ELECTRONIC SYSTEMS

Luca Colombo | Politecnico di Milano | a. a. 2022/2023

DISCLAIMER

These notes cover the arguments of the course 'Electronic Systems' held by Professor F. Zappa at Politecnico di Milano during the academic year 2022-2023.

Since they have been authored by a student, errors and imprecisions can be present.

These notes don't aim at being a substitute for the lectures of Professor Zappa, but a simple useful tool for any student (life at PoliMi is already hard as it is, cooperating is nothing but the bare minimum).

Please remember that for a complete understanding of the subject there is no better way than directly attending the course (DIY), which is an approach that I personally suggest to anyone. Indeed, the course is really enjoyable and the professor very clear and helpful.

In any case, if you found these notes particularly helpful and want to buy me a coffee for the effort, you're more than welcome: https://paypal.me/LucaColomboxc

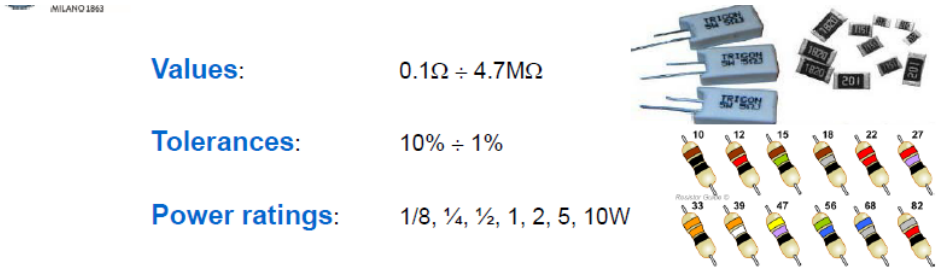
lucacolombo29@gmail.com

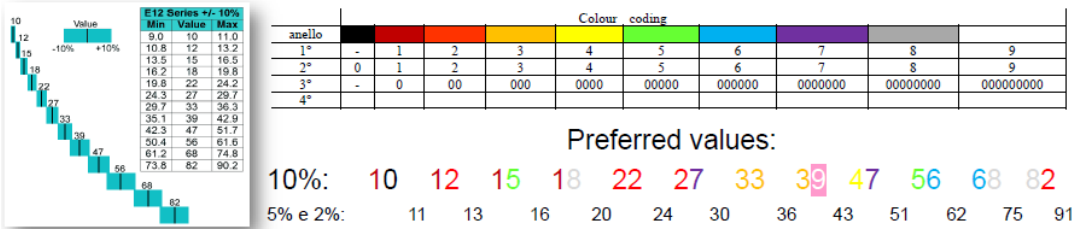
BASICS ON ELECTRONICS

RESISTORS

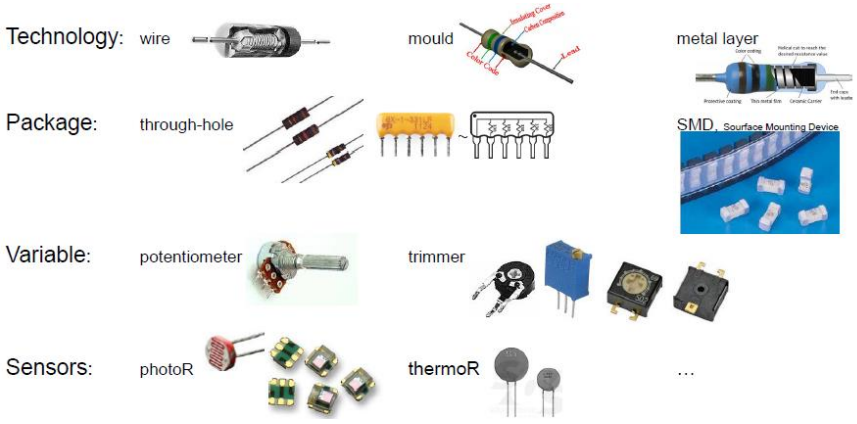
Typical resistors' values spans from 0.1 Ohm up to few MOhms (eventually also Giga-Ohm). Then all components have their own tolerance, because every manufacturing mechanism has its tolerances. Cost of course changes a lot with tolerances.
Moreover, we have also power dissipation; we have also to specify the power rating of a resistor. The power rating is the power the resistor is going to dissipate under operation. It is obtained with V*I, that if V = 5V and I = 1mA is e.g. 5W (NB: the equation P = V^2/R = R*I^2 can be used only if the component is a resistor).

I have to specify the power rating because if I use the same resistor in a different circuit with different voltages or currents, the power dissipation is different. If the amount of power to be dissipated is too high, the resistor could break → in principle I would need a resistor with a higher power dissipation but with the same resistor value. Usually, the power value expected must be overrated by the resistor by a factor 3/2 (150%).





Moreover, we have only some preferred values of resistors, because other values can be reached with series and parallel between resistors. Moreover, due to tolerances, it might happen that, for instance, the 2k resistor and the 2.5k resistor have the same value (2.2k). So values of manufactured resistances' values are chosen so that the spreads due to tolerances don't overlap one with the other passing from one value to the other → also better to design electronics that doesn't work only with a specific value of resistance, or that doesn't change if the resistor drifts.



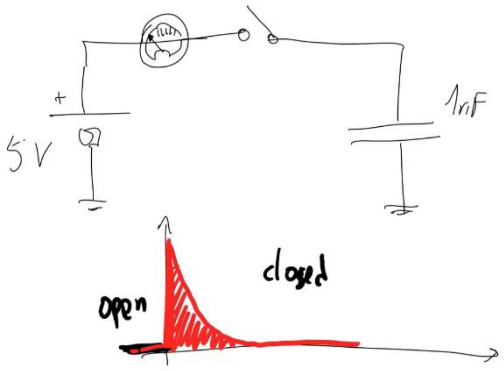
Metal layer technology is more precise than the mould and wire technologies.

CAPACITORS

Typical values ranges from 1pF to 1000uF. Asian coding uses figures, the American one uses figures and the "." notation or nothing.

A capacitor is a wire and we want the two wires to see very well the other side, so we use to metal plates. On the metal plates charges will accumulate. So the capacitive effect requires to wires to be very close one to the other. **C = epsilon* S/d**, where S is the surface and d the distance.
Moreover: **C = Q/V** (charge accumulates equally on the two plates).

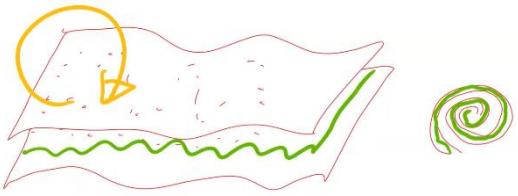
Charge of the electron: **q = 1.6*10^-19 C**.



The red area of the curve that is the integral of i(t) in dt over time is the charge Q that is C*V. In DC regime there is no more current in a capacitor. In AC, the capacitor keeps changing the charges on its plates because the direction of the current keeps changing, so the capacitor will behave as a shortcircuit. These are two asymptotic behaviours → every signal below fpole for the capacitor can be assumed as DC and the C open, while if higher the capacitor can be consider a shortcircuit.

With the capacitor, in DC (in DC there is no current in the capacitor) we can set the value of C we want and the tolerance, but not the power anymore, because the capacitor will never dissipate power, since it is a reactive component (current and voltage are shifted by 90 degrees and there is no current through the capacitor). **We have to specify the maximum voltage the capacitor can withstand**. Again, the safety rule requires 150% more than the desired value.

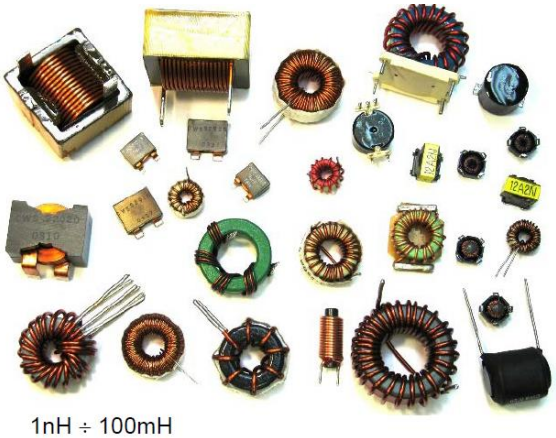
To create capacitances with a very big value, we place two metallic plates one close to the other, we create oxide on the plates and we place the two plates one on the top of each other and we roll the plate. The two foils are isolated one from the other. Thus we achieved a very huge S and very small d, so the C is huge.



The issue is that the two thin foils are so close to each other that it can withstand small voltages, otherwise the disruptive discharge could make a hole in the capacitor creating a shortcircuit.

NB: if a units of measure is not specified for a capacitor, it means pF.

INDUCTORS



Values: 1nH ÷ 100mH

Tolerances: 20% ÷ 5%

Current ratings: 1mA ÷ 1A

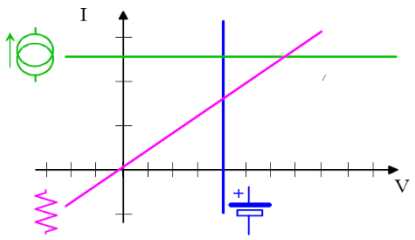
An inductance is a wire with 0 resistance that is in a solenoidal configuration across a nucleus. When the current flows in an inductor, we generate a magnetic field that constantly changes direction because of the AC current across the coil. This field self-induces a counter force on the windings itself.
The inductor behaves as the opposite of the capacitor. In DC it behaves as a shortcircuit, whereas in AC it becomes as an open circuit.
As soon as we try to push current in the inductor, it reacts trying to reduce it. The signal at which we discriminate if the signal is AC or DC depends on the fpole → **tau = L/R**.

In an inductor we have to specify the current value (there is no voltage across it in DC).

NB: in general, the current-voltage dependency is not mediated always through a resistor, but there are nonlinear behaviours.

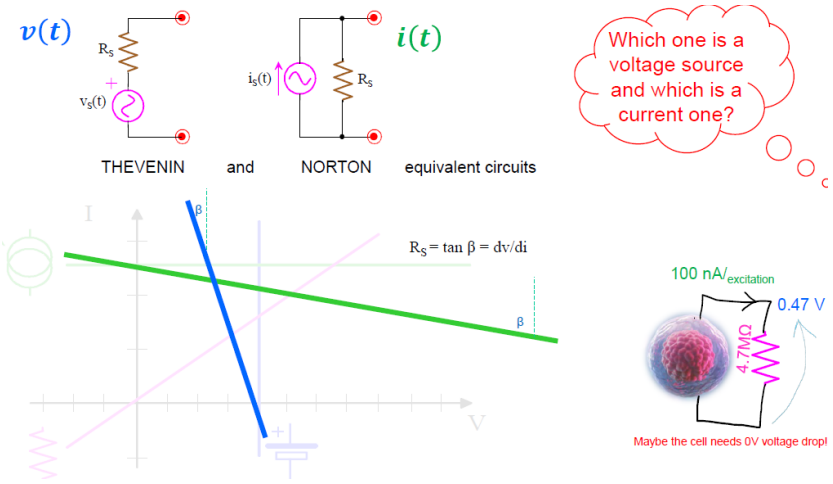
IDEAL VOLTAGE AND CURRENT SOURCES

A current source provides a constant current whatever the voltage across it, while a voltage source provides a constant voltage no matter the current flowing through it.



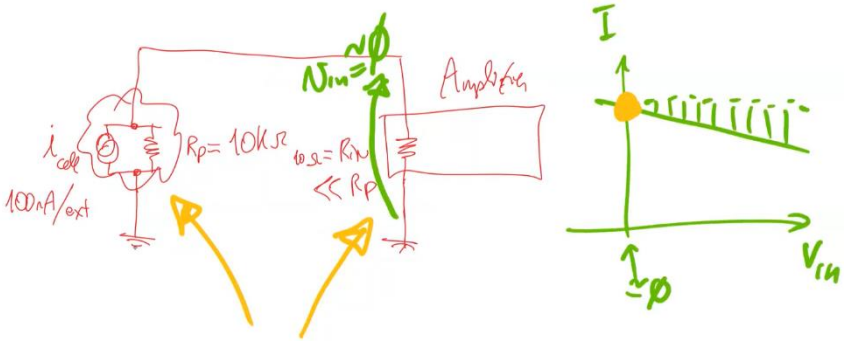
REAL VOLTAGE AND CURRENT SOURCES

The line is not exactly straight. If there is a voltage drop across the current source, the current through the current source starts decreasing → the real generators are modelled with Thevenin and Norton equivalents that take into account the voltage drop.



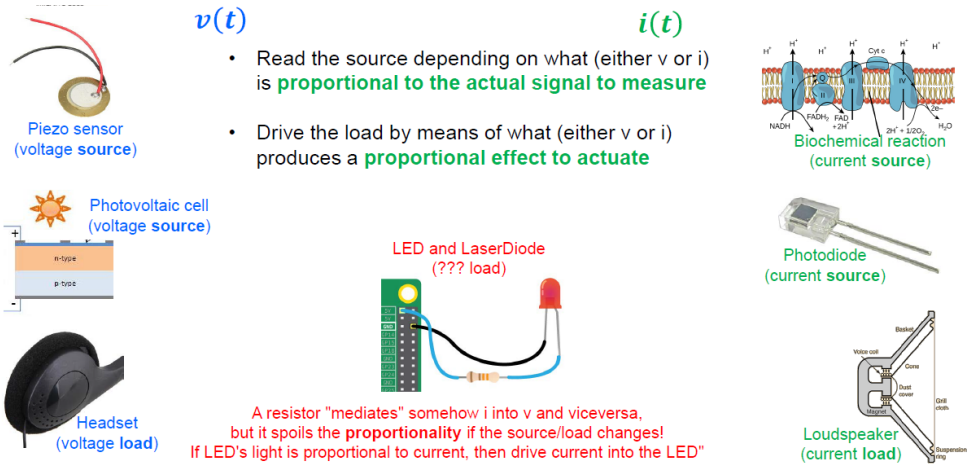
When a source is a voltage source or a current source. We cannot say it from the plot, but it is depending on the physical mechanism that underlines it behaviour. For instance, if we have a biological cell that generates 100 nA per excitation, if we want to measure the signal, we should measure the current. But sometimes the instrument can measure the voltage and not the current → we connect the cell to a resistor. However, this approach with the resistor is completely wrong, because we are converting a current into a voltage and we are reading the voltage but this voltage could change the intrinsic characteristics of the cell. The voltage drop across the resistor is reflected back on the cell, so the excitation current will be different (decreased).

So if the source is a current source better to read just the current and try to maximize the current, without developing any voltage across the current source. The amplifier to read the current should have an input impedance much smaller than the parasitic resistance of the real current generator. If so, there will be almost no voltage drop across the input impedance of the amplifier so the Vin across the cell itself will be almost zero, we don't loose any current due to the parasitic resistance.

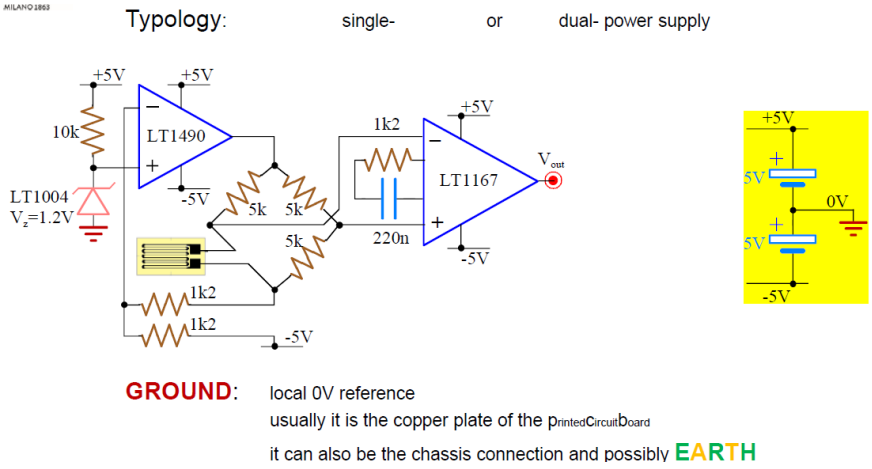


The slope of the I-V curve is the equivalent resistor of the Norton equivalent circuit. The ideal current generator has an equivalent resistance in the Norton model that is infinite. So for a current source the best amplifier is the one with a zero input impedance, so that there is no voltage generated at its input.

As for the voltage source, the amplifier to amplify the signal should have an Rin much greater wrt the Rs of the Thevenin equivalent of my voltage source. In this way there is no current flowing in the circuit, so the voltage source will provide the voltage signal without any drop.



POWER SUPPLY

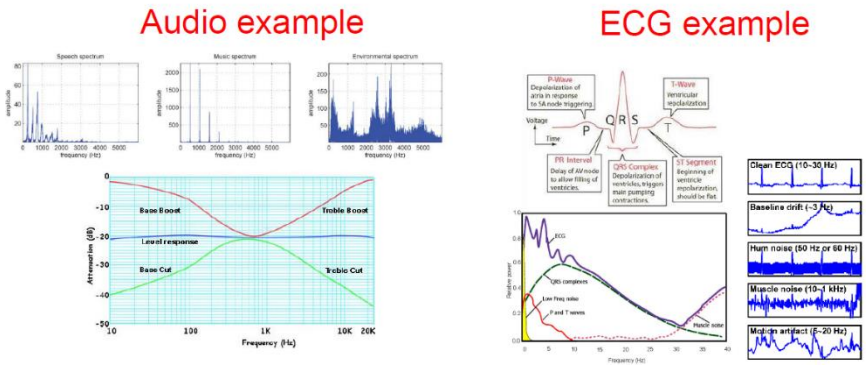


We have operational amplifiers that need power supply. The power supply refers to how we bias the circuit to have the circuit that operates. In the case of the image we have a dual power supply, meaning that there are two batteries that provide +-5V. To do so, we connect them in series and we consider the intermediate node as ground; if so one node will be 5V above ground, the other will be 5V below it. So ground is not necessarily 0V, but the signal we want.
Of course, the KVL must always be respected, so the current that flows out the most positive power supply must be equal to the one flowing in the most negative power supply. Then the difference must flow through ground.

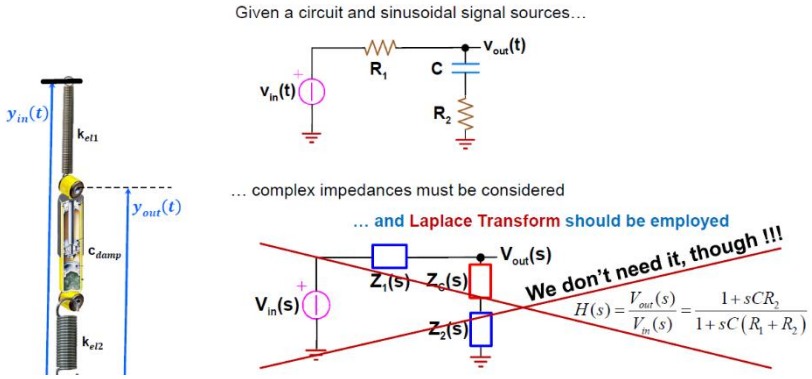
FREQUENCY RESPONSE

We have to know which are the frequencies of the signals we are working with and we want to amplify.

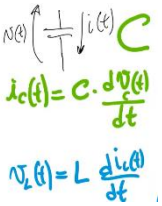
Not just to Low-Pass, Band-Pass, High-Pass, Notch, but also for Integrators, Derivators…



Frequency analysis



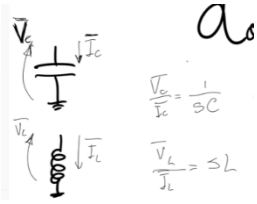
Let's start from a very simple schematic where we have a voltage source, two resistors and a capacitor. We want the output voltage in the time domain and frequency domain. We have seen the asymptotic behaviours of an inductor and a capacitor in DC and AC, but in middle frequencies, the current that flows in a capacitor has a specific value (the same for the voltage across the inductor).



If at DC the variation of voltage over the capacitor is 0, then the current is zero. Similarly, in DC if the variation of the current in the inductor is 0, then the voltage across it is 0. Conversely, if the variation of voltage becomes very high across the capacitor, the current becomes infinite. The same and opposite reasoning for the inductor.

Capacitor and inductor can also be seen in the frequency domain.

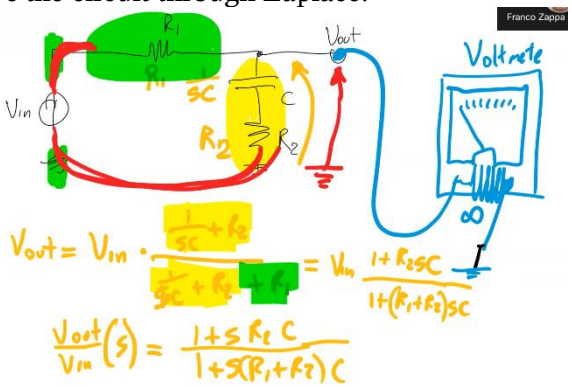
However, if the system has not a constant and finite frequency, but it varies the specific of the signal, with DC parts, AC and transients, we need to introduce the Laplace transform. It is another way to describe our system in the frequency domain (s = alpha + j*2*pi*f).

