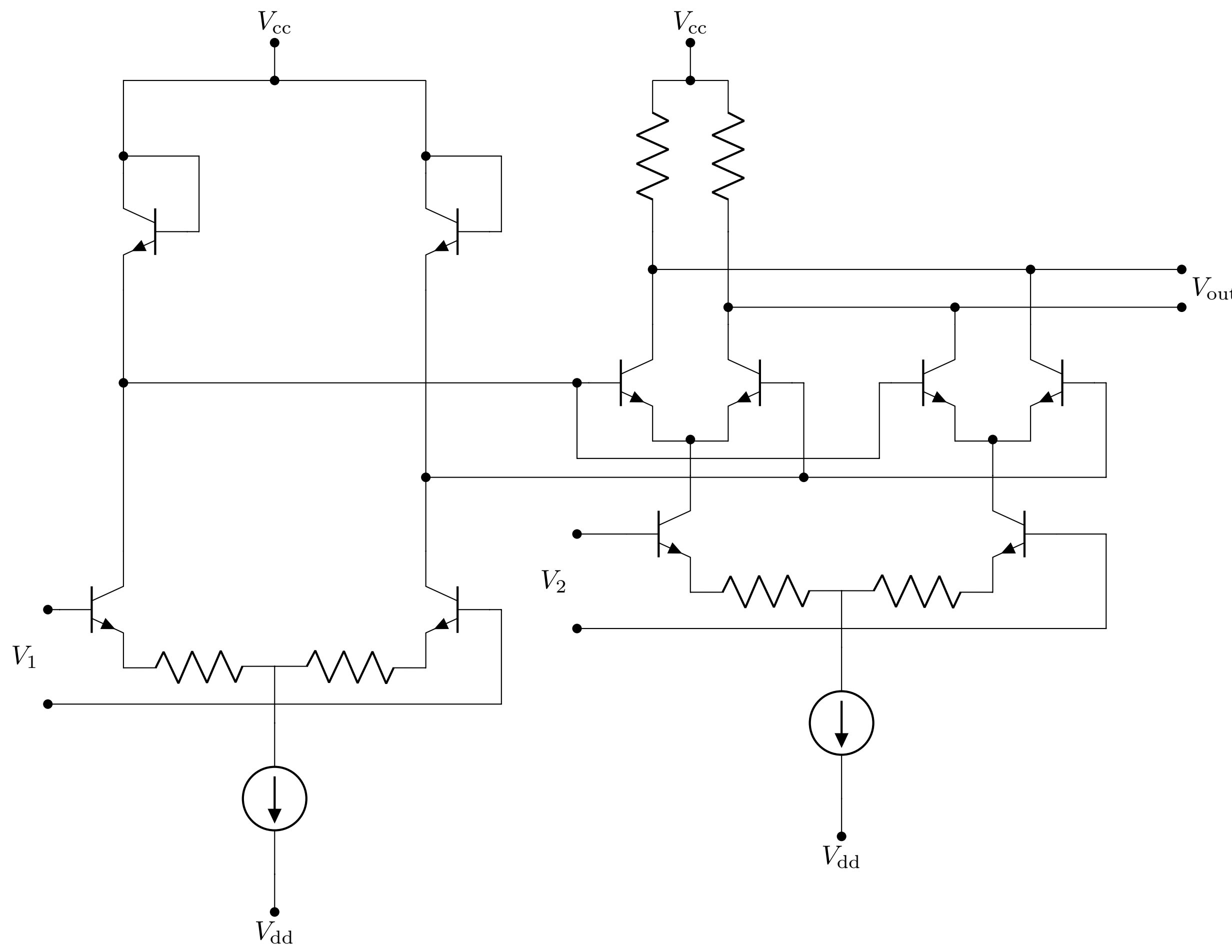
Suppose we want to detect the amplitude of interference fringes. We modulate one arm of the interferometer with a piezo. The intensity variation will depend on whether or not we are near a peak of a fringe. However, if we choose the modulation amplitude properly, we can avoid this issue. Let our modulation be $\varphi = A \sin(\omega t)$. When $A = 2.8$, the RMS photodiode signal $\langle V^2 \rangle$ is a measure of the fringe amplitude independent of the phase. Barrie Gilbert developed multipliers.

Suppose we have a set of differential pairs:



We have V_1 across the two upper differential pairs and V_2 controling the current sources. When $V_2 = 0$, then the upper differential pairs are identical, so V_{out} is zero. If $V_1 = 0$, then the differential pairs both produce zero, so V_{out} is zero. Thus, we have no linear terms and at least one product term.

To get rid of higher order terms, we add a logarithmic converter at V_1.



This circuit is very good as a multiplier, but has a low signal to noise. The AD734 is one of these.

7 Circuit Analysis

7.1 EOM Driver

Today, we will be discussing schematics from `https://arxiv.org/pdf/physics/0506050.pdf`, designed by our own Holger Müller. These are high voltage amplifiers for driving EOMs.

7.1.1

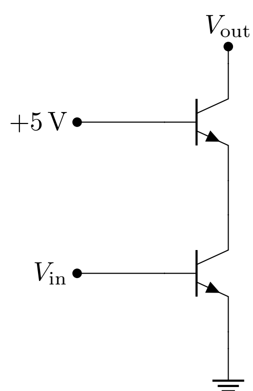
First, we will consider the circuit shown in Figure 1 of the paper. An EO is a crystal with a capacitance $10 - 100\,\mathrm{pF}$. We are considering a higher capacitance

EO, so about 100 pF. Note that a +300 V supply is presumed to be connected at the top of the schematic.

The EO modulator is connected between AP1 and AP2. Note that neither of the EO pins are grounded. The first advantage of this is that each side only needs to output 225 Vpp, instead of having one side output 550 Vpp. Thus, we can use lower voltage transistors, which are faster. The cable will have 100 pF/m for RG-58 cable (or 45 pF/m for RG-62 cable). We need some current to charge all the capacitors in the system. By having the signal ungrounded, the cable capacitors are charged to only half the voltage, and thus need less current.

T7 and T10 form a differential pair. The input is on the base of T7, the base of T10 is grounded. The portion of the circuit below that forms the current sources for the differential amplifier. T7 and T10 are BFW16A transistors, which are too low voltage to drive the output directly. They are very fast, and have substantial gain at high frequencies.

T3 and T4 form a second stage, amplifying the voltage but not the current. They are at a constant base voltage of 10 V. 10 V is used to keep the collector voltage of T7/T10 high, which reduces their capacitance. In the common base amplifier, the current gain is $1 - 1/\beta$. Due to the phase shift, at frequencies where the gain would be small in a common emitter amplifier, β has become imaginary, so the gain in the common base amplifier is still 0.7. Note that the small dependence on β means that the transistors do not need to be as well matched. We are using the cascode connection:



The lower transistor can be fast and low voltage. The upper transistor can be high voltage and slow, but it is in a common base configuration, so it will still work to a high frequency. In addition, the capacitance from collector to base of the upper transistor can be bypassed to ground through a capacitor from the +5 V to ground. This mitigates the Miller effect. Recall that the maximum collector-emitter voltage of a transistor is lowest for a disconnected base, and highest when the base is shorted to ground. This situation is equivalent to the base being shorted to ground, since leakage current flows into the supply. Thus, we have the highest voltage rating.