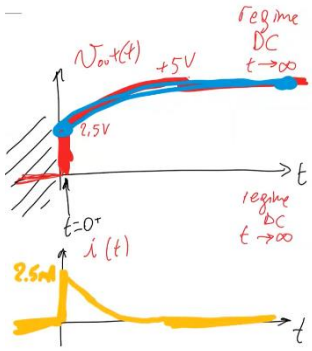


However, we are no longer using the Laplace transform because it's too complex, so we try to solve the circuit in another way.

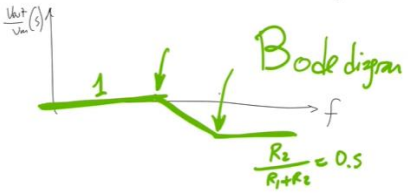
Coming back to the previous circuit, we want to compute Vout given Vin. When we buy a voltmeter, we know that it has two wires and an infinite input impedance. The voltmeter reads 0 current, so it doesn't steal any current.

Let's solve the circuit through Laplace.

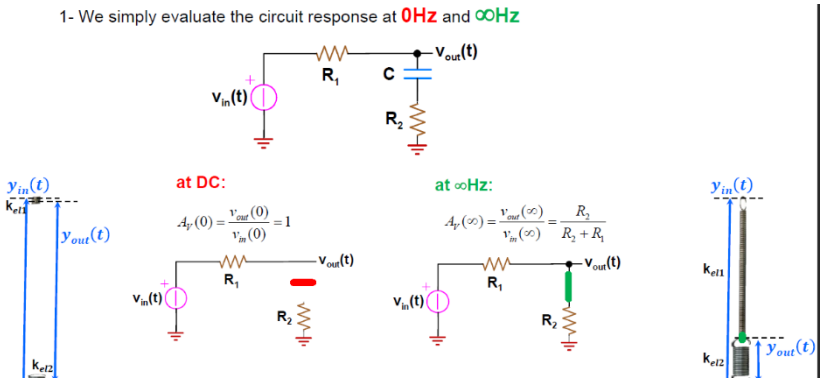




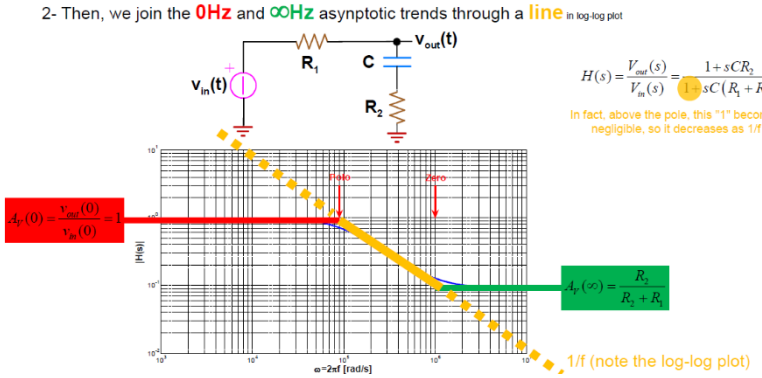
Now we want to plot is how this gain varies with frequency f.



However, we can forget about the Laplace transform and study the frequency response of our circuit by using an asymptotic approach. Let's consider the circuit at DC (constant voltage applied at the input) or in AC (AC voltage at the input) and let's compute the gain.



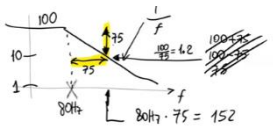
Now we have to compute at which frequency the gain drops, that is when the reactive components start to display their behaviour. To do so, we have to compute the pole of the circuit, that is 1/tau, where tau is the time constant of the circuit, the speed at which the circuit relaxes down to a quite state. Imagine charging the capacitor at whatever voltage we wish. Then we switch off the voltage source



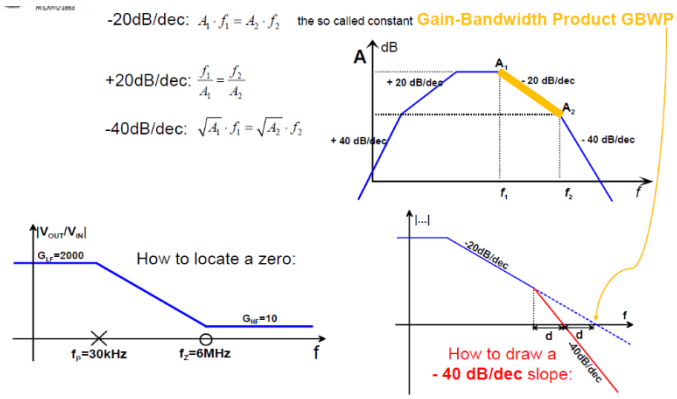
(shourcircuited), the capacitor will spontaneously discharge because a current will flow through R1 and R2 and charges will recombine. This occurs at a speed that depends on the time constant tau of the circuit. In our case, the total R for the tau is R1 + R2 → tau = C*(R1+R2). The transient is typically over after 3 to 4 tau.

As for the frequency of the zero, we can compute it with the GBWP.

In a Log-Log plot, the trend of increase or decrease is increasing or decreasing because as we can see in the Laplace transform, positive slope are associated to s (or f), negative to 1/s (or 1/f). This means that if we have a decrease of 75 in gain, we will have a shift in frequency of 75.



Wrt the right Laplace equation we commit some errors where we have the poles and zeros, and they are of 3dB. In fact, **the pole is the frequency at which the gain drops by a factor 3dB**.



The inverse of seconds is Hz or rad/s? Given the period, f = 1/T is in Hz, but if we calculate f = 1/tau, in this case we are in rad/s. So one over period is actually in Hz, while since tau is not a period, 1/tau is in rad/s. 1/(2*pi*tau) is actually in Hz instead.

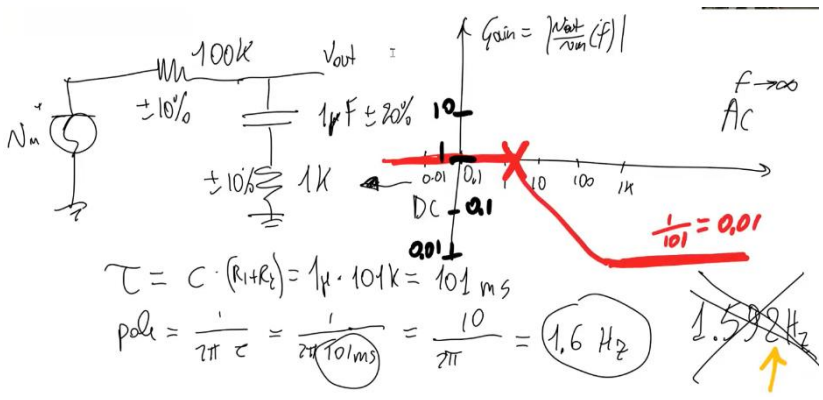
NB: the decibel (dB) is a ratio of something, e.g. Vout/Vin and applied like this.

$dB = \text{ratio} = \frac{V_{out}}{V_{in}} = Gain$

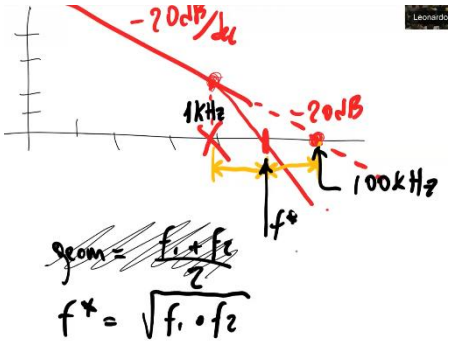
$dB = 20 \cdot \log_{10} \frac{V_{out}}{V_{in}}$

The -20dB/dec slope can also be called a 45° slope because the slope has an angle of 45°.

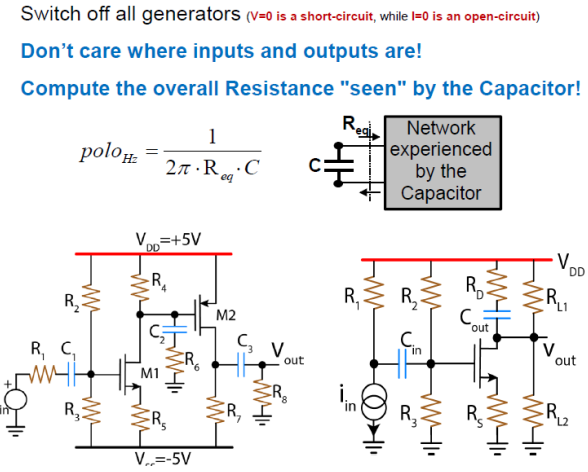
Example



How to compute the frequency when we have a -40dB/dec slope:



How to compute poles



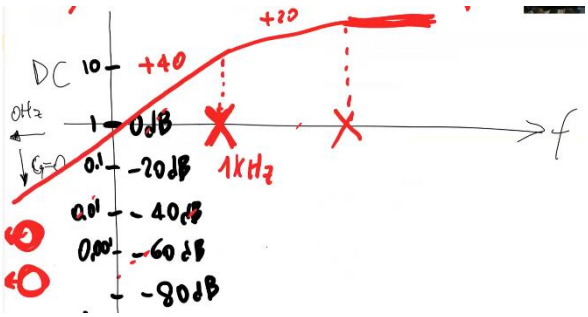
In order to compute the pole, we choose one capacitor, we remove it and look at the equivalent capacitor it sees. The hope is that one capacitor doesn't meet another capacitor because in this case we don't know how the other capacitor is affecting the one of which I'm computing the equivalent resistance. Hence so far we assume capacitors that are not interactive, like in the circuit on the left.

First of all, I have to check if the capacitor is actually introducing a pole in our circuit, and this happens if in going from the input to the output we pass through a capacitor. Moreover, if there are capacitors (e.g. 3 of them) connected in parallel, they don't give rise to three poles, but just one pole given by the parallel of the three.

To compute the pole of C1, we remove it and see the resistance we see from it. When we compute the pole all the voltage/current sources must be off, either power supply or input signal. Hence the equivalent impedance seen from C1 is R1 + R2||R3 (in an ideal MOSFET the impedances seen into the gate and into the drain are infinite, whereas into the source I see 1/gm).
As for the capacitor C2, the Req = R4 + R6. For C3, Req = R8 + R7 (if the transistors are ideal).
In a MOSFET, if there is no degeneration at the source, we see r0 from the drain, but if there is degeneration, hence a source resistance Rs, we don't see r0 but more (not r0 + Rs, because I have a feedback action in the transistor itself, it's much higher).

Hence the circuit has three poles, but we are not done, because there may be also some zeros. Moreover, we also miss the value of the DC gain.
As for the DC gain, all capacitors are open. Hence the DC gain is clearly 0, because the C1 is open; for sure C1 introduces also a 0 in the origin, but then also a pole later. In a Bode plot, gain equal to 0 means -inf. Before the pole of C1 the signal is not passing, then after the pole of C1 the signal passes.

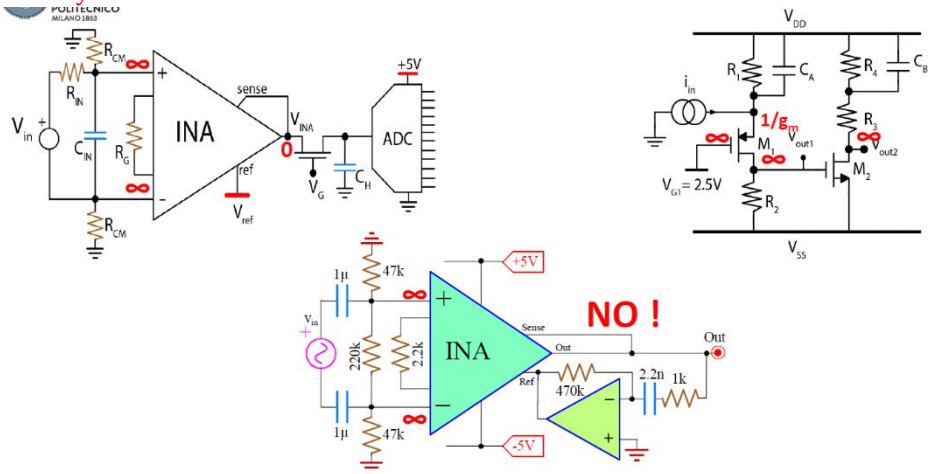
Furthermore, also C2 introduces a zero in the origin, by the same reasoning. So we will have a +40 dB/dec slope at 0 Hz.



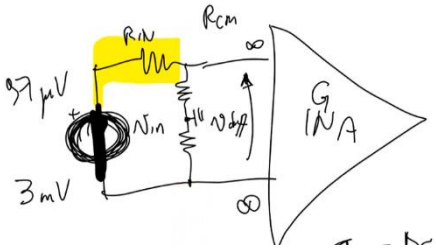
Above the two poles, the C1 and C3 will be shorted for the signal, so we can compute the flat gain of the circuit when C1 and C3 are shorted. Firstly we compute the transfer from the input to the gate of M1 through the resistive network, then gain of M1, from gate to drain is -Rdrain/Rsource, so -R4/(1/gm + R5) because C2 is open in this regime. Same reasoning for M2, (R7||R8)/(1/gm).
Then we have t o compute the pole of C2. If it is after the poles of C1 and C3 I'm ok and happy, but the problem is if the pole is lower than the other two. If so, the assumption we made on having the capacitors open or closed must be redone with the correct involved capacitors.

If by chance the pole associated to C2 is in the middle between C1 and C3, we cannot compute the gain going through the network because C3 would be open. But I know the value the gain could have with C2 after C3, I know the position of the first pole and the distance between C1 and C2, and with GBWP technique I can compute the gain.

Poles in Electronic systems



For the upper left circuit the pole of Cin is due to Rin||2*Rcm. Cin is introduced so that at DC it is open, and Vdiff at the input of the INA is not Vin, because we have the resistance Rin. We can rewrite the network as aside.



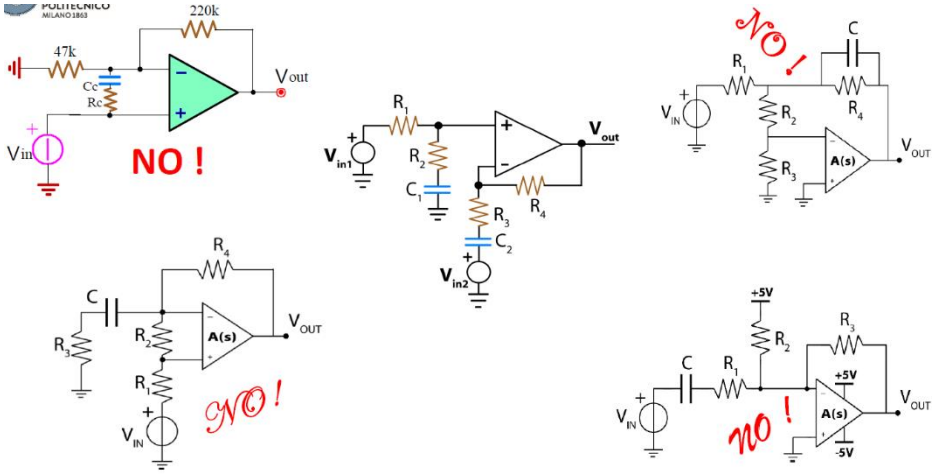
Vdiff = Vin*(2*Rcm/(Rin + 2*Rcm))

So we have a flat DC gain. At infinite frequency, the Cin is shorted and Vdiff is 0 and hence the output is 0 → gain is 0 and it dies after the pole associated to Cin. Hence the Cin is introduced to perform a LP filtering action.

If we now move to the bottom circuit, the pole is computed considering the two capacitors in series because when the generator is off they are connected, so we have just one pole. Req = 220k || 2*47k.

As for the gain, at DC the two capacitors are open and the gain is 0, so we have a zero in the origin of the Bode plot. So the capacitors act to introduce a HP filtering.

Feedback action on poles

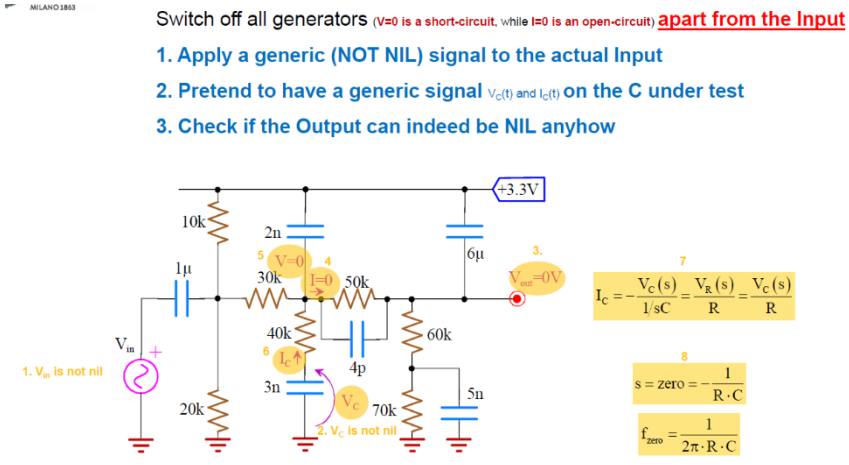


Feedback is useful but when it works something happens. Feedback means that if we have an amplifier and we apply a signal in input, the amplifier amplifies the difference, so the output moves but then, thanks to the network, also the other pin moves and hence the input difference decreases and the output decreases.
For instance, in the central circuit the – terminal is virtual ground, it doesn't change its voltage. If so, the pole associated to C2 is just associated to the R3 resistance.

For the upper right circuit, again the – terminal is virtual ground, the + terminal can move (but we decided to keep it fixed at ground). If so, R3 is in between VG and GND, so no current in it, so no current in R2, so no voltage drop across R2, so no voltage drop across R1 and hence C sees just R4.

For the bottom right circuit for the pole we have C*R1 (same concept on VG).

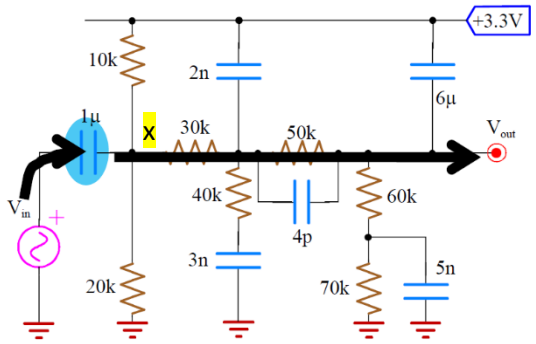
How to compute zeros



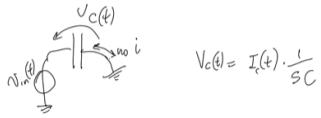
The capacitor introduces a zero if the output stays at zero when we have an input signal. If the output is 0, there is no current through the 6u capacitor and so no current in the parallel configuration at its left, so no voltage drop across 50k, so no current, so no current in 2n and 40k and so on. If I want to know if the capacitor introduces a zero I simply pretend that Vin moves, voltage across the capacitor moves and the output stays at 0, and node x stays at 0. *Is it possible that Vin moves, Vc moves and node x does not and there is no current?*

A Capacitor along the signal path: a zero at the origin

$f_{zero} = 0$



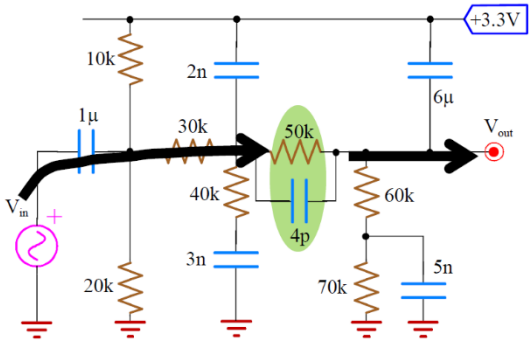
Let's draw the simplified circuit and the equation of the capacitor.



It's actually possible to have Ic equal to zero whatever Vin is if s = 0. This that we found is the frequency of the zero (in this specific case it is a zero in the origin) → whenever we find a capacitor connected to the path of the signal we have a zero at f = 0.

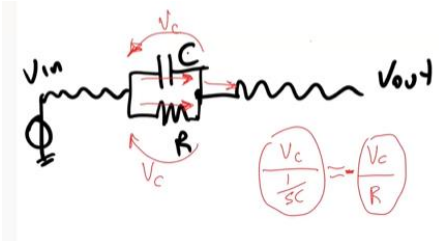
The other possibility is the following.