Ignoring the inductors for now, we have a differential amplifier with 10 kΩ output resistors and a 500 Ω resistor between the collectors. Thus, the gain is

$g = \frac{20\,\mathrm{k}\Omega}{500} = 40$. Assuming the upper cascode transistors have 10 pF capacitance to ground and our EO has a capacitance of 100 pF, we have an effective capacitance of $C_{\mathrm{eff}} = 10\,\mathrm{pF} + 2 \times 100\,\mathrm{pF}$. Thus, our maximum frequency is

$$f_0 = \frac{1}{2\pi 10^4 \Omega \times 210\,\mathrm{pF}} \approx 80\,\mathrm{kHz}$$

This is not very fast. The speed could be increased by reducing the resistors, but this would lead to more current (too much) being dissipated in the resistors. For 10 kΩ resistors, the power dissipated is

$$P = \frac{(300\,\mathrm{V})^2}{10^4\,\Omega} = 9\,\mathrm{W}$$

Increasing this is difficult. Thus, we want to add another stage to increase the current.

The current is increased with a pair of complementary emitter-followers (T1, T2, T5, and T6). These do not need voltage amplification, but do need current amplification. Suppose the emitter-follower transistors have emitter to base capacitance of 3 pF. Since there is no voltage gain, we can ignore the collector base capacitance. We have three transistors, and some capacitance to ground, so our total capacitance is 12 pF. Our upper frequency limit is thus

$$f_0 = \frac{1}{2\pi \times 1.2 \times 10^{-13} \times 10^4} = 1.4\,\mathrm{MHz}$$

Our emitter-follower transistors have $f_T \approx 100\,\mathrm{MHz}$, so they maintain a current gain of 70 at 1.4 MHz. Thus, the capacitance of the load appears to be 3 pF at the input of the emitter-followers. Thus, it will not change the speed considerably. The diodes between the pair of bases biases the emitter-follower transistors so that they are slightly open even with no signal. This reduces crossover distortions.

Finally, inductors is added between the output resistors and the emitter-follower inputs. The inductor reduces the impedance at high frequencies, allowing the speed to be pushed to 2 MHz. The inductor should be chosen to avoid creating a resonant.

f_T for the lower cascode transistors is 1.7 GHz, so they are always operating at a high gain.

A 47 Ω resistor is in front of every base (51 Ω for the inputs, which was chosen to match impedance). This is to interrupt high frequency parasitic resonances. They are added just in case, they may not be needed. Fast transistors tend to oscillate at any parasitic resonance frequency. Base current is usually low, so it is safe to put a resistor there. One should put these resistors in unless there is a good reason not to.

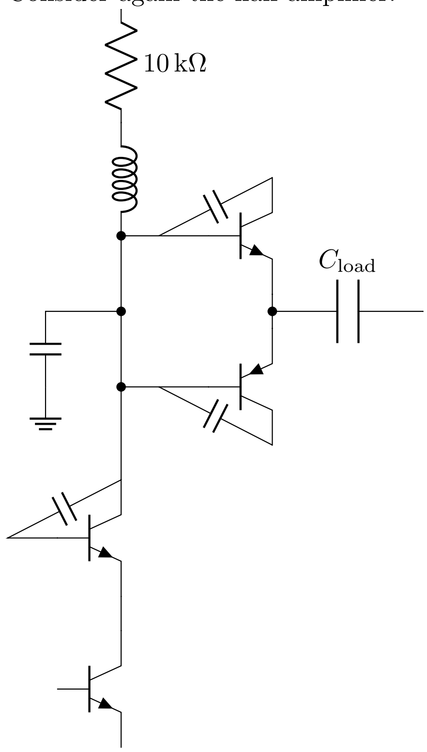
There is a capacitor in parallel to the resistor across the bottom of the differential amplifier. This reduces the impedance at high frequency, which increases the gain at high frequency.

The circuit behaves similar to what we would expect. For some reason, it had a gain of 100 in practice (the 500 Ω resistor was probably smaller).

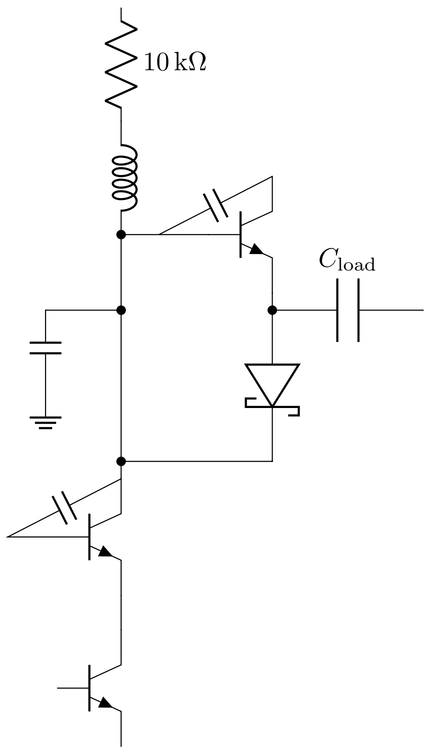
The main problem with this circuit is obtaining BF759 transistors. These used to be used for driving cathode ray tubes, which are no longer common.

7.1.2

Consider again the half-amplifier.



The speed is given by the four capacitances, one to ground and three transistors. If we could reduce the number of transistors, we could decrease the capacitance. The input transistor can be used directly instead, decreasing the total capacitance.



The second cascode transistor can now pull current from the output through the Schottky diode. For a 360 Ω resistor and faster transistors (BFQ262) with a capacitance of 1.5 pF, assuming 2 pF for wiring, the speed is

$$f = \frac{1}{2\pi 5 \times 10^{-12} 360} \approx 100\,\text{MHz}$$

This circuit has an asymmetry. The current to discharge the output needs to flow through the lower cascode transistor as well now. This can be done with a 3.3 kΩ resistor from the output to the lower cascode transistor to provide feedback. An emitter follower is then used to drive the input of the cascode pair.

The entire circuit is shown in Figure 2 of the paper. This was used to drive an EO with four crystals. It runs a bit slower than expected, only 75 MHz.

7.2 High Voltage Circuits

7.2.1 EO Driver versus Piezo Driver

EOs can operate at 100s of megahertz. Fiber EOs are generally not difficult to drive, as they need only a few volts. Free space EOs require these specialized drivers.

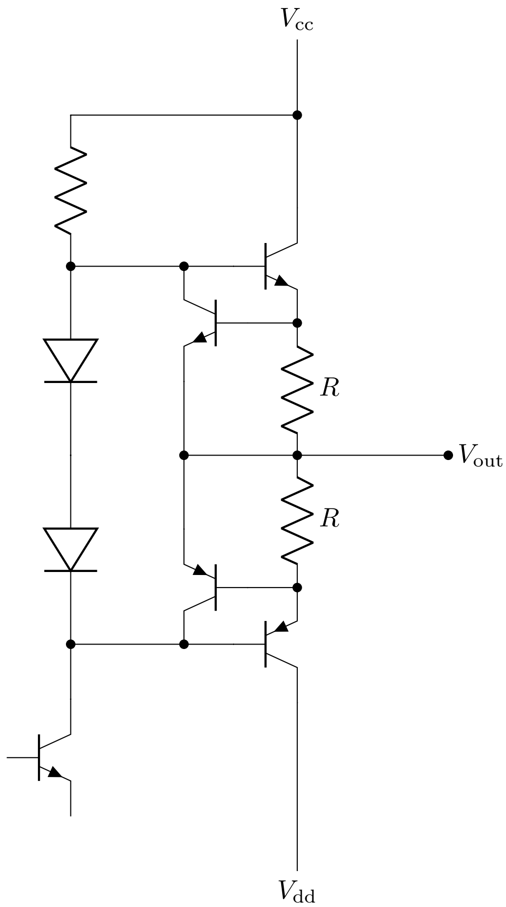
Piezos are also capacitive loads requiring high voltage. They are much larger capacitors than EOs. However, they have mechanical resonances before 1 MHz.

Thus, the driver is rarely the speed limiting factor.