An RC-shunt along the signal path: a zero at finite frequency



$$f_{zero} = \frac{1}{2\pi \cdot R \cdot C}$$

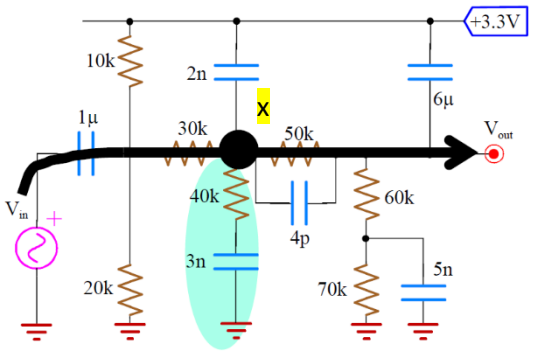
If we have a RC network along the path of the signal, we have a zero at infinite frequency. Again, if the output is 0, we have zero current flowing in the **RC shunt** from the output but at the same time a voltage across it because we have the input that is on. So we have actually a current from the input side of the RC shunt. This current recirculates in the RC shunt. *At which frequency the voltage across the capacitor causes a current through the resistor that is equal to the one through the capacitor?*
At infinite frequency, s = -1/RC. This is the zero of the network at a finite frequency.



Of course the pole of the capacitor 4p is not due to RC as the zero, but the resistance in the pole is the equivalent capacitor.

Then we have also another possibility, due to a **RC series** along the path of the signal.

An **RC-series hanging at a node** along the signal path: a finite zero

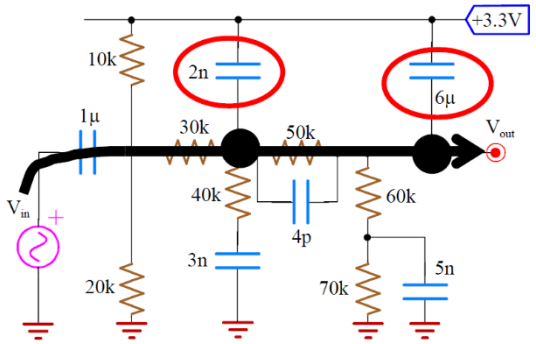


$$f_{zero} = \frac{1}{2\pi \cdot R \cdot C}$$

If we have an RC to ground we have a zero at finite frequency. If the output is 0V, there should be no current up to node x, because we have no voltage drop across the 50k resistance. *May we have a voltage across the 3n capacitor but 0V at node x?*
Yes, it happens if the current flowing through the resistance in series gives a voltage that is the opposite with respect to the one of the capacitor. Again, this happens if s = -1/RC. Again, related pole is different.

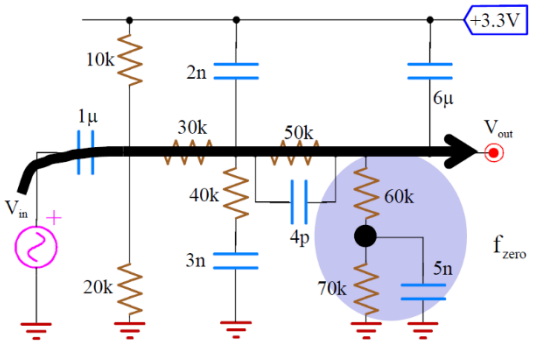
There may be, in the path of the signal, also capacitors attached to ground. These capacitor are not giving any zero, because it is not possible for the output to be zero if there is a voltage across the capacitor if we have just a capacitor to ground.

A Capacitor hanging at a node along the signal path: NO ZERO



If we have a capacitor that is on another signal path, it introduces a zero at finite frequency.

A **Capacitor hanging at a node off the signal path**: a finite zero



$$f_{zero} = \frac{1}{2\pi \cdot (60k\,//\,70k) \cdot C}$$

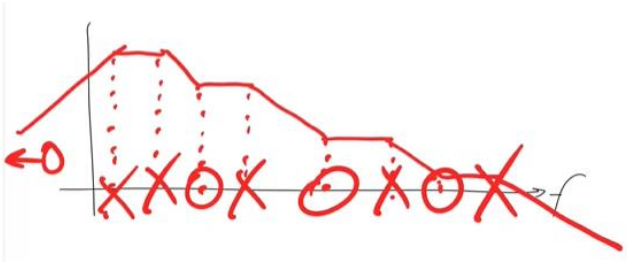
To summarize, a capacitor along the path of the signal introduces a zero in the origin. If the capacitor is from one node to ground we have no zero; if we have a RC to ground or a RC shunt we have a finite zero, and also if we have more complex networks attached to the signal path. In the network in example we have for sure 4 zeros, of which one is in the origin.

As for the poles, capacitors here are interacting. So we cannot compute the pole assuming that all the other capacitors are not existing when computing the pole associated to one specific capacitor. In this

case, I cannot know the precise frequency of the pole, I'm just giving an estimation based on reasonings on resistances and having capacitors shorted or not.
The number of poles that we have is exactly the number of capacitors we have? At the most, they are 6, and we don't have any parallelisms or series of capacitors.

But are there any dependent capacitors (so do we have a loop formed just by capacitors)?
Yes, the 2n, 4p and 6u capacitors form this loop. So the other capacitances interact one with the other but they are independent one from the other, exception given for these three.
So we have just 5 poles, because the 3 dependent capacitors behave like two poles, because only 5 out of the 6 capacitors are independent (but still interacting).

The final Bode diagram will be for sure something like below.

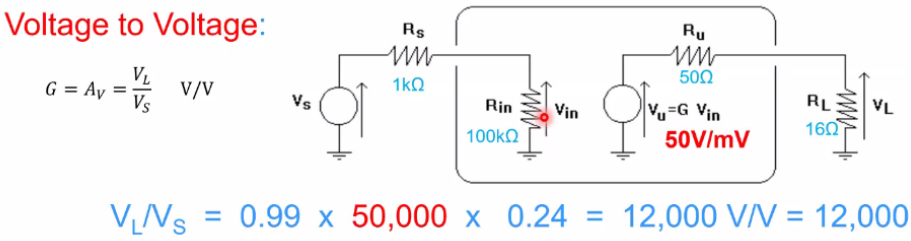


NEGATIVE FEEDBACK

Negative feedback improves the gain, the stability, the input impedance, output resistance and so on. We need to choose the proper amplifier depending on the input that we have.

VOLTAGE AMPLIFIER

In the example of the image, the input source provides a voltage that is proportional to the input signal. E.g. in input we might have a piezoelectric microphone that produces a signal proportional to the voice. We model the input source with the input resistance Rs that considers the voltage drop across the microphone if we force the microphone to provide a current.



Unfortunately there are **voltage drops**:

$$V_{in} = V_S \cdot \frac{R_{in}}{R_{in} + R_S} \qquad V_L = V_u \cdot \frac{R_L}{R_L + R_u}$$

Hence a good **Voltage Amplifier** must have:

$R_{in} >> R_S$ very high input impedance $\qquad$ $R_u << R_L$ very low output impedance

So we have our microphone, we model it with a voltage source v_s(t) and we have across the mic, if we change the current of the mic, a voltage drop that decreases with an internal source resistor. The smaller the resistor, the more vertical the I-V characteristic, so more ideal.
We need to understand how to properly match the source with the load, and the load is modelled with a resistor.

In the first amplifier we want to amplify the voltage and provide the voltage to the load. It Is the load that provide a physical reaction proportional to the voltage and not the current. To model the amplifier, we use two pins, the input and ground and the output and ground at the output.
We can model the amplification as a voltage dependent voltage source, that is a voltage source whose voltage v_u is proportional to the input voltage, so the output is G*Vin.

Then we can model also the real behaviour of the output voltage source, so we place a resistance in series to it and also we introduce an input impedance. Let's see what happens due to the presence of an input impedance that, unfortunately, is not too high and an output impedance that, unfortunately, is not nihil. If the output impedance of the amplifier is not negligible with respect to the load impedance, if we compute Vload/Vs, the gain is not the gain of the amplifier but it is reduced due to the input and output impedances of the amplifier. From Vs to Vin we have the first path, then the amplification and then the final partition due to the output stage, due to Ru and Rl (load). If we compute Vin/Vs, it is like in the image (voltage partition).

In the final stage we have a great loss of signal because the output impedance of the loud speaker (load) is much lower than the output one of the amplifier. Hence, if current needs to flow, the current will cause

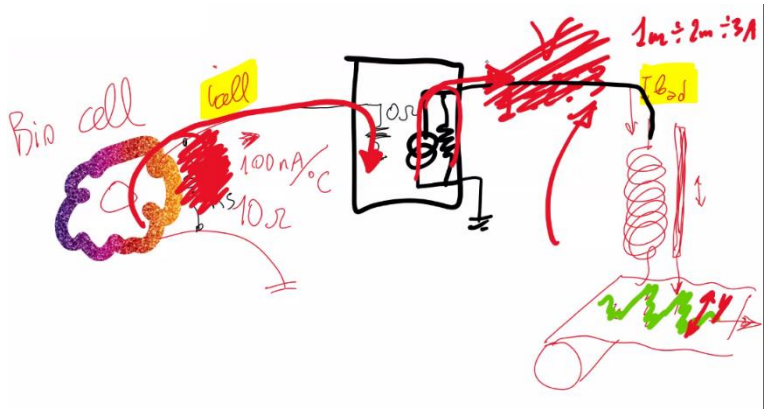
a voltage drop inside the opamp that will be higher on the load, simply because the output impedance of the amplifier is higher.

We realized that we loose signal not at the input, but in the output partition. To maximize the voltage gain, we should improve matching, to have an input impedance higher of the amplifier with respect to the voltage source input impedance, and the best way to provide a voltage is to have a load impedance that is much higher than the output impedance of the amplifier.

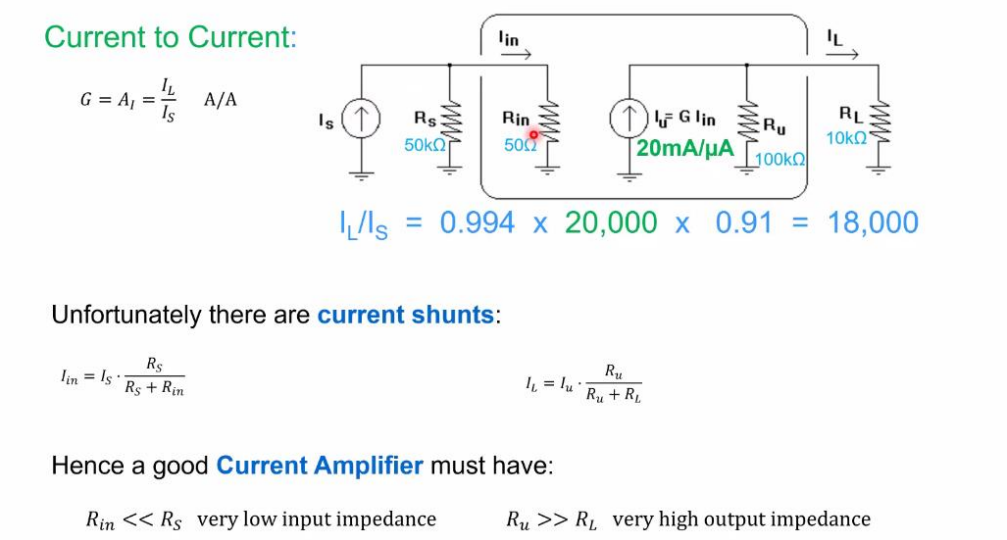
Since we know the amplifier that we bought, we have to choose a mic and a load coherently with respect to the already bought amplifier.

CURRENT AMPLIFIER

We need to amplify a current and provide a current. The source and load are both currents; the amplifier requires an input impedance as smaller as possible and an output impedance as bigger as possible, so that the current in input flows in the amplifier and not in the input impedance path related to the source (in the image a solenoidal load to be driven with a current, not a voltage).



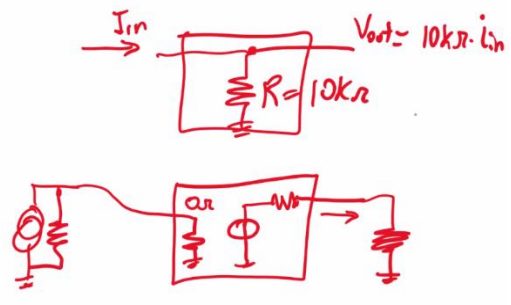
We have a current source Is, the amplifier with input impedance Rin, the output impedance Ru and the load Rl. Again, the source current Is that enters in the amplifier is a current partition; if Rin is small, almost all the current will flow in Rin, and a small part in Rs (input impedance of the source).



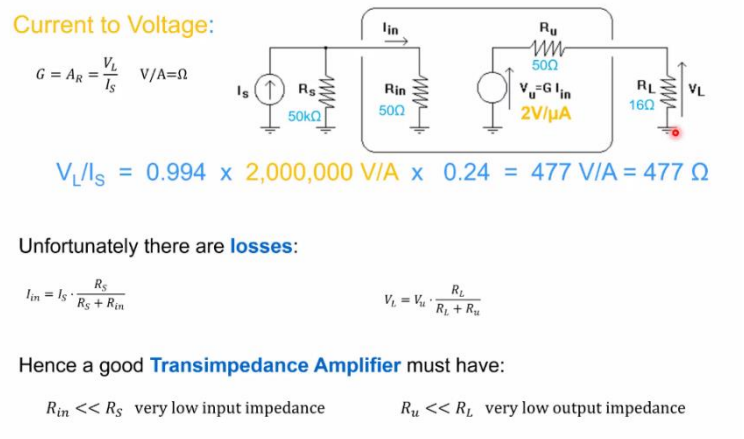
As for the output, the current that flows in the output speaker is still given by the partition. If Ru (output impedance of the amplifier) is bigger than Rl, it is good. It is exactly the opposite with respect to the voltage amplifier.

TRANSIMPEDANCE AMPLIFIER

Now for instance we have an input current and an output voltage to drive a load. Now the gain is Vout/iin, so the amplification has a size of Ohms → transimpedance amplifier because the gain is no more a pure number but an impedance, but it is not a resistor. It is not as the upper image, but as the lower image. The output should provide as much current as the output load wants, not just like as if it was a resistor.

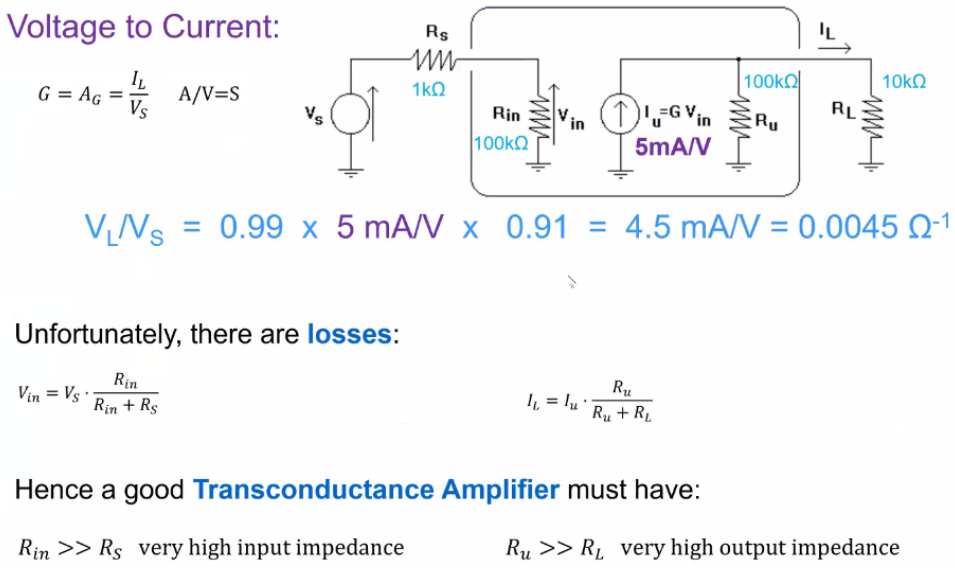


The input is modelled as a current source, so Rin much smaller than Rs. In output Rload must be much bigger than Ru. If not, we could have a bad matching in the output stage as in the image.



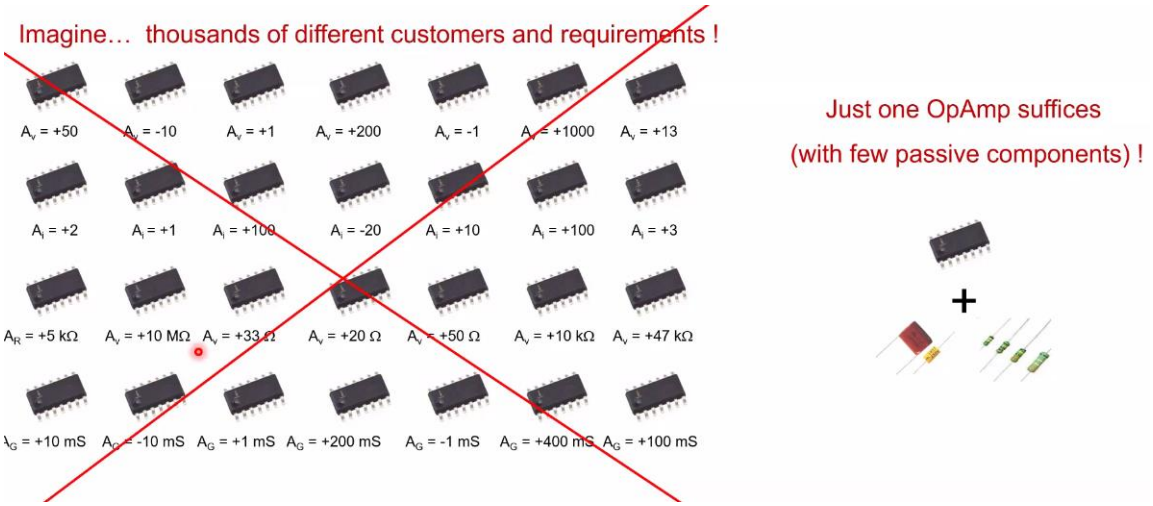
TRANSCONDUCTANCE AMPLIFIER

It is exactly the opposite of the previous one.



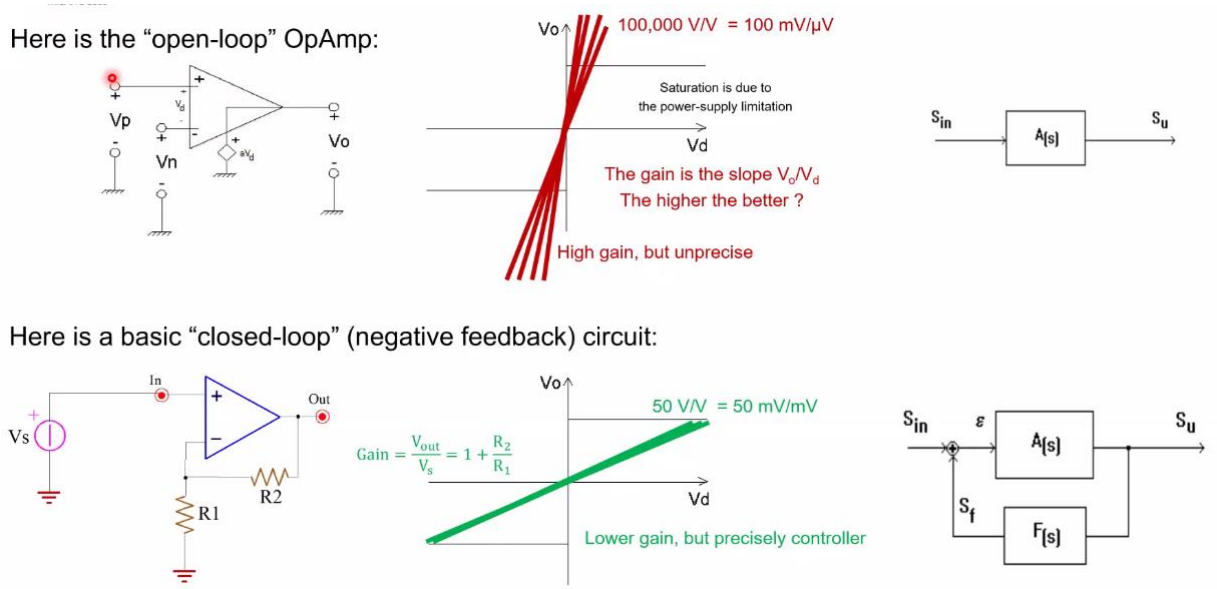
The gain is now in siemens (S).

So to achieve these results we can simply use an amplifier matched with the correct passive components and feedbacks.



OPAMP WITH NEGATIVE FEEDBACK

The opamp is not used in open loop configuration (OL), because always feedback will be given to it.



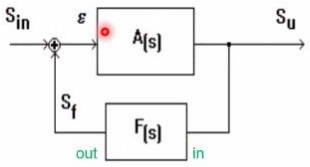
The idea is to buy one amplifier and use one input of it to give the input, and the output to provide the output. The gain will decrease if we use components in feedback to the amplifier with respect to the OL configuration. I like having a strong amplifier, but it could also amplify the noise and if it gets old or heats up, its gain will change a lot so it won't be used as before. Hence let's not use the opamps in OL configurations, but let's introduce feedbacks. If we subtract the feedback signal to the input signal, we get an output error signal that will be amplified in a lower way than before. So we can set the gain by choosing the attenuating network Sf.
The gain will be much lower but more stable and controllable and selectable with the passive components.

Mathematics

Let's model the amplifier with the Laplace transform of its voltage gain, A(s). Let's use an attenuating network F(s) and something known at the input, that is the sum of the two signals.

If A(s) is a non-inverting amplifier (positive gain) and the feedback network is not inverting, at the input I need to subtract, otherwise the overall feedback would be positive. Or I can put a + and having an inverting amplification.

Closed-loop OpAmp with feedback:



$$S_u = A(s) \cdot \varepsilon$$
$$\varepsilon = S_{in} - S_f$$
$$S_f = S_u \cdot F(s)$$

Real (closed-loop) gain:

$$G(s) = \frac{S_u}{S_{in}} = \frac{A(s)}{1 - A(s) \cdot F(s)} = \frac{A(s)}{1 - G_{loop}} = -\frac{1}{F(s)} \cdot \frac{1}{1 - \frac{1}{G_{loop}}} \rightarrow -\frac{1}{F(s)}$$

Real gain — Correction factor — Ideal gain

Loop gain G_{loop}: measures the **feedback effectiveness**, the higher the better !

if G_{loop} is very high, ideally infinite, the "error" ε gets nil

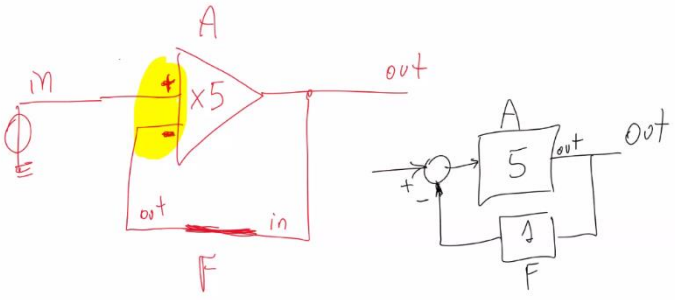
the negative feedback forces S_f to mimic S_{in} !

Indeed, the **ideal gain** does not depend on the OpAmp!

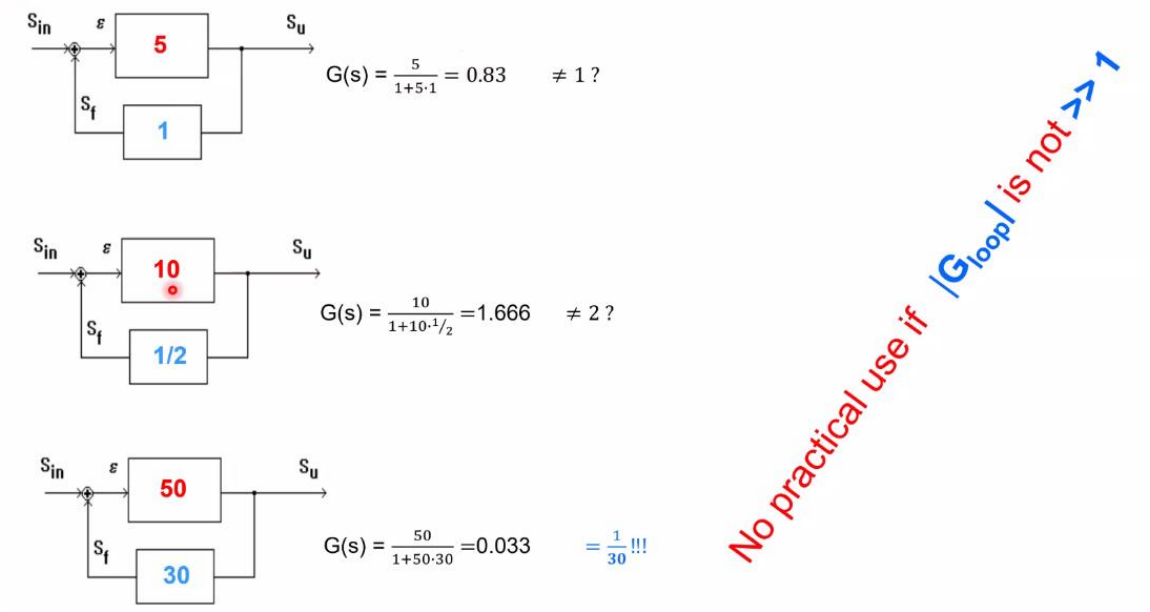
Su (u = uscita) is epsilon multiplied by A(s) OL gain of the amplifier (error in the slide, epsilon = Sin + Sf with Sf negative).
The overall actual gain of the amplifier is x. Since I'm interested in negative feedbacks, either in input I have a minus, so the – Gloop will have a Gloop negative to have a negative feedback so that 1 – Gloop is a big positive number.
Thus, the gain drastically reduces. Let's divide now the numerator and denumerator by Gloop, that is A(s)F(s). -1/F(s) is the ideal gain. 1/Gloop must be smaller with respect to 1, so that if Gloop is strong the gain drops and the real gain is just 1/F(s), that is determined by the feedback and it won't be affected by temperature, age and so forth and so on like A(s) → the feedback must be precise to have a precise gain.

In the end the real gain is not the gain of the amplifier, but reduced by the negative feedback. Gloop represent the effectiveness of the feedback, if it is very high it is perfect, because it means that the output is bigger and in feedback there is the inverse of Sin, so that if the input is 2V, it returns in Sf = 2V, so 2-2 = 0 so epsilon = 0, hence the feedbacks works in a way that the output will be moved by the feedback so that we will have the same signal as the input, still maintaining the output bigger than the input, thanks to the feedback.

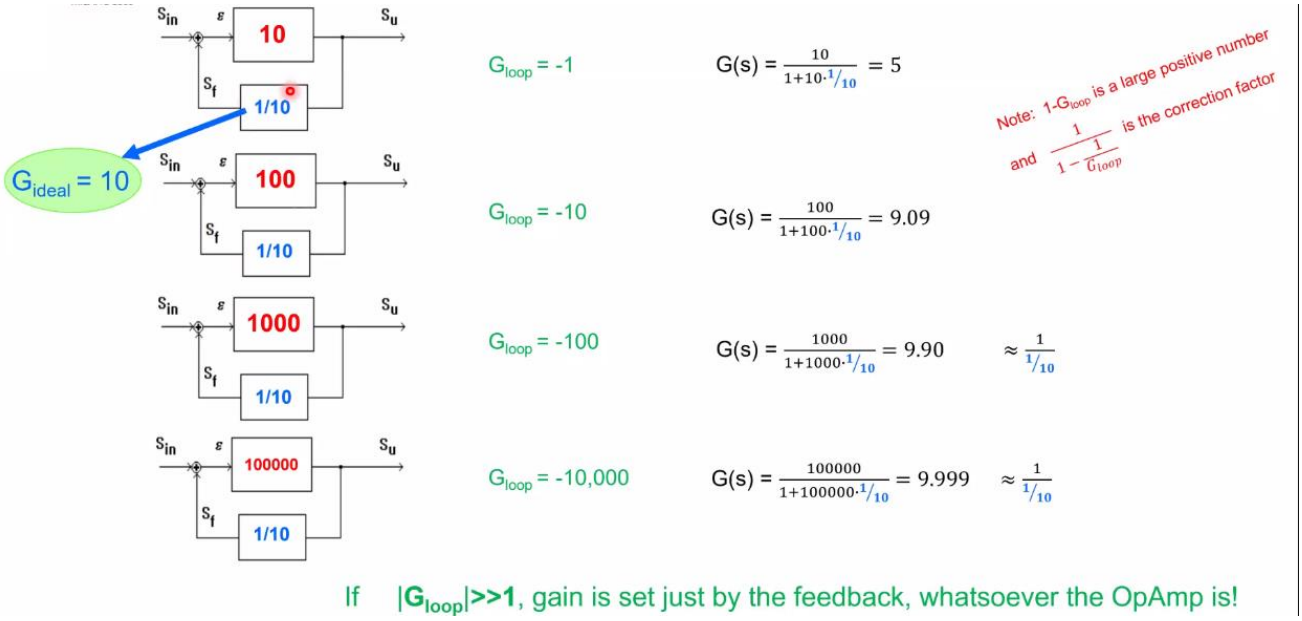
Example



The final gain out/in = A/(1 – AF) = 0.83. Other examples in the next images.



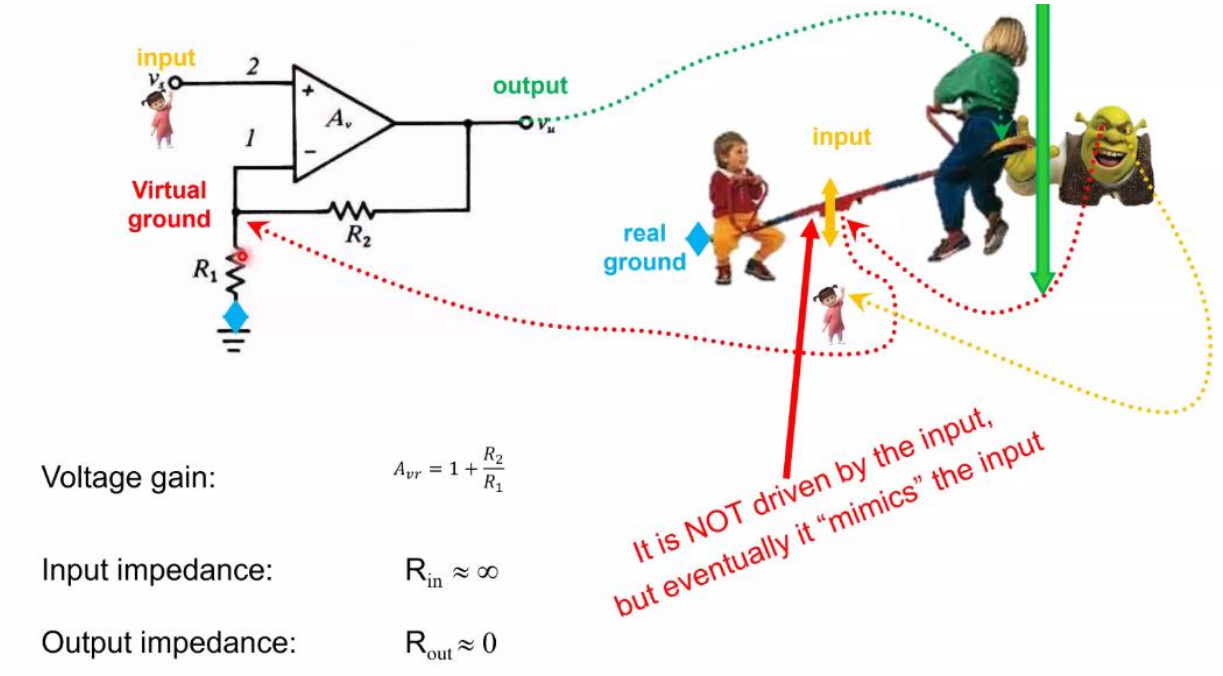
Gloop was poor in case 1, 2, but strong in case 3. Hence 1/Gloop is small with respect to 1 and the gain is the inverse of the feedback network. If Gloop is much higher than 1, it makes sense to use the feedback, otherwise we should not use the feedback.



Now we have case 1, the gain is 5, and so on. Again, if Gloop is poor, it is not an advantage, while if we increase it, the gain arrives close to 10, that is the ideal gain.

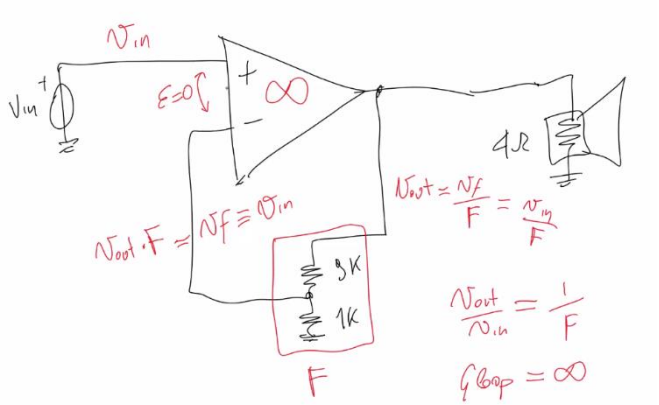
So for negative feedbacks it is important to have whatever amplifier but a strong Gloop, which means much higher than 1 (e.g. a hundred).

NON-INVERTING AMPLIFIER



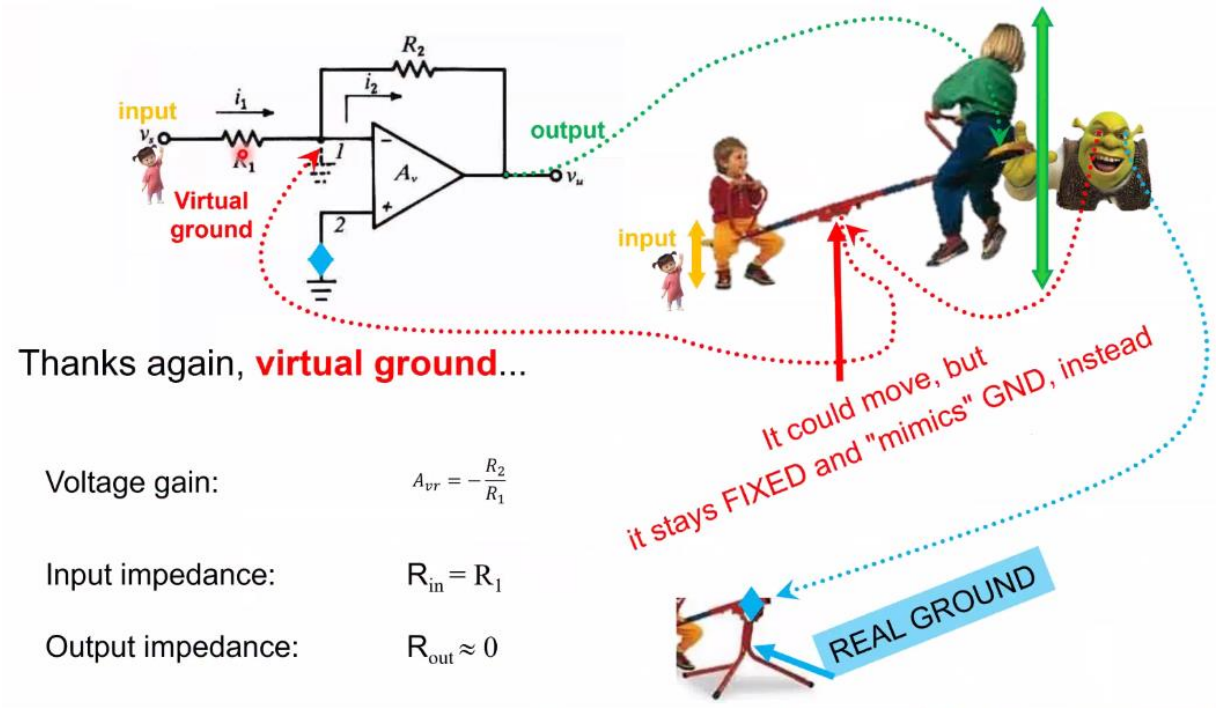
We have a very small voltage source whose series resistance is very high (very bad source) and we want to 'move a load'. The amplifier needed is the one in the middle. The input source drives the input of the amplifier, then we use two resistors. The opamp doesn't require current from the input source. The feedback will be ok when the signal between input and feedback is 0. The current in feedback flows in R1 and R2.

To have a 0 error in input (epsilon), we need Vin = Vfeedback, which means that Vout = Vf/F = Vin/F if Gloop is infinite.

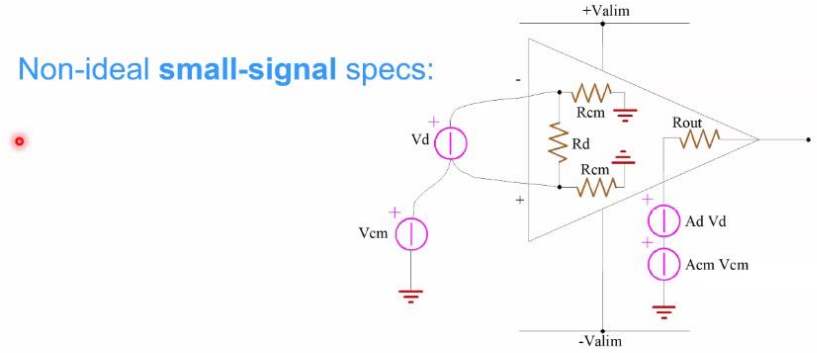


Even if Gloop is not infinite, it should still be bigger than 1 to have a good feedback → we need resistances with small tolerance, 1% and not 10%. To the load, in this configuration, we can provide whatever current we want, because the opamp does it, either voltage or current. We don't need current in input because the opamp can provide whatever current we need without the need of it in input.

INVERTING CONFIGURATION

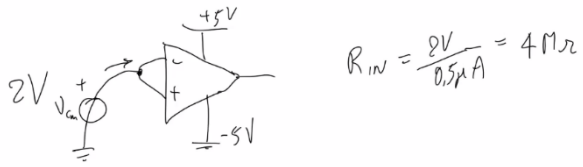


REAL ELECTRICAL PERFORMANCES

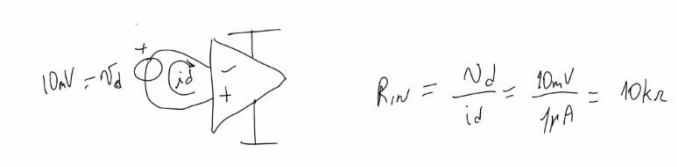


· Differential Gain (A_d)	> 100'000	> 100dB
· Common-Mode Gain (A_{cm})	< 10	< 20dB
· Bandwidth (BW)	10Hz ÷ 1kHz	
· Differential input impedance (R_d)	> 100kΩ	
· Common-Mode input impedance (R_{cm})	> 1MΩ	
· Output impedance (R_o)	< 4kΩ	
· Temperature drifts	some %/°C	

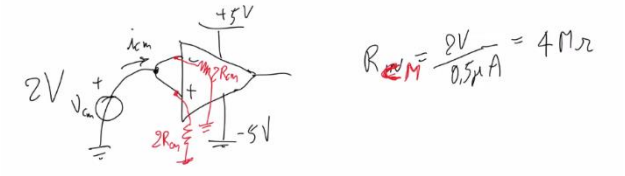
Ideally, I want an output voltage of the amplifier proportional to the voltage input difference through the differential gain Ad. The I want a Rout close to zero, to be ideal. We are considering an amplifier that will be biased with e.g. +- 5V, and it is made of different transistors. Depending on the number of components inside, the gain can vary. I hope to have a 0 output impedance, but it will be 1/gm of the output transistor. I want the input impedance to be infinite, but it won't be in the real case.
To measure if the input impedance is infinite or not, we can connecte + and – pins together and then apply a Vcm signal. If a current goes in, it means that the input impedance is not infinite.



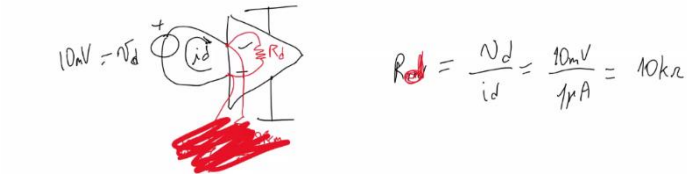
I can also take the opamp and apply a differential voltage. If I measure a differential current Id, then as in the image below.



To model Rin in the two cases, in the case of common mode I should model it as if there were two resistance to ground, each equal to 2 times Rcm (they are in parallel).

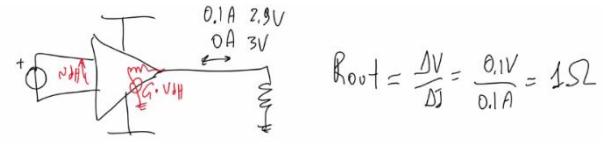


In the case of the differential source, I cannot use the same two resistors as above, but I should use a differential impedance Rd.