For driving a piezo, we want protections rather than speed. We want to protect against short circuits (current limiting) and over-volting the piezo. For EOs, the driver has to be too fast to have these sorts of protections.

7.2.2 Current Limiting

For a complementary emitter-follower, we can implement current limiting like this:

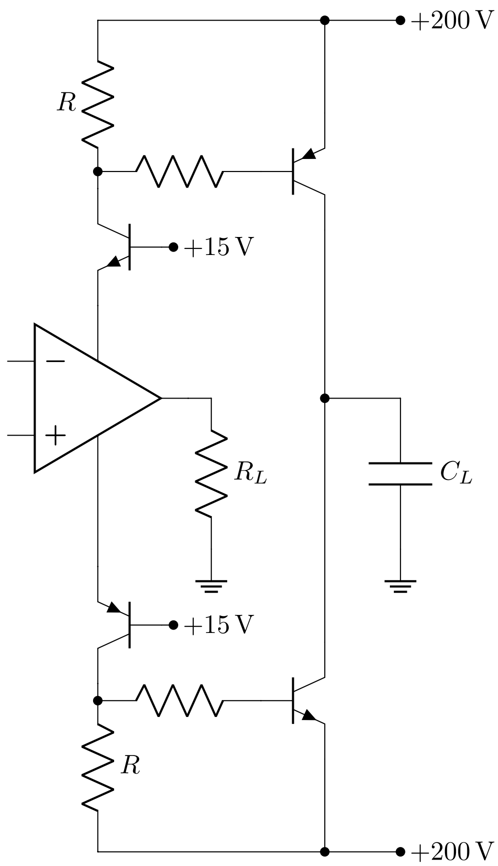


When the voltage drop is too large, there is a voltage drop across the resistors R. We choose

$$R = \frac{0.6\,\text{V}}{I_{\text{max}}}$$

This will make the transistors turn on as the current increases, which then turns the output off. This is usually included in a high voltage op amp. They can be obtained from Apex with for up to $\pm 1000\,\text{V}$ and $\pm 10\,\text{A}$. However, it is very important to have protection diodes. On the input, two diodes in opposite directions can be placed in parallel. ⟨Diagram⟩ Using JFETs instead of diodes will keep the current lost small.

7.2.3 High Voltage Op Amp

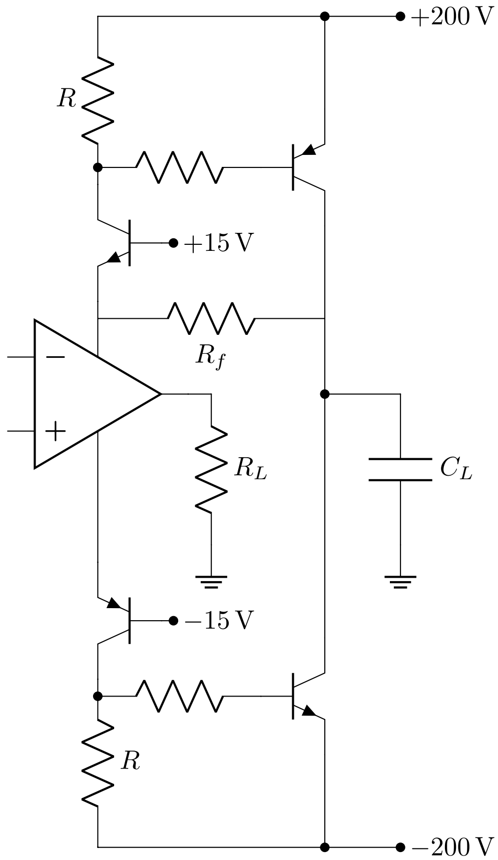


The current outputted by the op amp must be the output voltage over R_L plus some internal current used. We can use the voltage drop due to current drawn from the power supply by the op amp to drive the output complementary pair.

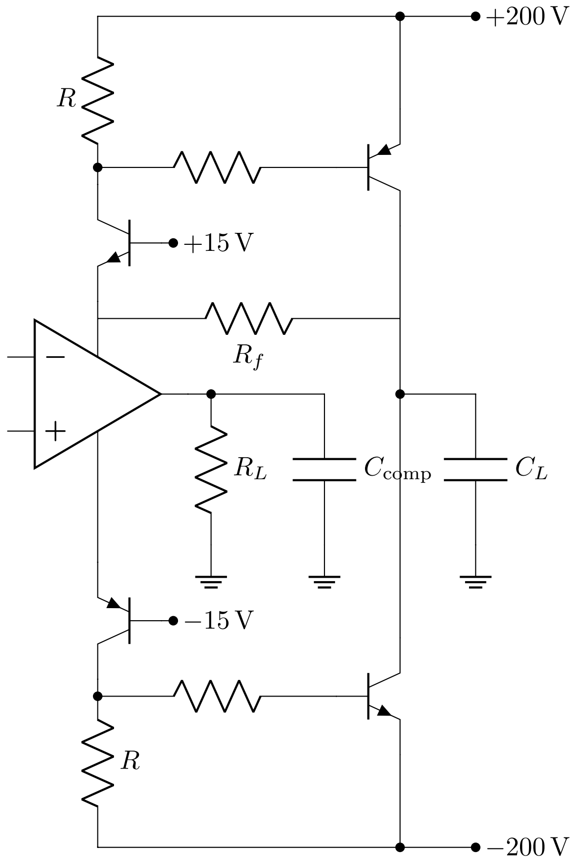
The base protection on the output transistors might be bypassed with a capacitor to improve high frequency performance.

R should be such that $R = 0.6\,\mathrm{V}/I_0$, where I_0 is op amp current needed with no load. The apparent I_0 can be increased by adding resistors from the op amp supply inputs to ground.

This circuit is not good for driving a capacitive load. We can speed up the output stage with feedback.



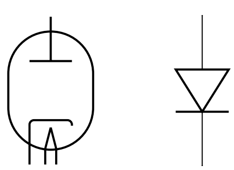
We can also increase the gain at high frequency by adding a compensating capacitor.



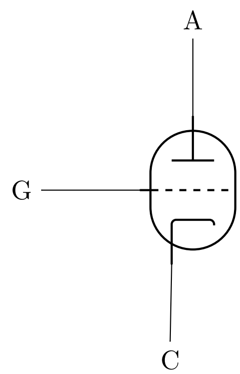
This increases the current outputted by the op amp in order to charge C_{comp}, which will increase the current used to charge C_L.

This circuit could be built with MOSFETs instead of BJTs, but the value of R would have to be adjusted such that it is able to turn on the MOSFETs.

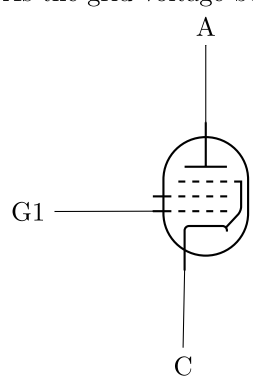
7.2.4 Vacuum Tubes



The upper node is the anode. The lower left node is the cathode. The pair in the lower center is the filament. This is roughly equivalent to a diode, but less optimal.