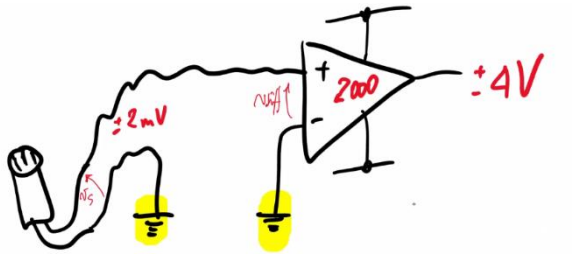
Output impedance

Finally, we have the opamp, we drive it with the load, the output will be e.g. 3V with 0A, 2.9V with 0.1A and hence we can know the value of Rout.



This is the reason why I should model in the real opamps 3 resistors in input and one in output.
Even if the differential signal Vd is 0, the output might be not 0 due to the common mode. So we need to consider a contribution of it to the output that is AcmVcm. I love to have a very high Ad but a very small Acm.

In fact, different grounds can be not at the same potentials due to parasitism, we might have a voltage drop that can be assumed to be Vcm.

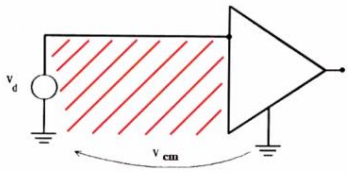


An ideal opamp should have Ad = inf and Acm = 0. I also would like an infinite bandwidth, but typically will be in the range of tens of kHz. Moreover, all these parameters will vary with temperature with a certain percentage.

Hence even if we buy a bad opamp with not infinite Rin, bandwidth and so on, the feedback will improve all the parameters.

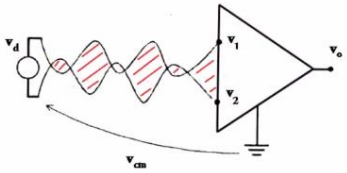
SINGLE ENDED VS DIFFERENTIAL AMPLIFIER

Single-ended input:



- GND bouncing undistinguishable from signal
- ElectroMagnetic disturbance coupled through loop

Differential input:



$$v_o = A_d \cdot v_d + A_{cm} \cdot v_{cm}$$

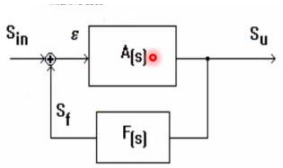
- let's amplify just Differential signal
- let's reject Common Mode disturbance

$$CMRR = \frac{A_d}{A_{cm}} \approx 80\text{–}100dB$$

I can twist the cable so that if there is any electromagnetic disturbance is not coupled with the loop of the wires.
Moreover, in case of common mode signals (e.g. between grounds), the Vcm is applied to both inputs, so this is the reason why residual Vcm will be rejected if Acm is 0, but this is not true in the real case. Acm is more or less 10 for instance, and Vd is usually negligible because small and needs to be amplified and Ad is high because of the amplifier, whereas the Vcm is very high but with a very small Acm. There is the risk that common mode signal overruns the useful signal → we have a strong disturbance.

It is introduced the CMRR that should be the highest possible.

Since we will never use the opamp in OL configuration, the feedbacks helps.



- feedback (G_{loop}) lowers the gain, but adds MANY advantages
- gain is independent of OpAmp and depends just on feedback
- stronger feedback, smaller correction factor !

Closed-loop gain:

$$G(s) = \frac{S_u}{S_{in}} = \frac{A(s)}{1 - A(s)\cdot F(s)} = \frac{A(s)}{1 - G_{loop}} = -\frac{1}{F(s)} \cdot \frac{1}{1 - 1/G_{loop}}$$

For example: G_{loop} = -10 correction factor = 0.9 G_{loop} = -100 correction factor = 0.99

Note: $1-G_{loop}$ is a large positive number and $\frac{1}{1-\frac{1}{G_{loop}}}$ is the correction factor between ideal and real gains

- and … many more further advantages …

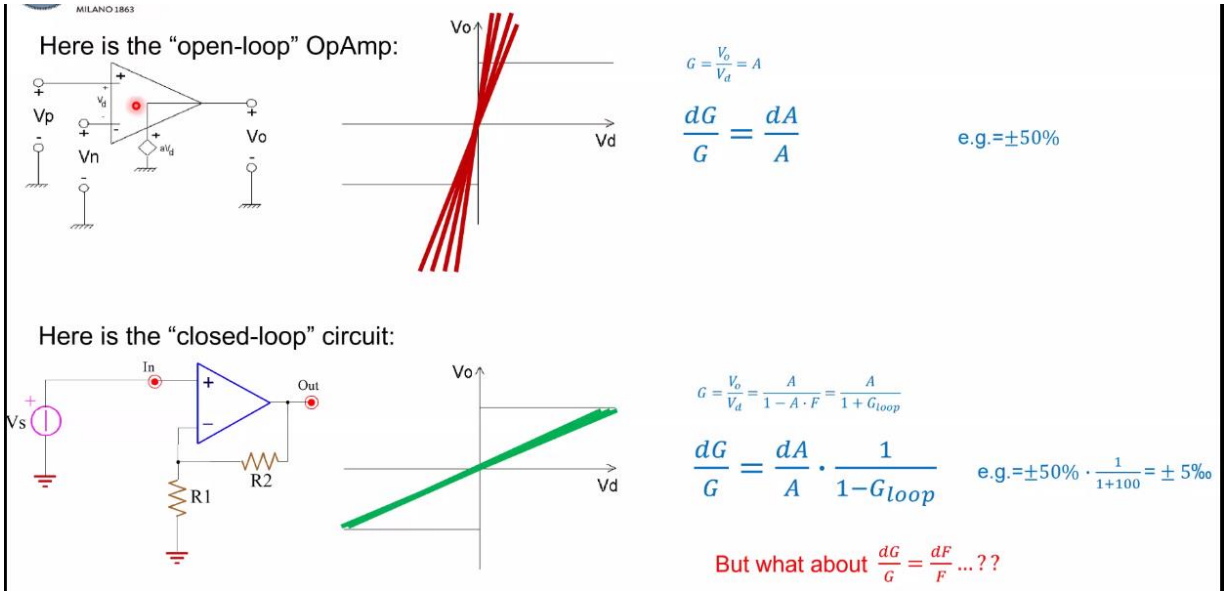
larger bandwidth	(excellent speed)
lower output impedance	(excellent voltage source)
either ∞ input impedance	(excellent voltage reader)
or *0* input impedance	(excellent current reader)

The gain is reduced by Gloop, however it is more stable.

IMPROVEMENTS GIVEN BY THE FEEDBACK

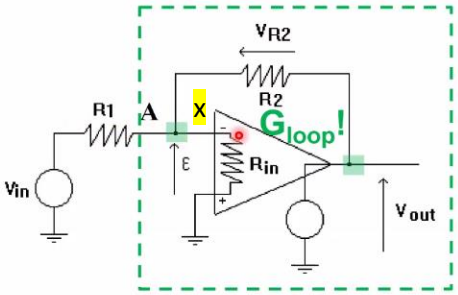
Tolerances

In the upper image the gain is high, but if A changes of 15% (e.g. due to age), the gain varies of 15%.



If we consider the same variation but with a feedback added, now it is better. The G (gain) is computed as always and now the variation of gain is dA/A divided by 1 – Gloop, hence if Gloop is strong, even if A varies, it is mitigated by Gloop → external effects won't affect the performances.
Now if I compute dG/G as a function of dF/F, that is due to tolerances of components, if F changes by 10%, G changes by 10% as well → the feedback reduces the dependencies on the tolerances of the opamp (forward dependencies), but not on the tolerances of the feedback components. We can buy a cheap opamp but we need very precise resistors, because they set the gain.

Input impedance



Input impedance:

$$R_{in} = R_1 + R_A$$

$$R_A = \frac{\varepsilon}{i_1} = \frac{\varepsilon}{i_{in} + i_2} = \frac{\varepsilon}{\frac{\varepsilon}{R_{in}} + \frac{\varepsilon - A(s)\cdot\varepsilon}{R_2}} = \frac{R_{in}\cdot R_2}{R_2 + R_{in} - A(s)\cdot R_{in}} =$$

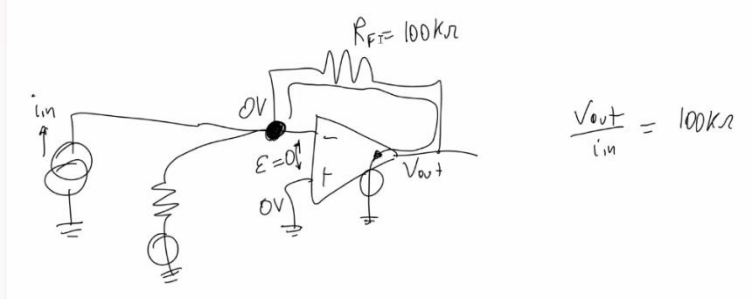
$$= \frac{R_{in}\cdot R_2}{R_{in} + R_2} \cdot \frac{1}{1 - A(s)\cdot\frac{R_{in}}{R_{in} + R_2}} = \frac{R_{in} \| R_2}{1 - A(s)\cdot\frac{R_{in}}{R_{in} + R_2}}$$

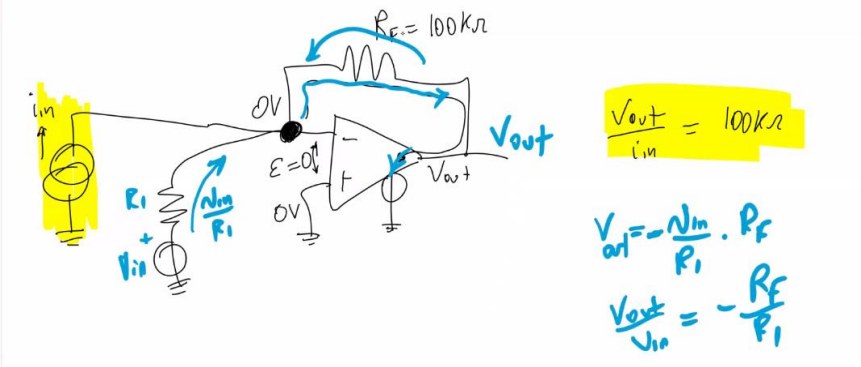
The feedback is negative because if I have a positive increase, + terminal will go positive, hence if the output stays 0, the output will start to go negative so that node x is free to move up and down, and it is controlled by the feedback.

Every time we see a node in a feedback loop, the loop will try to maintain that node without voltage variations.

Again, since the + terminal is grounded, the amplifier will provide an output so that node x will be always at 0V. So Vin will se R1 and then VG, not actually a ground, because driven by the amplifier, and the amplifier drinks the current that the input is providing.

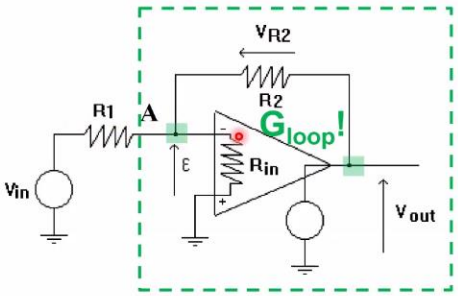
Every time we have an opamp and the output of the opamp goes back to the input, if we enter with a current source or a real voltage source, that node will stay at 0 because thanks to feedback the error epsilon between terminals + and – is kept to 0, the injected current flows through the output.





The last blue equation relies on the presence of a virtual ground. So even if we introduce a resistor between + and – terminals, it won't have any effect because it will have no voltage across it and no current across it and its impedance won't be sensed. The current that goes in the virtual ground flows in the resistor between + and – terminals and goes to ground where it ends.

Let's now compute the input impedance measured by the source.



Input impedance:

$$R_{in} = R_1 + R_A$$

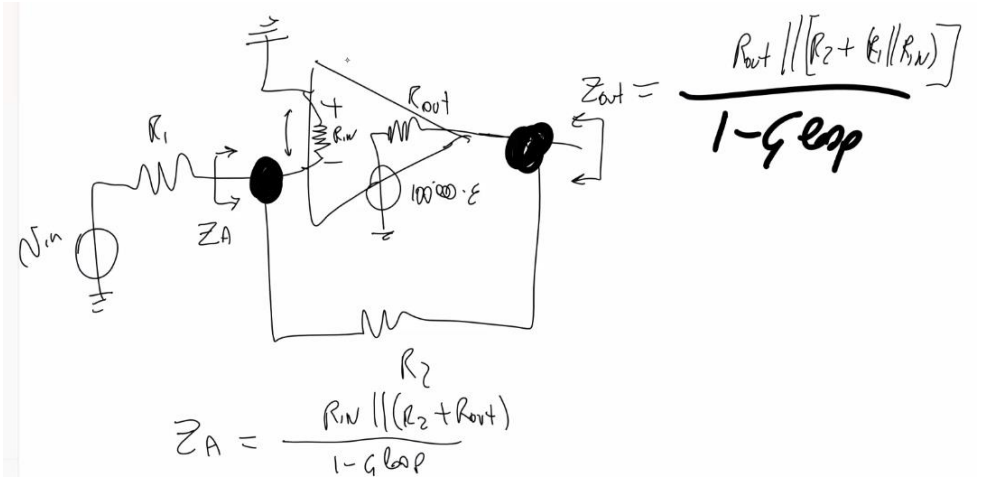
$$R_A = \frac{\varepsilon}{i_1} = \frac{\varepsilon}{i_{in} + i_2} = \frac{\varepsilon}{\frac{\varepsilon}{R_{in}} + \frac{\varepsilon - A(s)\cdot\varepsilon}{R_2}} = \frac{R_{in}\cdot R_2}{R_2 + R_{in} - A(s)\cdot R_{in}} =$$

$$= \frac{R_{in}\cdot R_2}{R_{in}+R_2}\cdot\frac{1}{1 - A(s)\cdot\frac{R_{in}}{R_{in}+R_2}} = \frac{R_{in}\,\|\,R_2}{1 - A(s)\cdot\frac{R_{in}}{R_{in}+R_2}}$$

The input impedance considers everything off. We have R1 and two paths then, one through Rin and ground, the other thru R2 and ground. So it seems that Rin and R2 are in parallel. But this is not true because R2 is not connected to ground, since as soon as A moves, the output node moves.

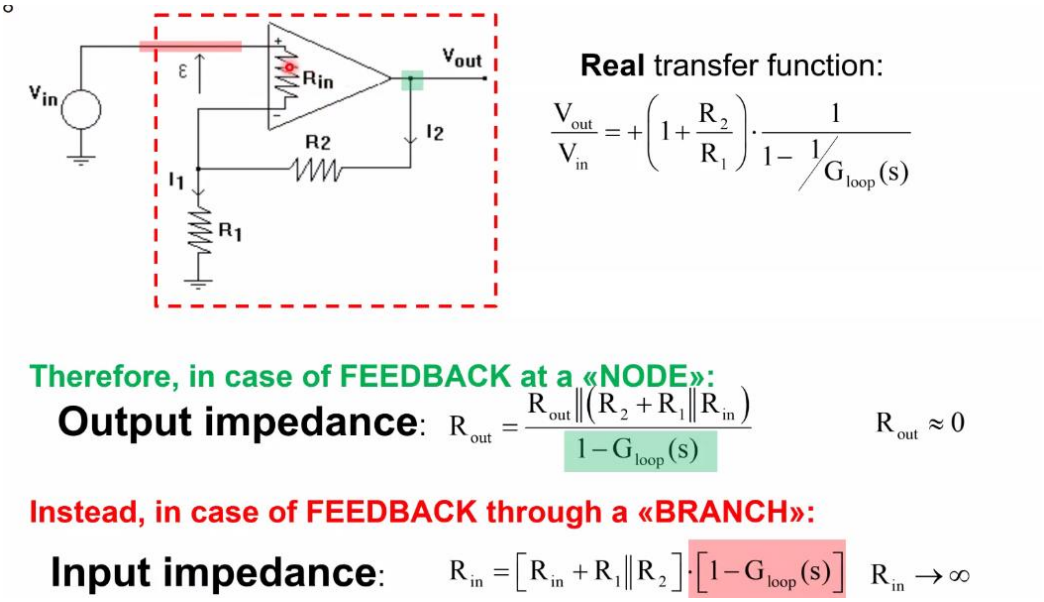
How much Ra is? It is epsilon divided by current i1 that flows, and that is iin + i2. Then the computations are like in the image.
In the end, the impedance into A, that is Ra is actually Rin || R2, divided by a correcting factor that is 1 – Gloop. Every time we have such a circuit, the input impedance seen by Vin is: firstly there is R1, then I need to realize that there is a negative feedback and if Gloop is infinite, the input resistance is Rin||R2. If it is not infinite, we need a correcting factor.

The same reasoning can be done for the output impedance. In the ideal case I see 0 Ohm, but in the case of a real opamp, we have a Rout in parallel to a parallel of resistances.



If Gloop is strong, the output resistance is almost null.

IMPEDANCES THROUGH BRANCHES



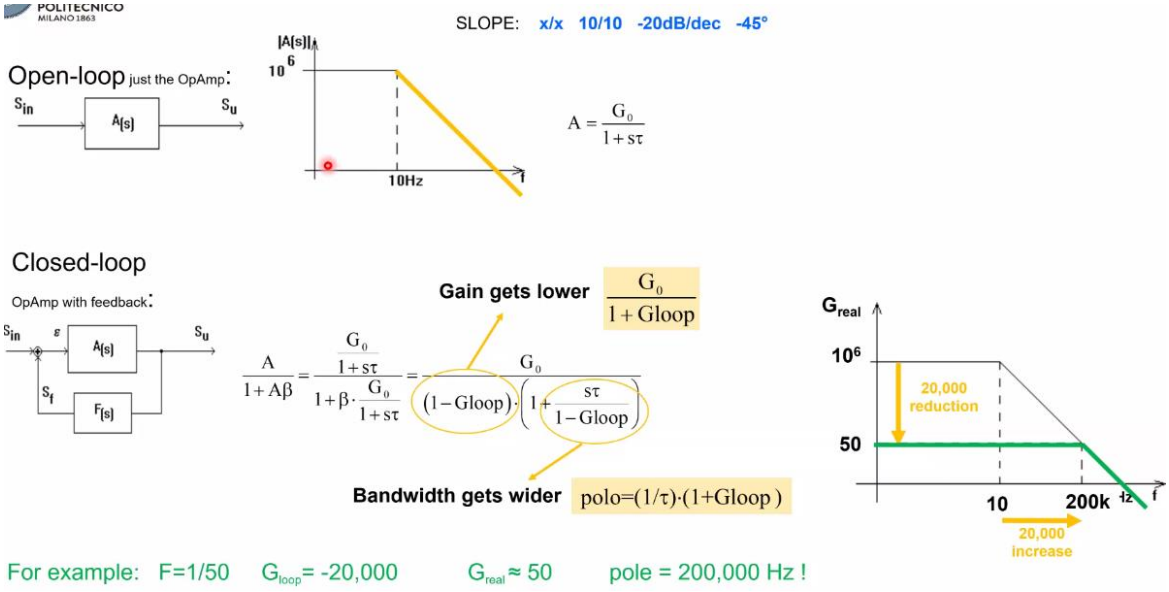
Let's consider the case of a real Rin (not infinite). If it is e.g. 1kOhm and I apply 1V, the opamp drinks mA of current from the source. But the loop gain Gloop will improve the impedance in input. In every

node of the loop the impedance is divided by (1 – Gloop), either the output node or the – terminal of the depicted circuit.
Thanks to the Gloop the input impedance is the dummy one magnified by Gloop, because in the loop branches we have a magnification.

NB: when computing output and input impedances the generators are off, and so also the voltage generator inside the amplifier, hence we have like ground at the output of the amplifier.

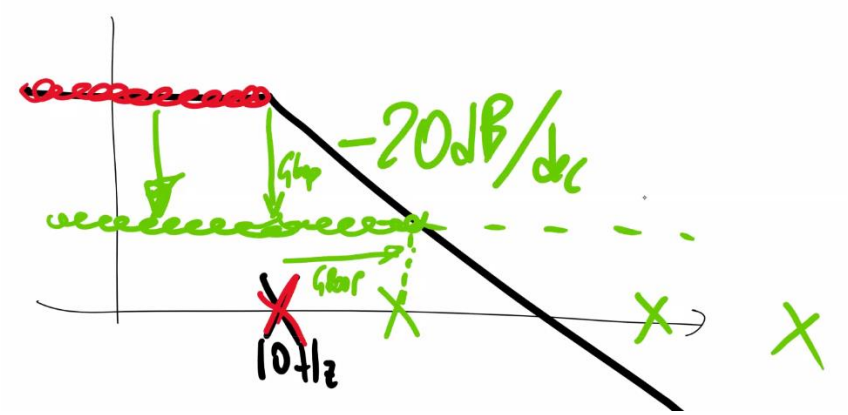
We end up with an equation that, rearranged, it showing an input impedance that is the easy one (Rin + R1||R2) multiplied by Gloop in the case of input on the positive terminal.
Again, the opamp gain is not needed to be infinite, it is sufficient to have good components.

FEEDBACK EFFECT ON BANDWIDTH



To have a very high gain, we need a lot of transistors in the opamp, but the higher is also the number of parasitic capacitances and hence the decrease in bandwidth. The bigger the amplification, the smaller the bandwidth in OL configuration of the amplifier.

If we use the opamp in OL, we are stupid, because we have a very huge amplification but with a very small bandwidth. To increase the bandwidth, we can shift the position of the pole with the feedback, because the tau is reduced with Gloop. In this way we kill the gain but we fasten the amplifier.
Hence Gloop kills the gain by a factor 1 – Gloop but also pushes the pole of a factor 1 – Gloop. Where the DC gain touches A(s), there we have the new pole.

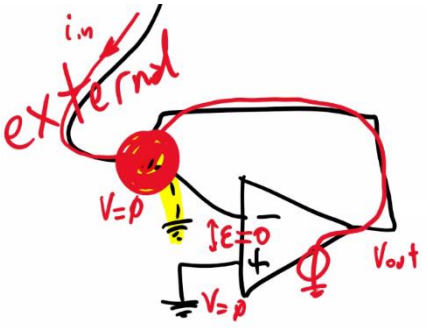


OPAMP STAGES

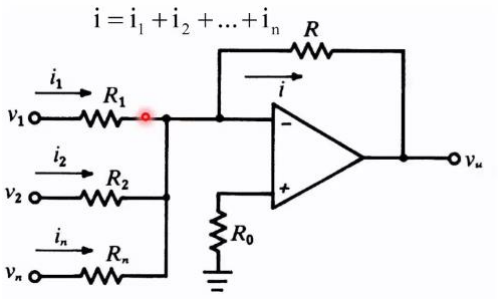
VOLTAGE AND CURRENT ADDER

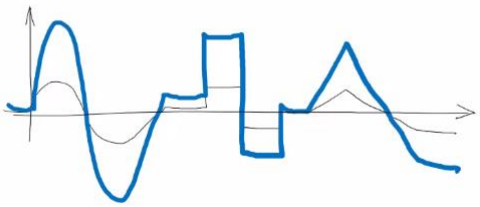
We want to add several current sources. Let's consider different current sources, maybe with different performances (related to the parasitic resistance, with the smaller the R the worst the current source). In order to add the current, let's have the current flow in a load resistor. It is not good, because Rload has to be small, and if the parasitic R is low, it is not good, Rload should be small for all the related resistance. If a voltage develops over Rload, it generates a current that steals current from the sources.
There is another possibility to have a resistor whose value is much smaller than all the others and it is the one with R = 0. In this case, iout is indeed the sum of all the currents. However, we sum the currents but we have no signal to amplify.

Hence we need a ground, but not a ground, something that behaves like it without being a ground. The best solution is to use virtual grounds. Hence we use an amplifier with one input grounded and then the other node will be at virtual ground. We introduce a feedback circuit to force one terminal to virtual ground. Thus the current flows in feedback to ground.



If we place a resistor in feedback, we will have a voltage drop on it, having hence Vout that is not 0, but Rf*iin (transimpedance amplifier).





Voltage gain: $v_u = -\left(\frac{R}{R_1}v_1 + \frac{R}{R_2}v_2 + ... + \frac{R}{R_n}v_n\right)$

Input impedance: $Z_{in\,i} = R_i$

Output impedance: $Z_{out} \approx 0$

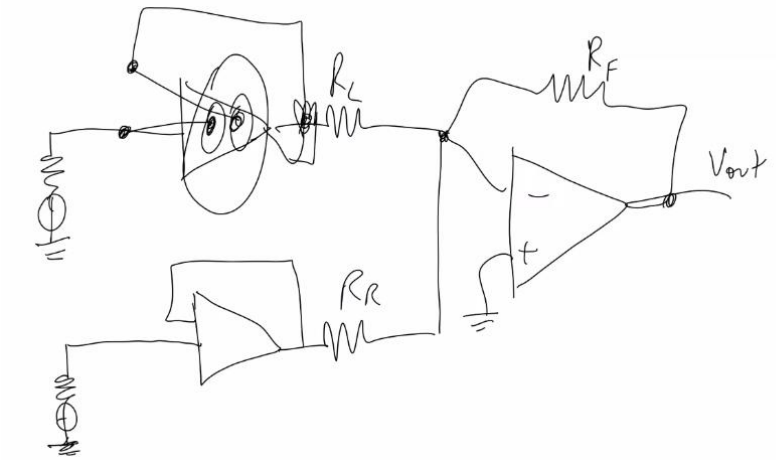
Vout is equal to the sum of all the input currents (Vi/Ri) multiplied by R in feedback. Hence having a resistor instead a virtual ground is bad because if we have different voltage sources and we want to sum them, if we use a resistor, we will have a voltage on the resistor that is also applied to the other sources, so we have a sort of cross-talk. The amplification for each line is R/Ri.

This might damage the voltage source. For this reason is much better to use a ground, and not a resistor.

The output impedance, if the opamp is ideal it is 0. If Rout is not zero, the impedance we see is not just Rout || R + (R1 || R2 || …), but it is this one divided by 1 - Gloop, with Gloop = A(s)*(R + (R1 || R2 || …)).

As for the input impedance, every generator sees R1, R2 etc, and then virtual ground. So it just depends on the input source resistance. This is not ideal because the feedback R is common to everybody, to change the gain relate to a source we need to change R1, R2 etc. To overcome these mismatches, we better use another amplifier, that is the following one.

The source is not applied straight to the R1, R2 etc, but it is applied to an opamp whose gain is equal to 1, that is a buffer. An inverting buffer, that inverts the signal, has always the feedback on the – terminal.

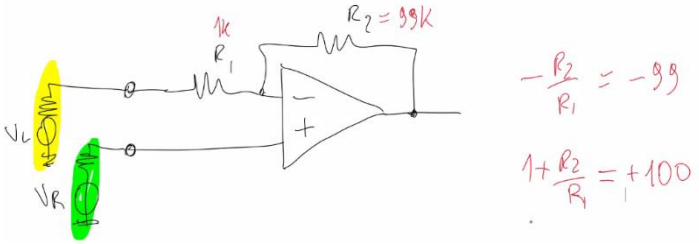


With the added buffers, the impedance in output of it is 0, whereas the input impedance the sources will see is infinite.

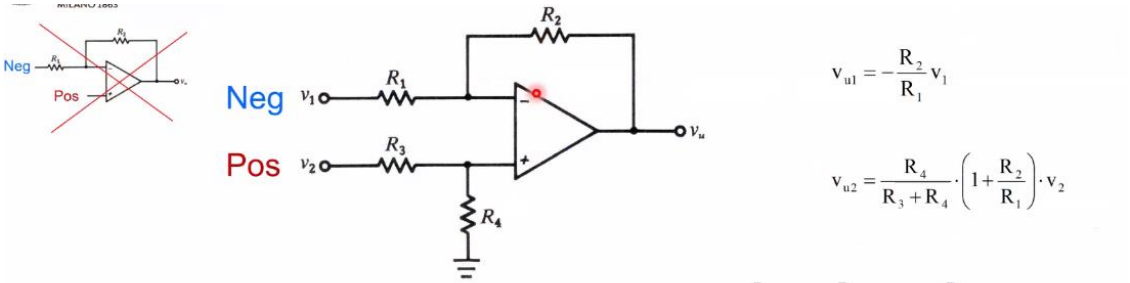
VOLTAGE SUBTRACTOR

The non-inverting configuration has a gain of (1 + R2/R1), where R2 is in feedback → gain always >1, so we will always have an amplification and never a decrease in signal.
Another possible configuration is the inverting amplifier, whose gain is – R2/R1. The gain is always negative and can assume any value greater or smaller than 1.

We use one source to the + terminal, and one on the – terminal. We want to make the subtraction. The source Vl experiences a gain of -R2/R1, the other 1 + R2/R1, and this is not good because the gains are different, so we don't subtract the signals. Hence we need to introduce an attenuation.



With the attenuation.



$$v_{u1} = -\frac{R_2}{R_1} v_1$$

$$v_{u2} = \frac{R_4}{R_3 + R_4} \cdot \left(1 + \frac{R_2}{R_1}\right) \cdot v_2$$

Differential gain (when $R_2/R_1 = R_4/R_3$):

$$v_u = -\frac{R_2}{R_1} \cdot v_1 + \frac{R_4}{R_3} \cdot v_2 = -\frac{R_2}{R_1} \cdot v_{diff}$$

Common-mode gain:

$$v_u = 2 \cdot \frac{R_2}{R_1} \cdot v_{cm} \cdot tol$$

Input impedance: $R_{pos} = R_3 + R_4$

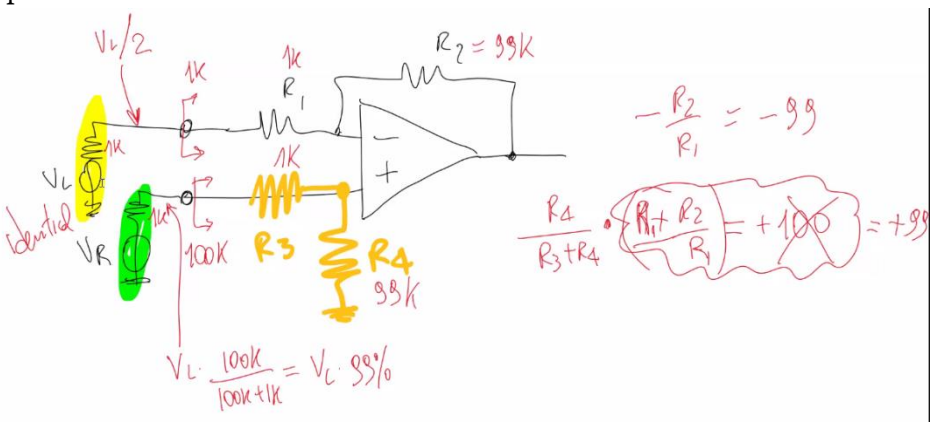
$R_{neg} = R_1$

$R_{neg} \neq R_{pos}$

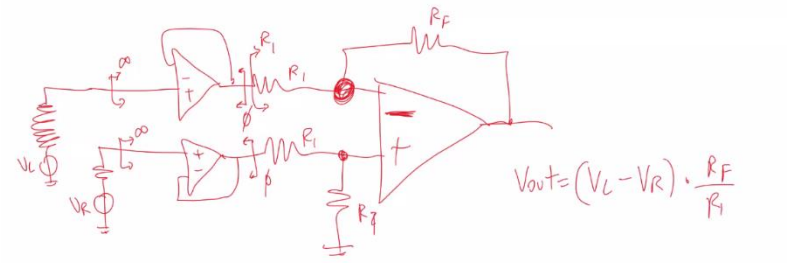
Let's solve this issue…

Output impedance: $R_{out} \approx 0$

The + terminal can move, whereas the - terminal is the slave (VG) and is linked to the + terminal. On the – terminal I see R1 as impedance, on the other branch R3 + R4. Hence even if the inputs source resistances are equal, the output is not 0 because of different input impedances on the two branches in input to the amplifier.

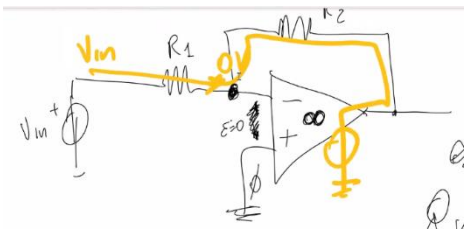


We solve this mismatch by using buffers in input. - terminal is still at virtual ground and now the output is perfect.



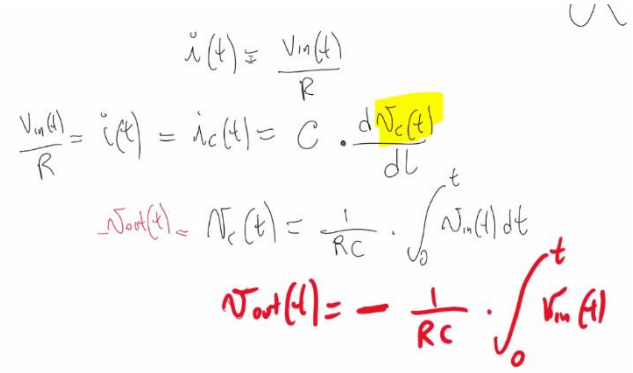
IDEAL VOLTAGE INTEGRATOR

We apply a Vin and node at terminal – will be fixed if I consider a strong feedback. The output is always the one that allows the virtual ground to be actually a virtual ground. We apply whatever Vin we wish but then the current that flows in must flow in the feedback.

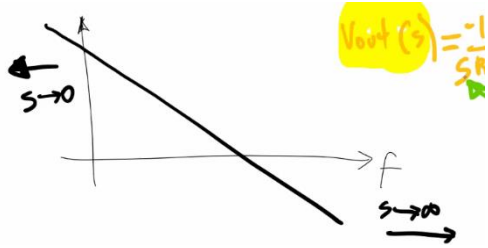


Now in feedback we use a capacitor. In the capacitor, Vin(t)/R = ic(t) = C * (dVc(t)/Dt).

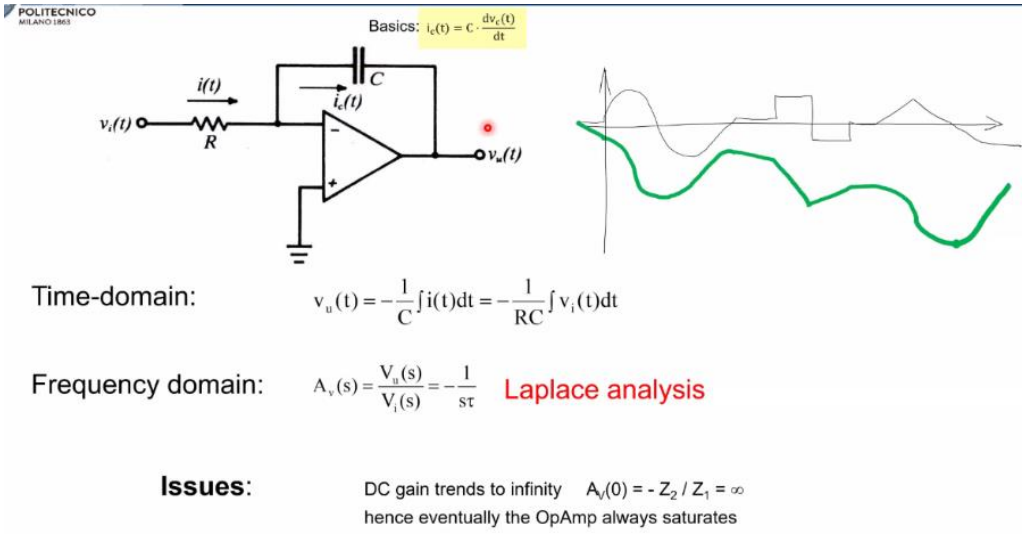
Vc(t) = (1/RC) * Integral(Vin(t)dt) = -Vout. Vc is the opposite than Vout.



The same reasoning can be done in the Laplace domain.
Vout(s) = (1/sRC) * Vin(s).
Since there is the s = alpha + j2pi*f, the plot is the one below, with a decrease of -20dB/dec.

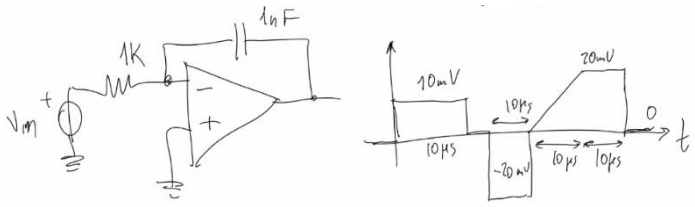


Every time we encounter this decrease, it means we have an integrator. This also means we have a pole in the origin (0 Hz) and no zero. The placement of the slope depends in the frequency where the plot intercepts the x axis, that is 1/2*pi*RC. It is not a pole, but the frequency at which the gain is 1.

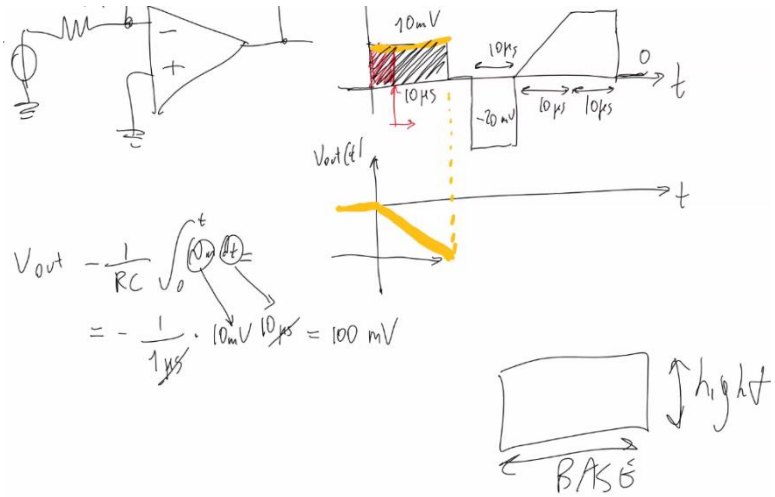


The gain is – (1/RC), that is not a pure number, but 1/[s], 1 over seconds, so that in output we still have a voltage. In the upper right plot we can see that we are integrate the signal in black resulting in the signal in green.

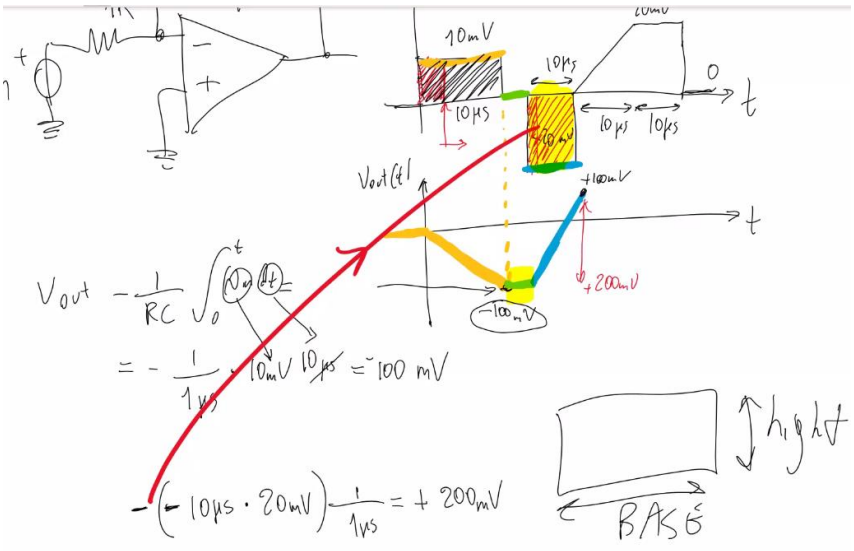
Example



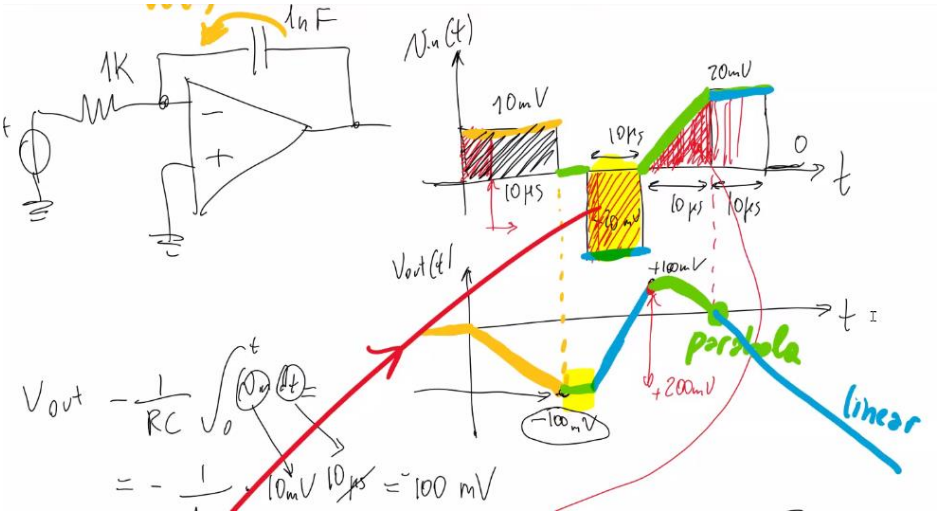
The input is the black signal.
Let's now compute Vout. We need to know the charge accumulated in the capacitor. At time 0, it Vc = 0V, since Vout is ground and then we have the virtual ground. The output is the area of the voltage multiplied by the gain. So initially is 0, but then at a certain time we should compute the integral of Vin multiplied by the gain 1/RC. If time passes, the area increases in the first 10 us, and since there is a minus in the equation, the output will be negative; since the input is constant, the output will be a straight line. The total area of the first part is 10 us * 10 mV.