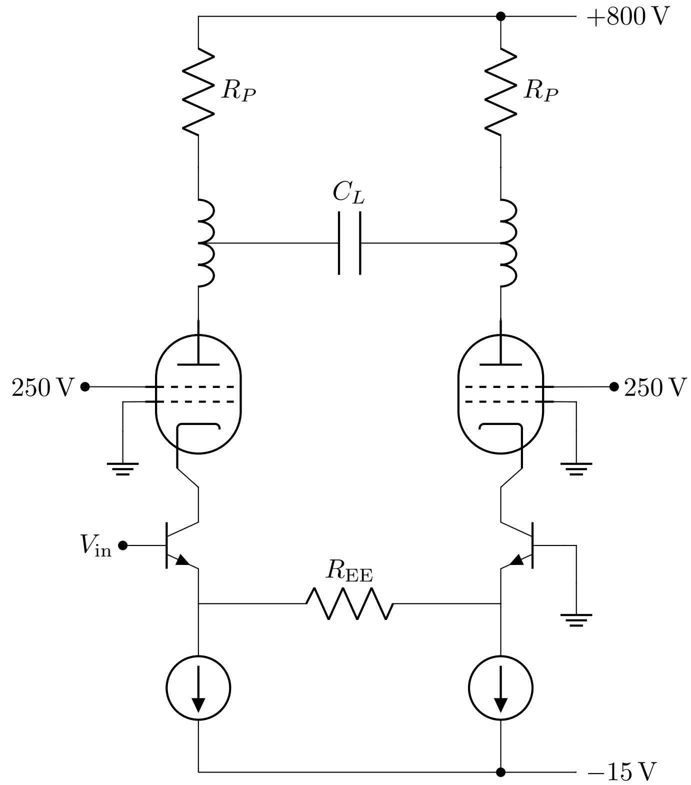
More useful than the vacuum diode is the triode. G is the grid, A is the anode, and C is the cathode. As the anode-cathode voltage increases, the current increases. As the grid voltage becomes more negative, the current is reduced. As the grid voltage becomes more positive, the current is reduced.



The pentode has two grids. The second grid is at a fixed offset voltage relative to the cathode. This has a characteristic curve similar to a field effect transistor. Like a JFET, G1 should be negative, which will keep the grid current zero.

The largest vacuum tube can run at 25 kV, 200 A, and 2.5 MW. For so much power, the capacitance is relatively small, 100s of picofarads. This is smaller than the equivalent constructed from parallel MOSFETs. Modern vacuum tubes are not glass, but tend to have integrated cooling.

We can built an EO driver with vacuum tubes. Suppose we replace the cascode transistor in the previous EO driver with a vacuum tube:



The vacuum tube is a kT120, which has a maximum power dissipation of 60 W. Thus, we want the anode current to be 111 mA. The voltage across the T-coil circuit is 450 V. Thus, $R_P = \frac{350\,\mathrm{V}}{111\,\mathrm{mA}} = 3.3\,\mathrm{k\Omega}$

We could use this to drive a Pockels cell, which requires 1.2 kV. The capacitance of the cell is 4 pF. Our total load capacitance (assuming 1 m of RG-62 cable) is $C_l = 2 \times (4\,\mathrm{pF} + 46\,\mathrm{pF}) \approx 100\,\mathrm{pF}$. We can replace our RG-62 cable with a twisted pair. This will still have electrical radiative shielding, but this falls off quickly. Using a twisted pair minimizes the magnetic radiation. Thus, our total capacitance is now $C_L \approx 30\,\mathrm{pF}$.

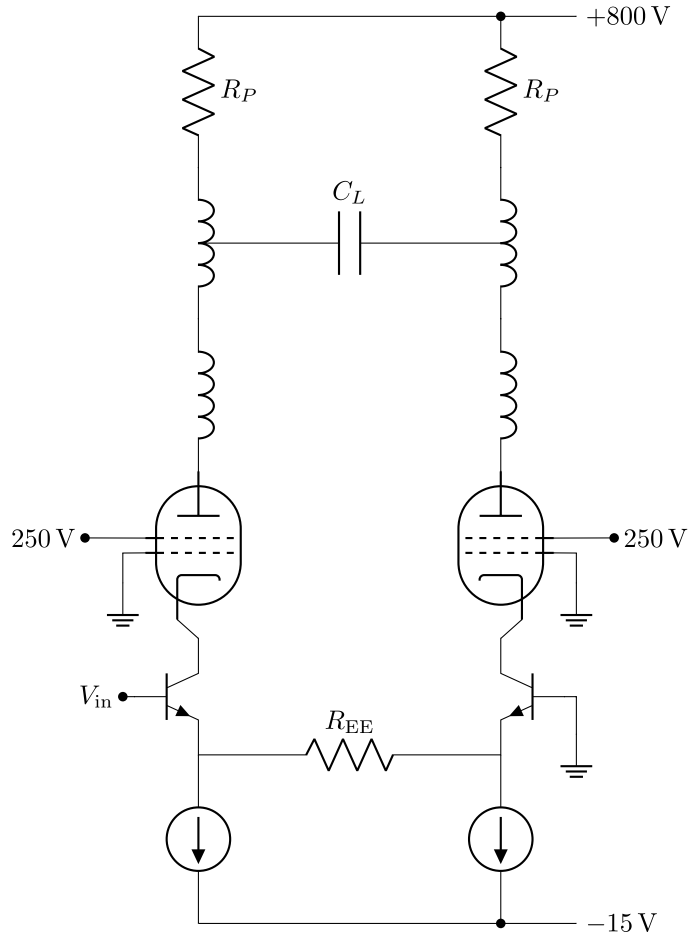
The upper speed limit, without the T-coil is

$$f_0 = \frac{1}{2\pi \times 3.3\,\mathrm{k\Omega} \times 30\,\mathrm{pF}} \approx 1.6\,\mathrm{MHz}$$

Since we have T-coil compensation, it is 2.8 times as fast, so

$$f_0 \approx 5\,\mathrm{MHz}$$

The plates of the vacuum tubes will have a spare capacitance of about 10 pF. However, we can add an inductor, which increases the speed up to 4:



The voltage gain is $\frac{2R_p}{R_{\text{EE}}}$.

The transistors do need to be 100 V, 2 A, but that is possible to find.

An equivalent MOSFET to the vacuum tube would have 100s of picofarads of capacitance, slowing down the circuit substantially. The vacuum tubes here provide voltage amplification rather than current amplification. The transistor only needs to supply 30 Vpp, which the vacuum tubes then amplify.

8 Noise

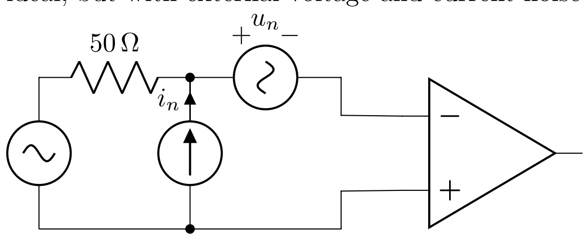
The electrical kind.

8.1 Op Amp Noise

An op amp datasheet will have two noise specifications. One is the voltage noise, u_n, in $\mathrm{V}/\sqrt{\mathrm{Hz}}$. This could be around $10\,\mu\mathrm{V}/\sqrt{\mathrm{Hz}}$. The other is the current noise, i_n, in $\mathrm{A}/\sqrt{\mathrm{Hz}}$. This could be around $1\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$.

Suppose we have a signal source with some internal impedance R that we want to send into the amplifier. Noise in the signal source is usually due to Johnson noise in the resistor. The voltage noise spectral density of an open resistor is $u_n = \sqrt{4k_BTR}$, where k_B is the Boltzman constant. The noise spectral density goes down for low resistances. It is about $1\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ for a $50\,\Omega$ resistor. The corresponding current noise spectral density is $i_n = \sqrt{\frac{4k_BT}{R}}$. It is about $20\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ for a $50\,\Omega$ resistor.

Suppose we have a $50\,\Omega$ signal source. We will consider our op amp to be ideal, but with external voltage and current noise sources.



The total noise is

$$u_{n,\,\mathrm{total}}^2 = 4k_BTR + u_n^2$$

The noise factor is the total noise over the source noise:

$$\frac{P_{\mathrm{noise}}}{P_{\mathrm{source}}} = \frac{4k_BTR + u_n^2}{4k_BTR} = NF = F$$

If our amplifier has noise $1\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ and we have a $50\,\Omega$ resistor, $NF = 2$ or $NF = 3\,\mathrm{dB}$.