Then we have a 0 to integrate → the output will not move, and then we have to integrate something that is negative. Since the gain is negative, we will find a transition that will be increasing. The computed area must add to the previous vale of Vout we had, that was – 100 mV, so since it is of 200 mV, we end up at 100 mV.

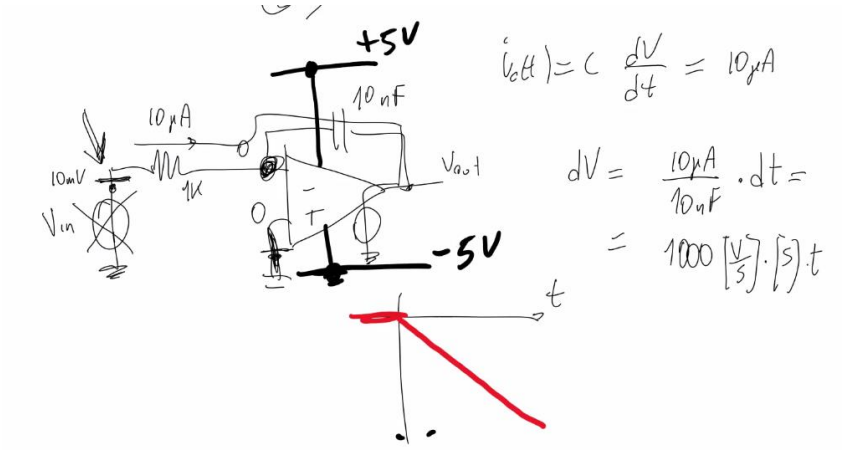


Then we have the triangle: - (10us * 20mV)/2 * (1/1us), so we must add -100 mV from the 100 mV point. Since the trend is linear, we will have a paraboloid decrease.



The problem of the ideal integrator is that the current is 0 in DC if the capacitor is not charging up, whereas it should be present a current if I apply a constant voltage at the input. The opamp tries to drink a constant current in DC, but a constant current through a capacitor in DC causes dV(t)/dt to be not nihil.

If in this integrator we have a constant signal due to the input or an offset, due to the virtual ground concept a current will flow through the 1k resistance, the current in DC flows through the capacitor, and if i is constant, voltage on the capacitor keeps increasing with an incredibly big rate, so the output starts from 0 and then when an offset is present the output continues to increase up to the saturation of the power supply of the opamp. Once it is saturated, it won't work anymore → <u>problem of saturation with constant currents</u>.

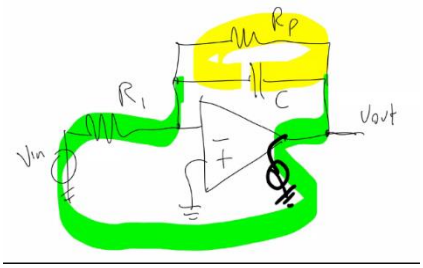


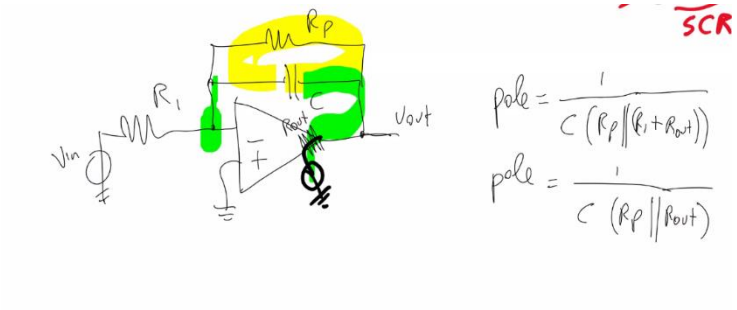
To avoid saturation, we should introduce a component that doesn't let the feedback to degenerate → we put a resistor in feedback, so that the gain doesn't saturate.

REAL VOLTAGE INTEGRATOR

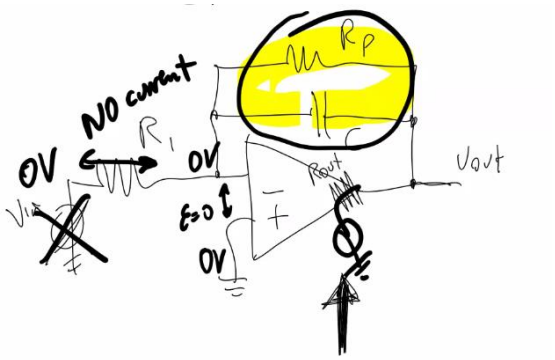
In DC the capacitor is open, so the gain Rp/R, whereas at HF it is 0 because the C is a shortcircuit. Hence we start flat with a constant gain and then goes to 0 → we have one pole, not at the origin.

The frequency of the pole is 1/2*pi*Rp*C. But let's compute it. We need to find the total R related to the C. C sees Rp parallel to the green path, that is made of Ri and Rout. Otherwise, I can see Rp and then virtual ground on the left and Rout on the right. But also this

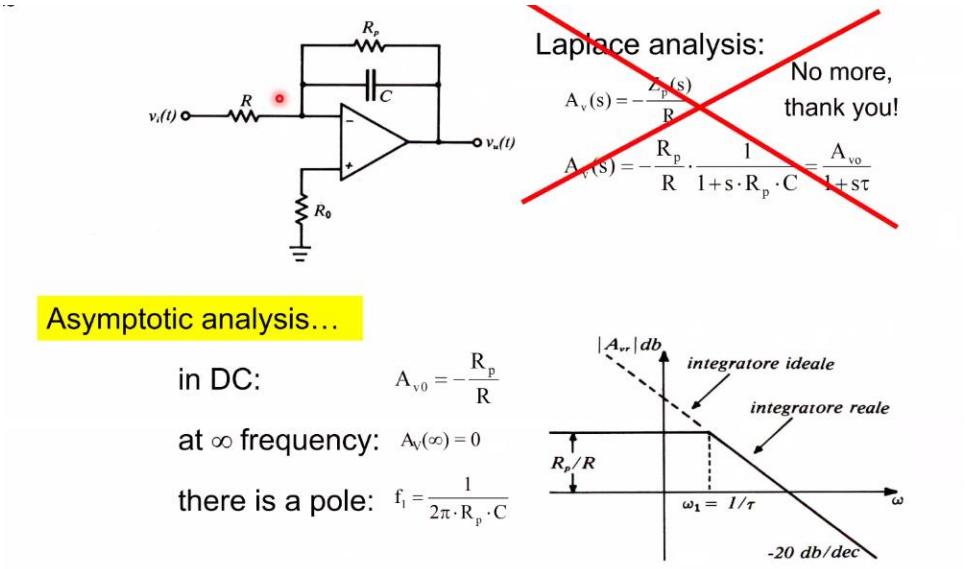




is incorrect.
When studing the pole, we switch Vin of, but not the one related to Vout inside the opamp. The correct analysis is the one below.

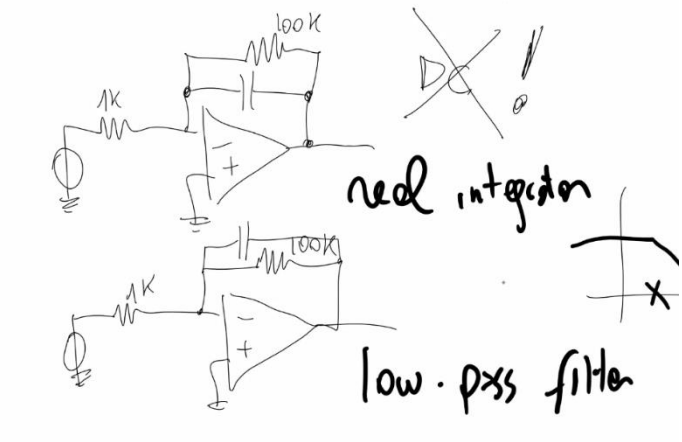


But then the capacitor cannot move as it wishes? One end of the capacitor touched the virtual ground, so as soon as the virtual ground moves, the opamp moves the output and hence the capacitor is in between constant and equal changes. So the output node can go up and down in order not to let the other arm (where the virtual ground is) change.



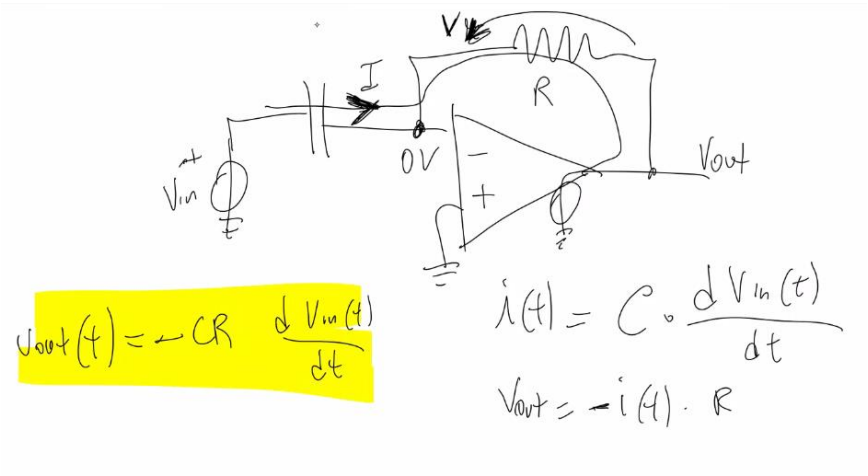
The frequency at which the gain is equal to one is given by the GBWP: 1/2*pi*R*C, that is the same frequency of the ideal case.
The ideal integrator saturates ad DC, so we added a resistor to get a real integrator. I could have also considered a normal amplifier and later added a capacitor, obtaining a low pass filter. Hence the LP filter and the real integrator have the same behaviour. But which is? It is an amplifier when the gain is flat,

hence from 0 to the pole, and then it is an integrator when the gain is no more flat, wherease it is a LP filter when the pole is in the middle of the frequency range.

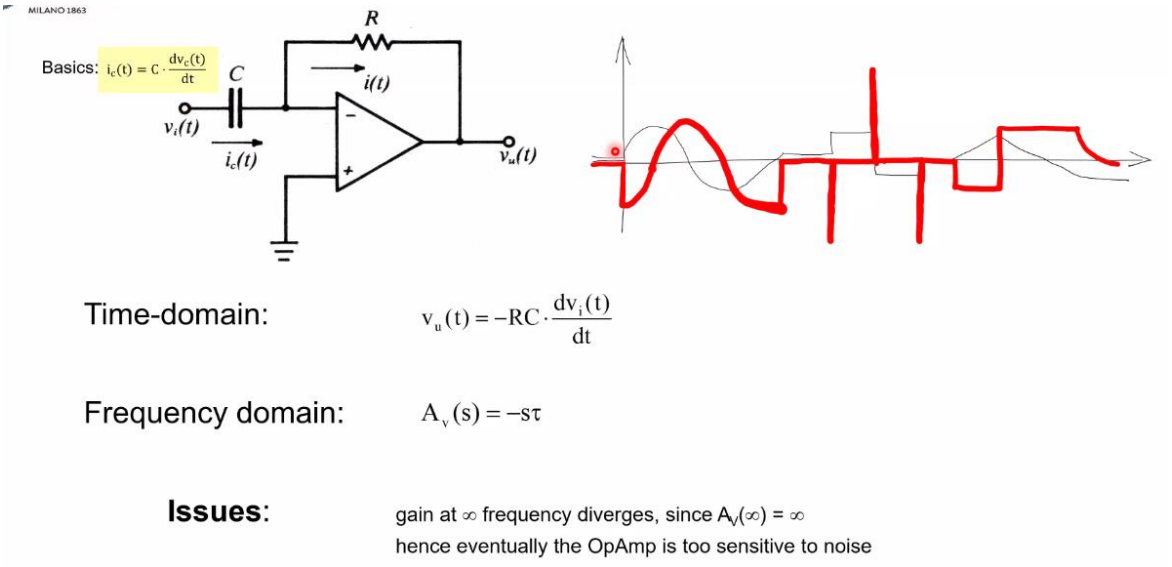


IDEAL DERIVATOR

In feedback we place a resistor.



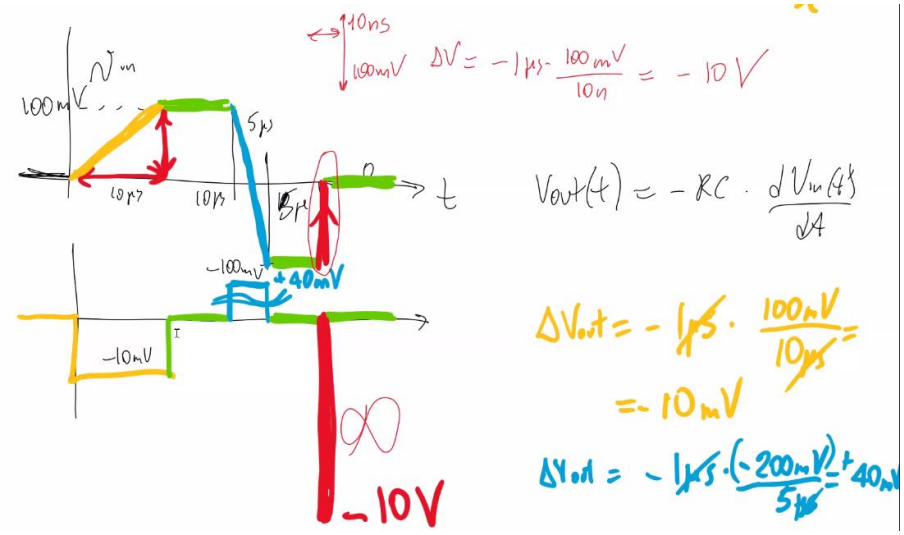
This time the gain is in seconds.



If the signal in input increases, the output decreases and vice versa. If the transition happens in 0 time, in output we will have a Dirac delta.

Example

The derivative in the first phase is the ratio between the two delta; since the slope is constant, the output will be constant. When the input is 0 the output is 0. When we have the sharp transition, the slope is

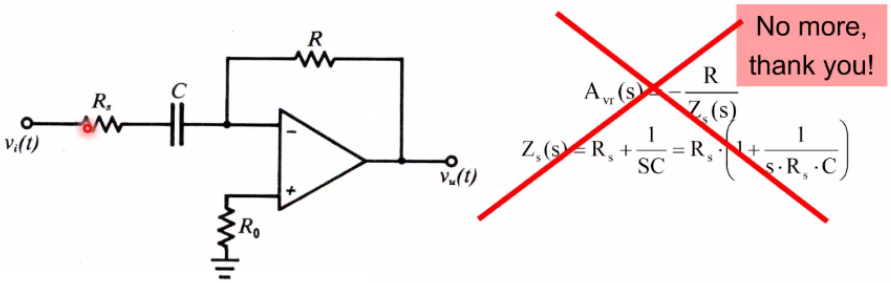


infinite, so we have a Dirac delta (lasting 0 time). However is not exactly infinite, but it goes to a finite value because it will, in reality, have a very very small duration.

The problem with this stage is the lack of control at HF, and we are shortcircuiting a virtual ground with an input. In the ideal derivator, we have a zero in the origin, with the frequency where the gain is 1 that is 1/2*pi*RC.

We want to force it in saturation soon or later with a pole, placing a resistor in series with the capacitor.

REAL DERIVATOR

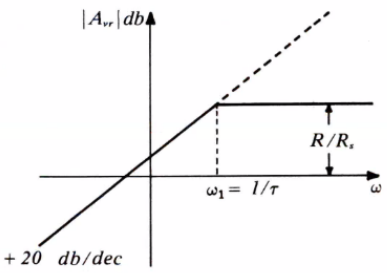


Asymptotic analysis…

in DC: $A_V(0) = 0$

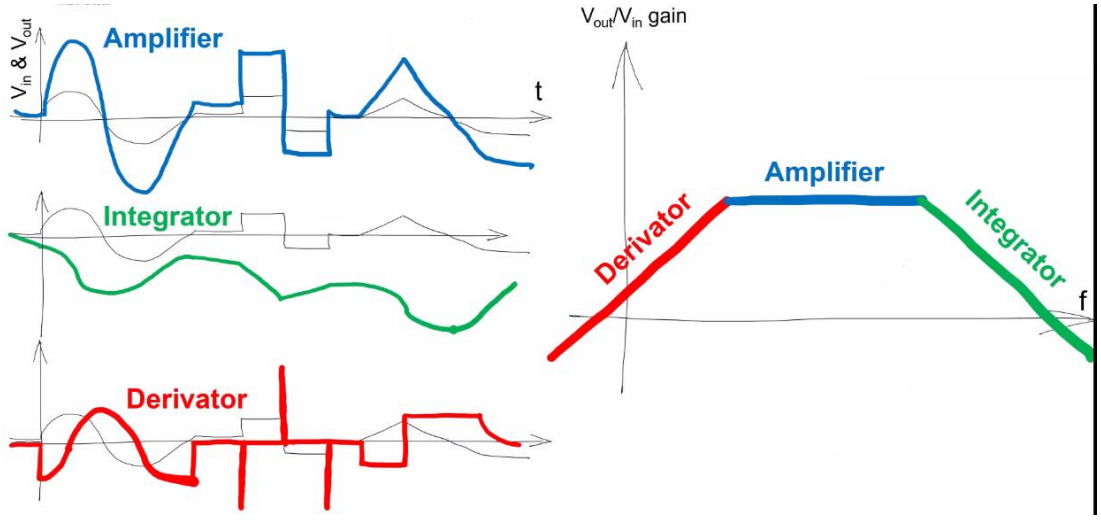
at ∞ frequency: $A_V(\infty) = -R/R_s$

bandwidth: $f_1 = \frac{1}{2\pi \cdot R_s \cdot C}$

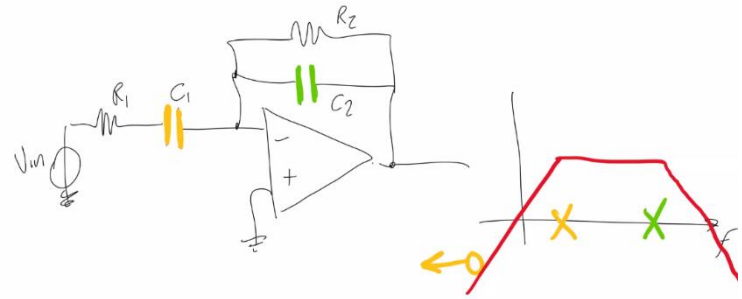


Now the pole is 1/2*pi*RC, whereas where the gain is 1 we have Rs. Also in this case we can consider either a real derivator or a HP filter. Again, the reasoning on when it is an amplifier or a derivator holds as before.

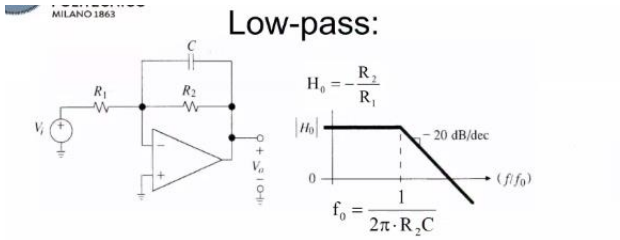
Recap

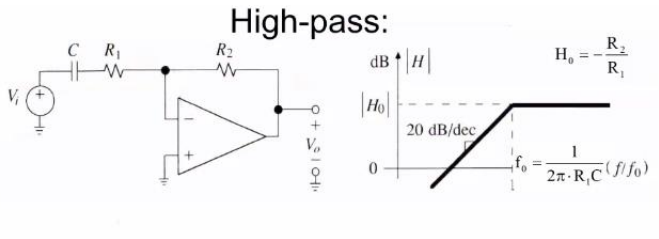


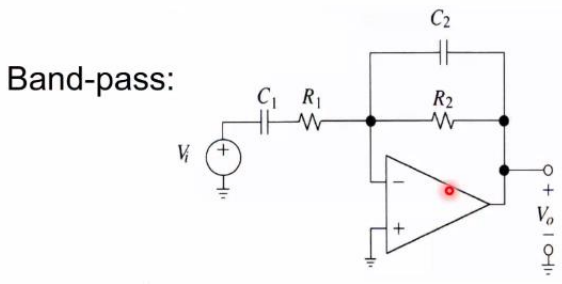
BANDPASS FILTER

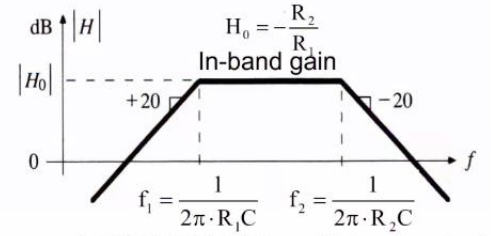


With symmetrical slopes in the Bode plot. The gain is -R2/R1, only if the input path has the pole at lower frequencies than the feedback pole. If we do the opposite, the capacitor C2 kills the signal with the capacitor C1 that instead lets the signal pass → the signal will die soon, so it's us that need to choose the pole correctly.