Suppose we also have current noise. This produces a voltage across R. Thus, our total noise is

$$u_{n,\,\mathrm{total}}^2 = 4k_BTR + u_n^2 + (i_nR)^2 \text{ and } NF = \frac{4k_BTR + u_n^2 + (i_nR)^2}{4k_BTR}$$

This function will have some minima in terms of R.

For low impedance signals, minimize u_n. For high impedance signals, minimize i_n (use a FET op amp).

8.1.1 FET Noise

It is useful to write u_n in terms of an equivalent resistor:

$$u_n = \sqrt{4k_BTR_{\mathrm{eq}}}$$

For FET transistors, $R_{\mathrm{eq}} \approx \frac{0.65}{g_m}$ (where g_m is the transconductivity). This holds best for frequencies above $1\,\mathrm{MHz}$. At lower frequencies, there is $1/f$ noise. JFETs have less of the $1/f$ noise, and thus should be chosen when minimizing $1/f$ is important. For a FET, $i_n = \sqrt{2eI_g}$, where I_g is the gate current. This is usually negligible. This equation can be used to determine the current noise from the input offset current. I_g is the input offset current for a op amp. The current noise is usually negligible, so it is good to use FETs when the signal impedance is large.

8.1.2 BJT Noise

For a BJT, $R_{\text{eq}} \approx \frac{1}{g_m}$. Unlike a FET, g_m varies with the collector current:

$$R_{\text{eq}} \approx \frac{1}{g_m} = \frac{1}{I_C/V_T} = \frac{V_T}{I_C} = \frac{e}{k_B T I_C}$$

(V_T is the thermal voltage). The current noise density is $i_n = \sqrt{2eI_B} = \sqrt{2eI_C/\beta}$, where I_B is the base current. Note that as the collector current increases, the voltage noise decreases. However, the current noise will increase. In total,

$$\begin{aligned}
\frac{u^2_{n,\text{ total}}}{u^2_n} &= \frac{4k_B TR + 4k_B T R_{\text{eq}} + R2eI_C/\beta}{4k_B TR} \\
&= \frac{4k_B TR + 4k_B T \frac{e}{k_B T I_C} + R2eI_C/\beta}{4k_B TR} \\
&= \frac{4k_B TR + 4\frac{e}{I_C} + R2eI_C/\beta}{4k_B TR}
\end{aligned}$$

For large current, the noise rises linearly. For small current, the noise rises as $1/I_C$. The only parameter we can choose is β. The noise is smallest when β is large. For this, we use a "super-β transistor". We then try to operate at the minimum noise I_C. β decreases as we approach the transition frequency, so the noise will tend to be higher at higher frequency.

For an op amp, the noise is typically given as a function of frequency in graphs. There will usually be a white noise region and a $1/f$ region. The noise at high frequencies will go up, but op amps are not fast enough for that to be relevant.

8.2 RF Component Noise

8.2.1 Attenuators

Attenuators add noise. The output of an attenuator has the full thermal noise, but decreases the signal. Thus, the NF of a 6 dB attenuator is 6 dB (assuming the input has only thermal noise).

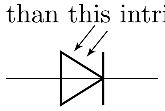
8.2.2 Mixers

A mixer usually has some conversion loss. Suppose we have f_{RF} at the input and f_{LO} at the local oscillator. Then we have $f_{\text{RF}} \pm f_{\text{LO}}$ at the output. Let us assume the mixer is ideal. If we use only one of the output frequencies, the mixer behaves as a 3 dB attenuator.

In practice, mixers are usually $4-5$ dB due to additional noise. This can be prevented with filters: the input and output can be sent through narrowband filters. A filter at the output will reflect the unwanted frequency back into the mixer, which will eventually convert it to the desired frequency.

9 Photodetectors

Light is made of photons. Due to this, there must be *shot noise* in any light detection scheme. The goal of a good detection circuit is to produce less noise than this intrinsic noise.



9.1 Shot Noise

The number of photons per second is $\frac{P}{h\nu}$ where P is the power and ν is the frequency. From Poisson statistics, the noise is $\sqrt{\frac{2P}{h\nu}}$. Noise appears as sidebands. When detecting, we detect intensity, which gives us a factor of 2 (due to sidebands combining). The power flucuation per $\sqrt{\text{Hz}}$ is

$$\Delta P = \sqrt{\frac{2P}{h\nu}}h\nu = \sqrt{2Ph\nu}$$

So, the total power flucuation is

$$\Delta P = \sqrt{2Ph\nu B}$$

where B is the bandwidth.

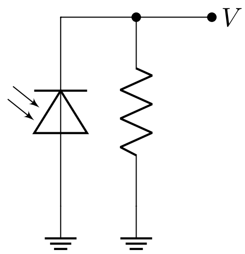
In an ideal photodiode, we would have one electron per photon. In reality, there is a factor η (the quantum efficiency) which is less than 1. However, we can neglect this without substantial changes.

So, the current flucuation is

$$\Delta I = \sqrt{2Ph\nu B e} = \sqrt{2eIB}$$

9.2 Simple Photodiode Circuit

One way to measure the signal is to simply add a resistor in parallel:



Then the signal is simply $V = I_D R$. The Johnson noise in the resistor is

$$i_n = \sqrt{\frac{4k_B T}{R} B}$$

The diode also has a capacitance, C_D, which limits the circuit bandwidth to

$$B = \frac{1}{2\pi C_D R}$$

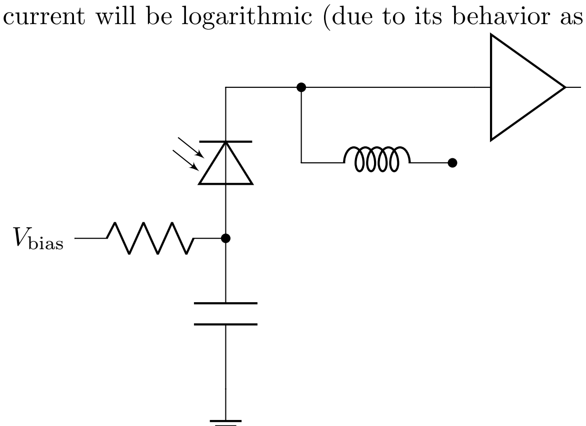
Thus, we have limited bandwidth before the Johnson noise becomes too large.

Consider the case where the Johnson noise equals the shot noise:

$$\sqrt{2eI_D} = \sqrt{\frac{4k_B T}{R} B} \implies \frac{2eI_D R}{4k_B T} = 1$$

Recall, that $\frac{e}{k_B T} = V_T \approx 25\,\text{mV}$. $V = I_D R$, so the Johnson noise equals the shot noise when $V = 2V_T = 50\,\text{mV}$. 1 mW at 1064 nm produces 1 mA, so for a 50 Ω resistor, we do get a roughly shot noise circuit.

We also want to bias the photodiode. Without biasing, the photodiode current will be logarithmic (due to its behavior as a diode).



The amplifier could be a minicircuits ZFL-500KN, which has a noise figure of 3 dB. Note that the 50 Ω resistor is inside the amplifier. The total noise is shot noise of $1\,\text{mA} = \frac{1\,\text{pA}}{\sqrt{\text{Hz}}}$ and Johnson noise of $50\,\Omega = \frac{1\,\text{pA}}{\sqrt{\text{Hz}}}$. We multiply the

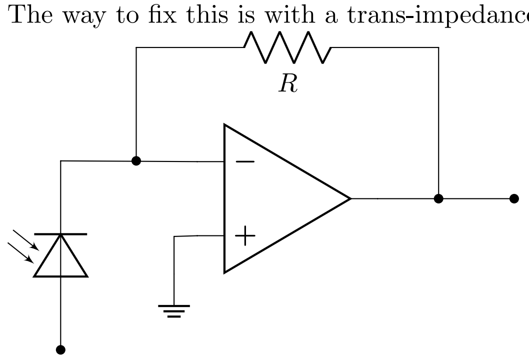
Johnson noise by the noise figure of the amplifier, $3\,\mathrm{dB} = \sqrt{2}$, so the total noise is about $1.8\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$. If we have 1 mW of power, it is not hard to be shot noise limited.

9.3 Low Frequency, Low Noise: Trans-impedance Amplifiers

Suppose we want to detect 1 nW. This would be 1 nA, so to get 50 mV, we would need $R = 50\,\mathrm{M\Omega}$. Let us choose $R = 100\,\mathrm{M\Omega}$ to reduce noise further. Suppose the diode capacitance is 10 pF (reduced from 20 pF using a bias of $-15\,\mathrm{V}$). Then our bandwidth is

$$B = \frac{1}{2\pi 10^{-8} 10^{8}} \approx 160\,\mathrm{Hz}$$

The way to fix this is with a trans-impedance amplifier:



This works because the op amp holds its $-$ pin at 0 V, so the photodiode capacitance is not charged. The new bandwidth is

$$V = \sqrt{\frac{1}{2\pi RC} \times f_T}$$

where f_T is the op amp transition frequency. So, for the same resistor and capacitance with an op amp with 10 MHz bandwidth, $B = 28\,\mathrm{kHz}$.

The total current noise is

$$i^2_{\text{noise}} = i^2_{\text{shot}} + i^2_{\text{johnson}} + i^2_{\text{op amp}} + \dots$$

Recall that $i_{\text{op amp}} = \sqrt{2eI_{\text{offset}}}$. To keep this small, we want to use an op amp with a FET input, as that keeps I_{offset} small. Choosing the right op amp, this can be negligible. However, the op amp also has a voltage noise u_n. We can model this as noise on the + input. The op amp will amplify this noise, producing it at the output. The current noise produced at the output is then

$$\frac{u_n}{R}\frac{R}{1/i\omega C_D} = u_n R i\omega C_D$$

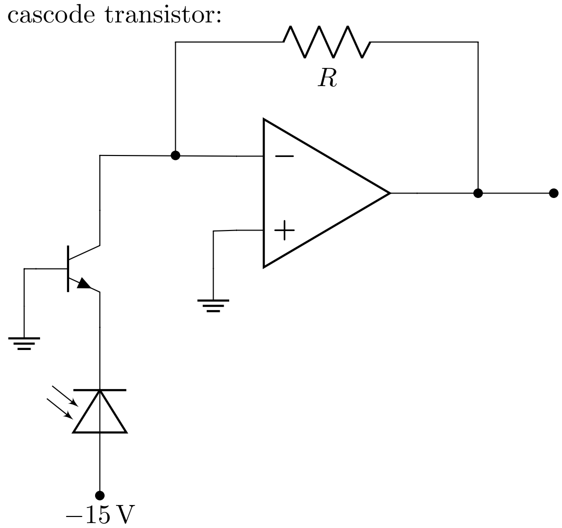
Our total noise is thus

$$i^2_{\text{noise}} = i^2_{\text{shot}} + i^2_{\text{johnson}} + i^2_{\text{op amp}} + (u_n R\omega C_D)^2$$

$u_n \approx \frac{10\,\text{nV}}{\sqrt{\text{Hz}}}$ for a FET op amp. The Johnson noise is $0.74\,\text{fA}/\sqrt{\text{Hz}}$, the shot noise is $1\,\text{fA}/\sqrt{\text{Hz}}$, $i_{\text{op amp}} = 0$, and $u_n\omega C_D = \frac{0.01\,\text{fA}}{\sqrt{\text{Hz}}}$ at 1 Hz, but $\frac{10\,\text{fA}}{\sqrt{\text{Hz}}}$ at 1 kHz, and $\frac{100\,\text{fA}}{\sqrt{\text{Hz}}}$ at 10 kHz. Thus, the op amp voltage noise dominates at high frequencies, and the circuit is not shot noise limited over most of its bandwidth.

9.3.1 Reducing High Frequency Noise

P.C.D. Hobbs developed a solution to this. It uses a structure similar to the cascode transistor:



We need a transistor with $\beta >> 1$ at I_D and a very small capacitance (this would be an RF transistor, with $f_T \approx 50\,\text{GHz}$). Although the diode has some capacitance, the op amp only sees the capacitance of the transistor, which is smaller. To keep the capacitance low, the transistor should be attached directly to the op amp pin, and should be fairly far from the ground plane.

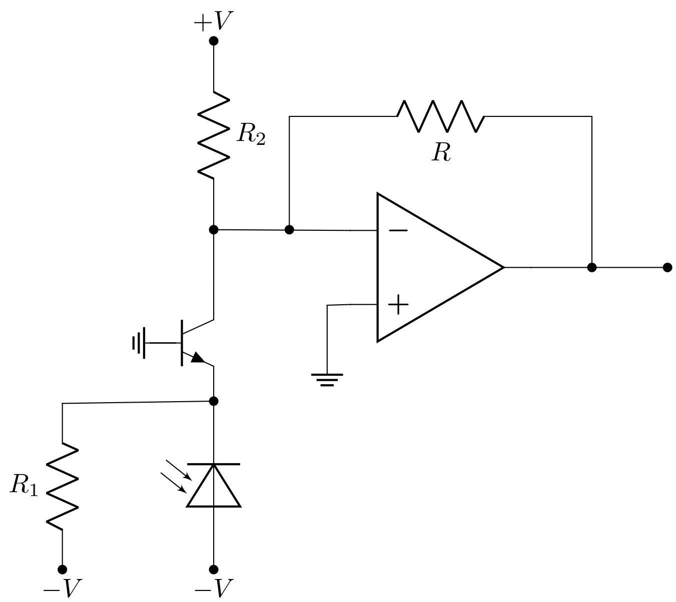
Now, the op amp only sees $C_T \approx 1\,\text{pF}$. Thus, the bandwidth is

$$B = \sqrt{\frac{1}{2\pi RC_T} f_T} = 120\,\text{kHz}$$

We also now have $u_n\omega C_D \to u_n\omega C_T$, which will be smaller, $0.5\,\text{fA}/\sqrt{\text{Hz}}$ at 1 kHz and $5\,\text{fA}/\sqrt{\text{Hz}}$ at 1 kHz. However, our bandwidth assumed that no current flowed into the diode capacitance. The transconductance is $g_m = I_D/V_T$. The bandwidth of the transistor circuit is $B = \frac{g_m}{2\pi C_D}$. For 1 nA, this is about 300 Hz.

9.3.2 Increasing Bandwidth, Increasing Transconductance

We can increase the bandwidth by increasing the transconductance. We increase the transconductance by increasing I_D:



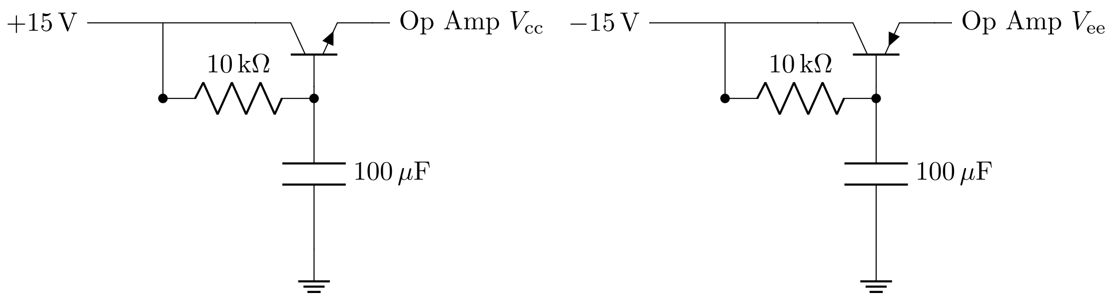
However, to keep the noise low, we want R_1 and R_2 to be fairly large, say $R_1 = R_2 = 500\,\mathrm{M\Omega}$. This would require a very high voltage to get 1 nA of current. However, this can be done with high voltage batteries. Since the current is small, they will last fairly long.

There is a way to produce the bias voltage without very high supply voltages, but we will not cover this for now. Just using an active current source does not help, as an active current source produces its own noise.

9.3.3 Practical Advice

Selection of the transistor is important. The transition frequency is usually not specified at low current. Thus, a very high transition frequency should be chosen such that it remains reasonable at low current.

This circuit is very sensative to op amp power supply pin noise. Usually, the op amp has a high power supply rejection ratio (PSRR). However, the PSRR assumes low impedance at the inputs. Thus, the PSRR is lower. The 78xx and 79xx regulators are too noisy in this application. To fix this, a low pass filter could be added after the regulator, but the resistor must be large, which limits the power. Thus, the low pass filter is used with a transistor.



The 2N3904 and 2N3905 transistors work well for this.

The capacitance parallel to the feedback resistor R should be minimized.

A $50\,\Omega$ resistor should be placed in series with the op amp output. Op amps do not like being connected directly to cables.

9.4 General Advice

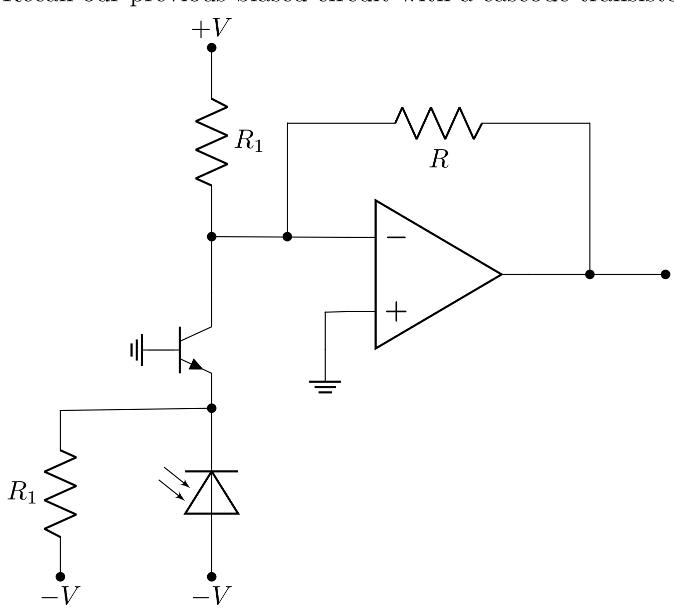
For $I_D \geq 1\,\mathrm{mA}$, just us a $50\,\Omega$ resistor.

For $I_D \geq 1\,\mathrm{nA}$, use a trans-impedance circuit, as discussed.

For $I_D \approx 1\,\mathrm{pA}$, we would need $R = 50\,\mathrm{G}\Omega$ to get $50\,\mathrm{mV}$. At this point, the shot noise is $3 \times 10^{-17}\,\mathrm{A}/\sqrt{\mathrm{Hz}}$. The op amp current noise is now relevant. The transistor trick is no longer useful, as the circuit bandwidth will be low enough that the noise increase at high frequencies is not important.

9.5 Further Improvements

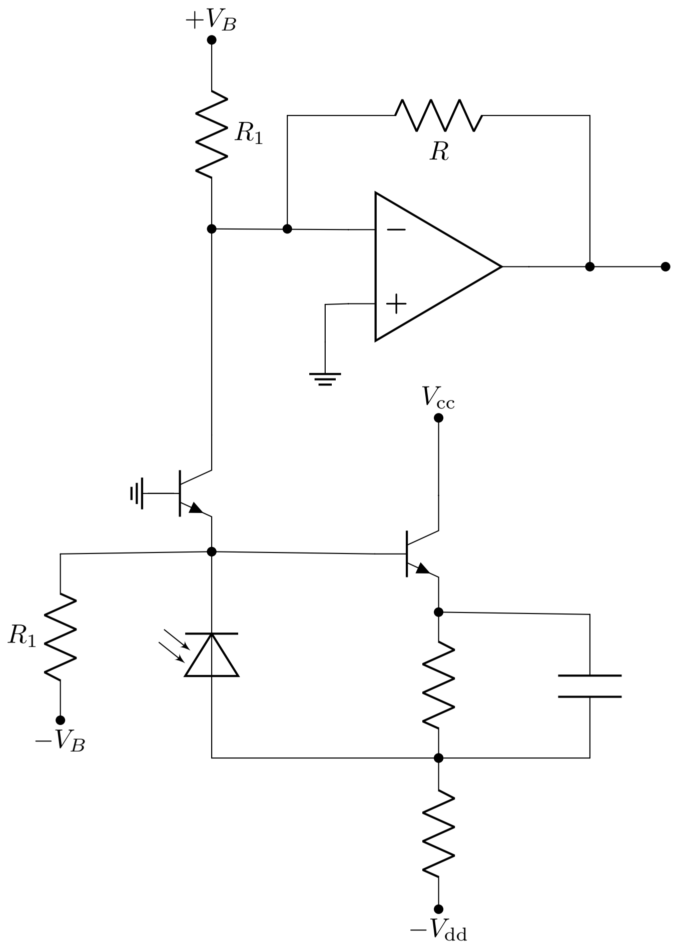
Recall our previous biased circuit with a cascode transistor:



Recall also that to reduce noise, we need $R_1/2 > R$.

The base current through the transistor is given by $\frac{1}{\beta}(I_D + I_B)$, where I_D is the current through the diode and I_B is the bias current. This will have full shot noise of $\sqrt{2e\frac{I_D+I_B}{\beta}}$. Most of the time, a current will not have shot noise. Transistors, diodes, and semiconductor devices have shot noise. Resistors do not.

Hobbs improved this further



The photodiode has some capacitance, which we ignored because it is never charged. The additional circuit is known as the bootstrap circuit, and it reduces the diode effective capacitance further. ⟨Explain why this is better⟩

The second transistor has some base current too, $I_{B,2}$. To avoid shot noise from this current being an issue, we need $I_{B,2} << I_D$.

What is the total noise? Define $R_1 \parallel R_2 = \frac{R_1 R_2}{R_1+R_2}$. Then the noise contains $\sqrt{\frac{4k_BT}{R\|R_1R}}$, $i_{n,\,\mathrm{op}}$, $u_{n,\,\mathrm{op}}$, $u_{n,\,\mathrm{op}}\omega C_T$, $\sqrt{2e\,(I_{B,1}+I_{B,2})}$, and $\sqrt{2eI_D}$.

A rule of thumb is that the overall bandwidth for two lowpass filters is

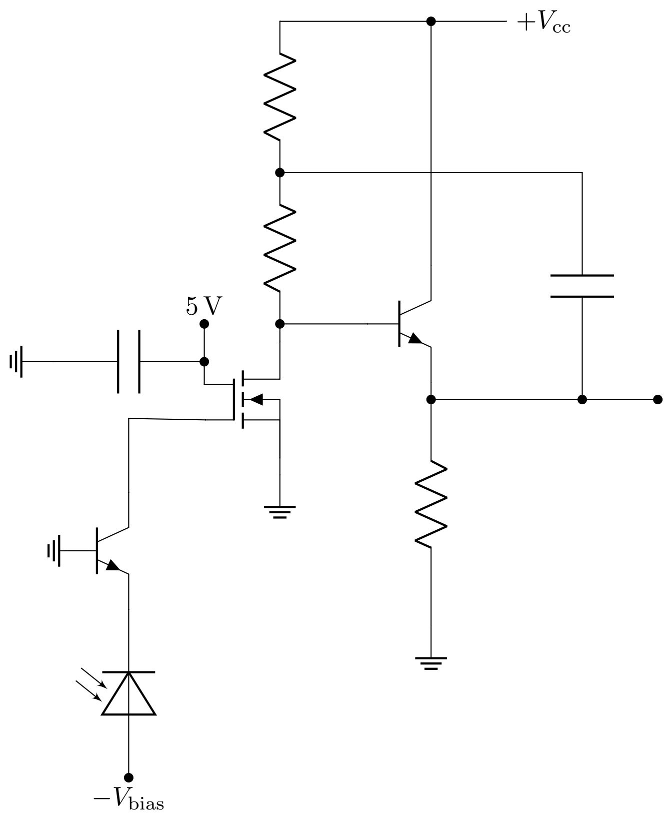
$$B^{-2} = B_1^{-2} + B_2^{-2} + \dots$$

This is exact when the response of the filters is Gaussian.

For this circuit, we have the bandwidth of the op amp: $B_{\text{op}} = \sqrt{f_T \frac{1}{2\pi R C_T}}$ (where C_T is the transistor parasitic capacitance, which does include the wiring), the bandwidth of the transistor: $B_T = \frac{1}{2\pi g_m^{-1} C_{D,\,\text{eff}}}$ (where $C_{D,\,\text{eff}}$ is the diode effective capacitance). $C_{D,\,\text{eff}} = C_D\,(1 - g_B)$, where g_B is the gain of the bootstrap follower. Since we are using a low emitter current, this may be lower than 1, perhaps only 0.9.

9.5.1 Improving Op Amp Noise

To improve the op amp noise, we can use a dual gate MOSFET with one gate set to a constant voltage.



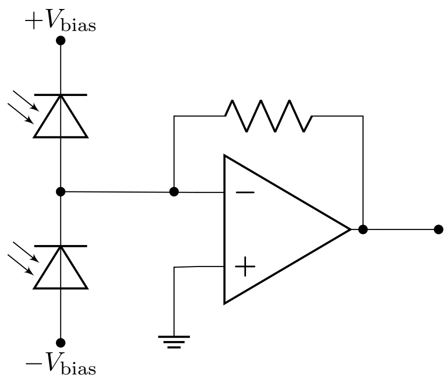
Having the second gate decreases the gate-drain capacitance. The gate replaces the − terminal on the op amp previously used. The transconductance of the MOSFET is g_m between 20 and 30 mA/V. The voltage noise will be less than $1\,\text{nV}/\sqrt{\text{Hz}}$. The transistion frequency could be around 2 GHz.

9.6 Detecting Intensity Ratios

We often want to detect the laser intensity after some sample. However, the laser may have some residual intensity noise (RIN). We could feedback on that noise to keep it low, but a better method might be to somehow determine the ratio of the signals.

Division can be noisy though. At low frequencies, RIN may be 40 dB to 60 dB higher than shot noise. Detectors and electronics have some frequency dependence. In order for the division to be reliable, these must all be matched at all frequencies. This is not possible.

Subtraction is easier.

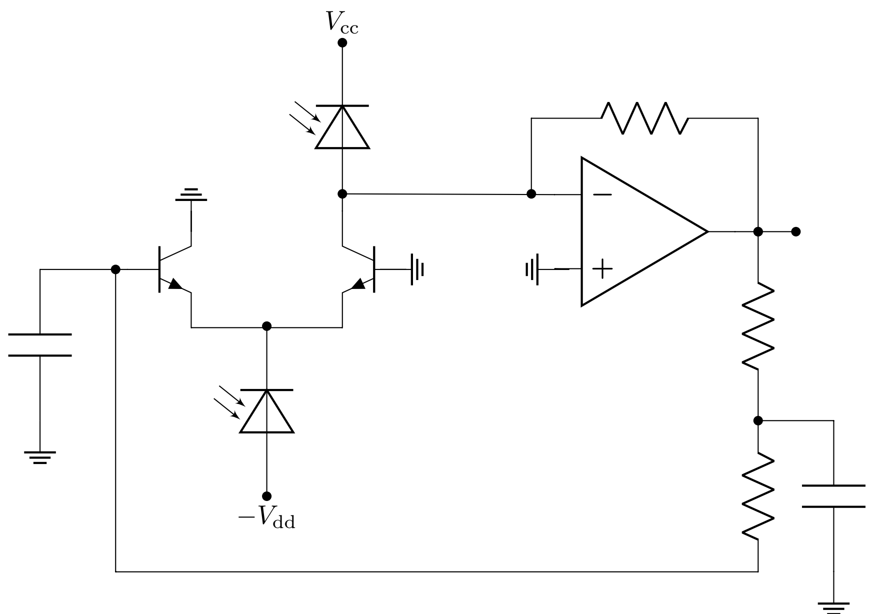


This does require that the intensities be approximately equal, but that can be done with wave plates and beam spltting cubes. The subtraction will work to quite high frequencies, as the op amp only sees the difference.

This circuit can be improved similarly to the previous circuit, except that the upper diode would need a PNP cascode transistor.

9.6.1 Auto-Balancing Differential Detector

It would be nice to have the circuit balance itself, so that we do not need waveplates. This could be done by diverting some current from one of the diodes. Phil Hobbs found a way to do this.



The transistors should be a matched pair.

The feedback balances the signals. The op amp voltage rises when there is too much current being pulled out of the op amp. This would be coming from the lower diode. To undo that, the left transistor needs to have more current, so the base voltage should be increased. Thus, we can connect the op amp output to the base of the left transistor.

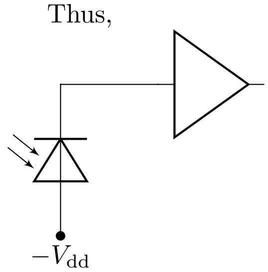
However, this requires that the signal be faster than the feedback, as otherwise the feedback will eliminate the signal. The low pass filter in the feedback path sets the feedback speed. The second capacitor is close to the base of the transistor to reduce inductance.

A PNP cascode transistor can be added below the upper photodiode.

9.7 Much Faster

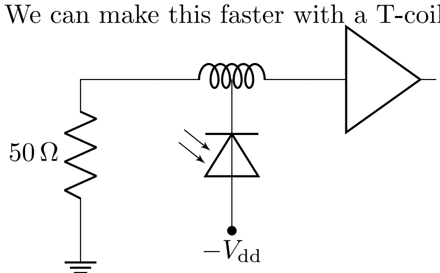
So far, we have assumed that the op amp is much faster than the signal. What if the signal is even faster, 500 MHz to a few GHz? Faster op amps usually require higher minimum gains.

Thus,



$$B = \frac{1}{2\pi \times 50 \times C_D}$$

We can make this faster with a T-coil circuit:



$$B = \frac{2.8}{2\pi \times 50 \times C_D}$$

To make things even faster, we need a photodiode with gain, such as the avalanche photodiode. These are operated close to where the diode breaks down, which allows it to amplify incoming photons. However, this adds additional noise and reduces the detective quantum efficiency. The temperature of an avalanche photodiode usually must be monitored.

Another option is the photomultiplier tube. These have very low dark current, but lower quantum efficiency.