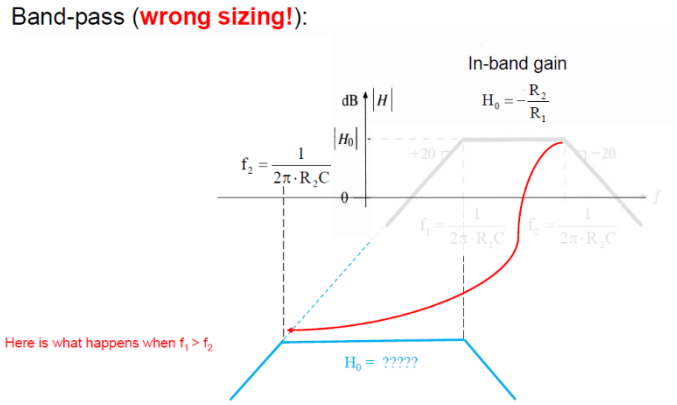


Lower band edge (lower corner frequency)

Higher band edge (higher corner frequency)

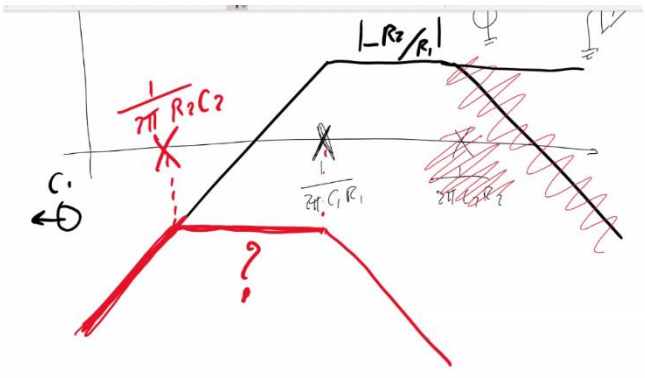
But only if $f_1 < f_2$, otherwise... what happens ?!?

Band-pass wrong sizing



We have to size the poles so that one happens at LF at the frequency we wish so that at midrange we are in a configuration where the input capacitor is shorted and the feedback one is open, so the gain is -R2/R1. If the pole related to the feedback capacitor is at LF, so we do the wrong sizing (as in the image above), we have an intermediate range very bad, because if we increase the frequency of the signal, instead of having the input capacitor shorted and the other one open, it is the opposite.
To compute the gain in the midrange in the case the pole are inverted, we can use the products. We know that in the case the poles are correct the gain is -R2/R1, and we have to put together the two trends.

Let's compute the height of the filter in case of bad conditioning, hence the pole related to C1 at higher frequencies than the one related to C2. In red what happens in case of bad conditioning, in black good conditioning.



The gain of the red flat part is not G = -R2/R1, but it is this quantity multiplied by a scaling factor.

$$G = \left|-\frac{R_2}{R_1}\right| \Big/ \frac{\frac{1}{2\pi C_1 R_1}}{\frac{1}{2\pi R_2 C_2}} = \frac{\frac{R_2}{R_1} \cdot C_2 R_2}{C_1 R_1}$$

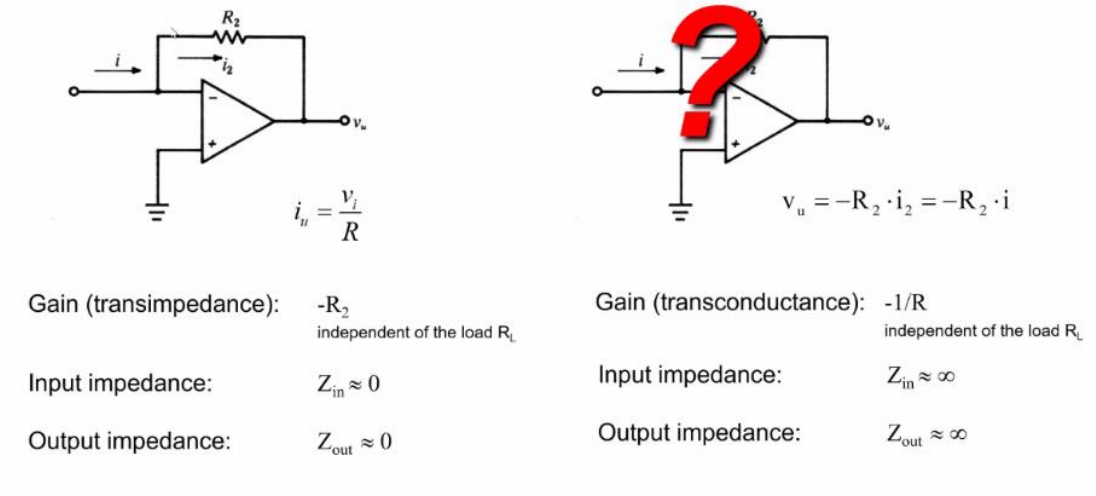
$$= \left(\frac{R_2}{R_1}\right)^2 \cdot \frac{C_2}{C_1}$$

Depending on C2 and C1 we will have a lower gain than expected.

TRANSIMPEDANCE AND TRANSCONDUCTANGE AMPLIFIERS

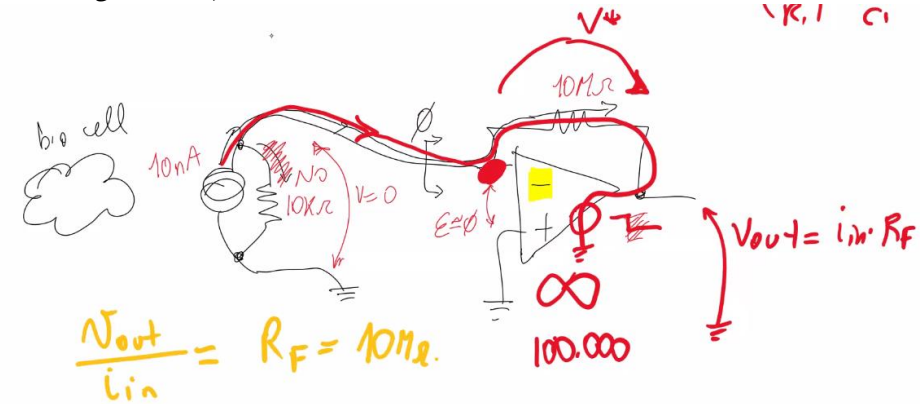
Transimpedance amplifier

The input impedance of the right circuit is 0 because we have a virtual ground at the – terminal. Also, the output is a very low impedance, and if we have an amplifier with low input impedance, we can connect a current source, whereas if the output impedance is low, it is a very excellent voltage source, so it is actually a I to V converter.



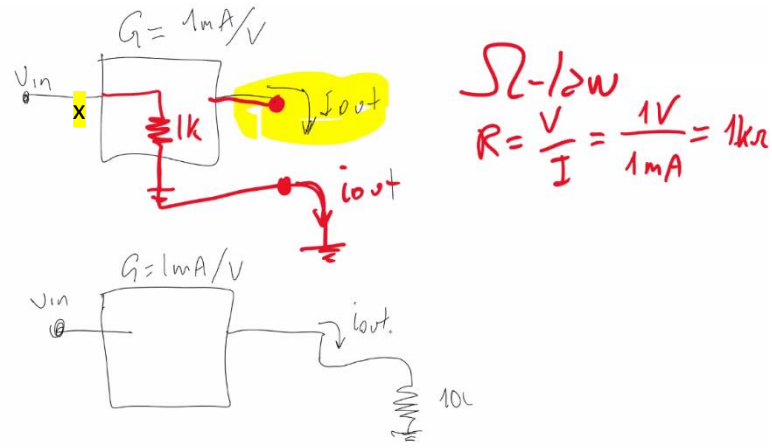
Gain (transimpedance):	$-R_2$ independent of the load R_L	Gain (transconductance):	$-1/R$ independent of the load R_L
Input impedance:	$Z_{in} \approx 0$	Input impedance:	$Z_{in} \approx \infty$
Output impedance:	$Z_{out} \approx 0$	Output impedance:	$Z_{out} \approx \infty$

The following an example. Since - terminal is virtual ground, no current is lost in the source (resistor in parallel to the current generator).



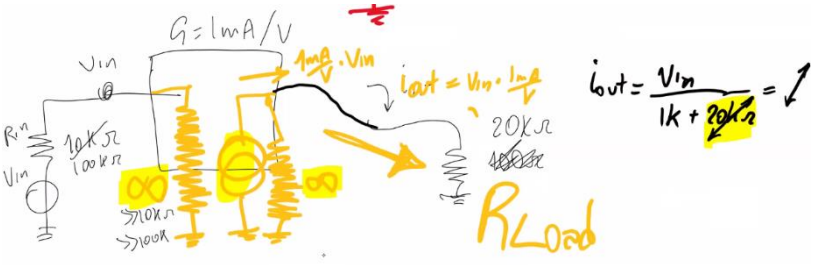
Transconductance amplifier

I want to convert a voltage into a current. I want an output current proportional to the input voltage, with a gain that is for instance 1mA/V. The typical erroneous solution is the ohms law, because it is not n amplifier, because I would have the output node for the current attached to ground if I use a resistor,



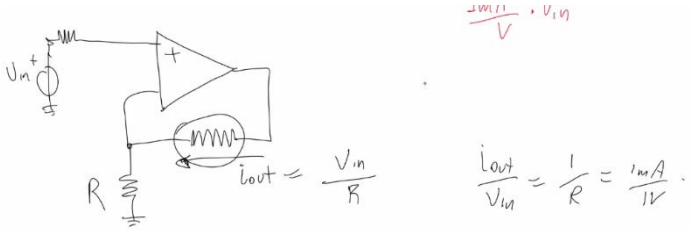
while now I want a floating output to be attached to whatever the customer wants, and if I use a simple resistor the current in output depends on the load. So x is wrong.

I want an amplifier with infinite input impedance so that if the input impedance of the source is whatever, the input impedance of the amplifier will dominate. Then I want the output to be a current generator, so the output impedance also infinite, so that it is a perfect current source.



If we consider an amplifier, connecting directly the source to the terminal of the amplifier is ok, because we verify the condition of infinite input impedance. However, in output to an OL amplifier we have 0 impedance, so it is a perfect voltage source, not a current source in output. So I put a feedback signal in output, as below. If Gloop is infinite, I return to the input what I've applied. Since the loop is negative, I have epsilon = 0 across the terminal, and so Vin is also applied to the output.
However, I want the current to flow through the output resistor of 1k and the load. Now, the current is however no more Vin/1k, but Vin/(1k + Rl). So how can I do it?

The current comes from the amplifier, that is a voltage source. To solve the problem, I attach the load to the output load and then I connect the output load to the feedback and the 1k resistor.



The load is in series with R as in the transimpedance amplifier, but the current is independent on Rl.

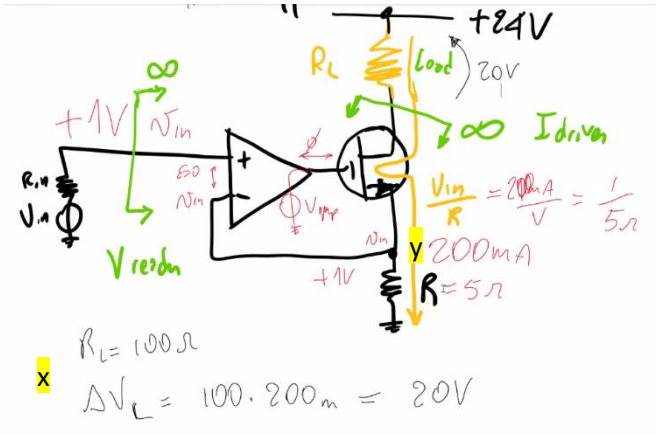
However, a limitation of this circuit is that (gain of 1mA/V), if Rl is huge, such as 100k compared to the 1k of R, the 1mA through a 100k resistor means 101V provided by the amplifier, and no opamp can provide such a high voltage.
Conversely, if Rl = 2k and we want a higher gain, such as 200mA/V, I can use a 5ohm resistance as R. But the 200mA should go through Rl and maybe the amplifier cannot provide it. Hence this stage is good but not always works due to the current and voltage limiting capabilities of the amplifier. However, it is good if we want to force a current in a cell.

How can the current be high in the load but not in the opamp?

Final implementation

I still need the resistor and I don't connect the opamp to the resistor, but I use a transistor in output, and the transistor will provide the huge current I need. The opamp provides no current but the voltage required.

The load is connected somewhere to ha high voltage. The current through R is fixed by the stage. So the input impedance is infinite, the output impedance seen by the load infinite, so it is a perfect V reader in input and I driver in output (24V power supply are given by the computation x).

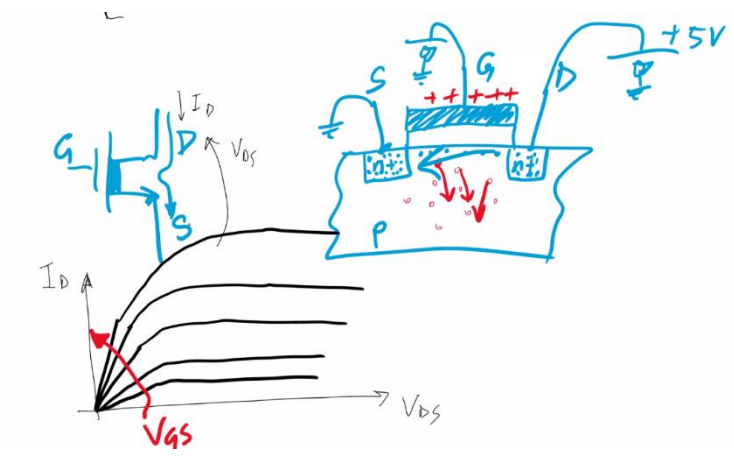


What is the voltage in output of the opamp?

The transistor is a very nice slave, so the buffer gives the voltage to it and the transistor gives the current that is ok with my requirement. If I apply Vin, the opamp output voltage increases, Vgs of the transistor increases and also the current flowing through it. If current increases voltage at – terminal increases. If – terminal voltage is different than + termina, the output is still high, up to the point the two input voltages are the same and the epsilon is 0.

Eventually, if Gloop is infinite, epsilon is 0, but if the gain is finite, epsilon will be e.g. 10nV, that are however negligible when transferred in output.

In output of the opamp, we need to consider MOS transistors. They are devices using a p-doped silicon with n+ doped introduced with diffusion. With no voltage to the gain, if we apply a Vds voltage, no current happens because we don't have current. So we need to apply a positive voltage to the gate so that the holes of the p-doped material are pushed down and electrons starts to pass across the material. The current is between drain and source (Id).



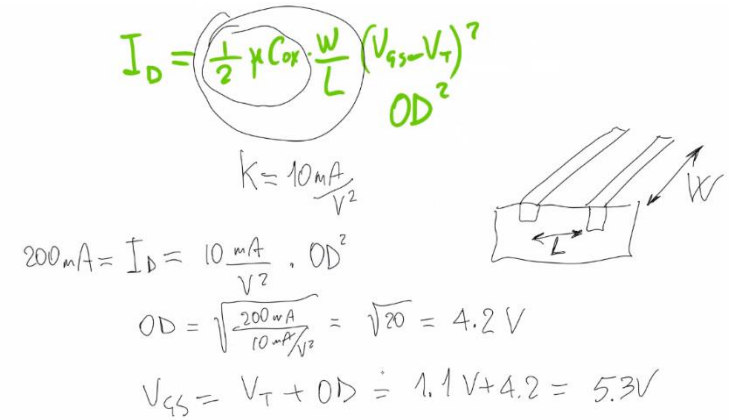
If Vds is increased too much, we suffer from a saturation of the current Id. The minimum voltage to be applied to the gate to have a current is called threshold voltage (Vt). The difference Vgs – Vt is called Vov, overdrive voltage.

The typical equation of a MOS transistor is the following, in the saturation region.

$$I_D = \frac{1}{2}\mu C_{ox}\cdot\frac{W}{L}(V_{GS}-V_T)^2$$

It is proportional to the squared Vov, while u is the electron mobility of the channel, Cox is the parasitic capacitance of the Silicon dioxide insulator, while W and L are the physical dimensions of the transistor.

So going back to the previous circuit transconductance amplifier, if we apply a Vin = 1V:



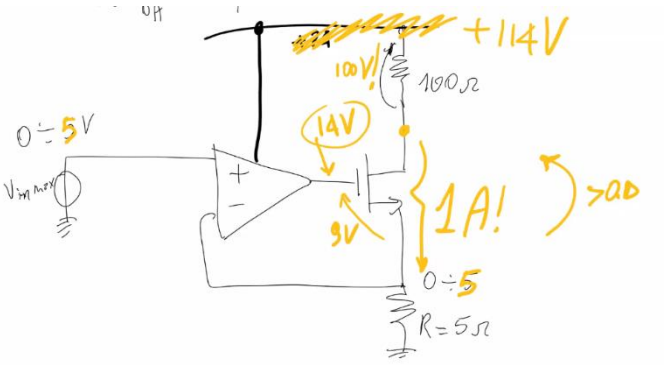
Hence between the upper terminal of R and the output voltage of the opamp we have to apply Vgs. Since we have Vin in the upper node of the resistor R, so we perform the Kirchhoff law and Vout of the opamp is 6.3V → we need a power supply higher than 5V for the transistor.

NB: if we apply a Vg and Vds = 0, we are in the ohmic region, where the current varies depending on Vgs and also on Vds and the channel is uniform. If now we increase Vd, e.g. to 2V and we apply 5.3V at the gate, we have a very high Vgs, but a lower Vgd (3.3V) → if we increase the drain voltage then we have channel at the source side but no more at the drain side → pinch-off condition if we reach Vt, and the green equation holds. So in the saturated region I MUST have channel at the source but NO CHANNEL at the drain.
Hence I should not have Vt, but a smaller value for Vgd. If Vgs > Vt and Vdg < Vt, it is ok and we are fine (channel at the source and no channel at the drain), and it is ok because the current doesn't depend on Vds. But we have to verify we are in this condition.
We want the two inequalities before, meaning that the condition to be satisfied is Vgs – Vt > Vov. The saturated region is perfect because if I change Rl, I don't want the current to change, so eventually Vgs changes but Id remains constant.

NB: we must verify the MOS in the saturated region in the end, that occurs if Vds >= Vov. Let's see from the example. Vs = 1V, Vov = 4.2V, so Vd must > 5.2V to have a working stage that acts as a current source.

If we let the transistor operate in the saturated region we are sure that it acts as a current source, so that if voltage at the drain changes the current will never move. But if for any reason we are in the ohmic region, than the current through the source won't be the value above, so if current reduces, voltage y would change and be lower than expected, so an epsilon will appear at the input of the opamp. So even if the transistor enters the triode region, since there is feedback the opamp will start to increase Vgs so that the current in the MOS is enough to enforce the overall feedback.

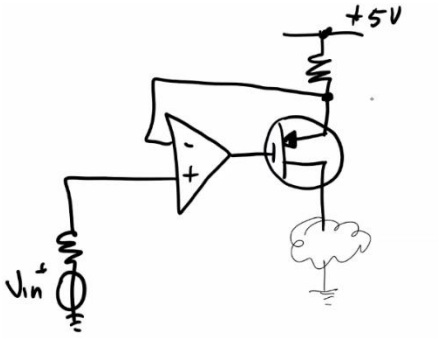
When we apply Vin,max, the circuit is the one aside. Vin can move from 0 to 5V, the bottom resistor is 5 ohm so with 1V in input we have 200 mA, but if we apply 5V the current should be 1A (5V/5ohm). If the current is 1A, we can compute the required Vgs, that is for instance 9V, so we run the risk that if we apply 5V in input the voltage at the output of the opamp should be 14V (KVL), so PS has to be higher than 12V. Moreover, if on the drain of the



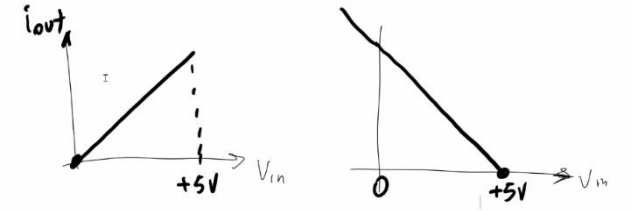
MOSFET we have our load that is for instance a 100ohm resistance, forcing 1A in 100ohm means that in the worst case the voltage drop across the load will be 114V, so PS of 24V is not enough.
Moreover, we have to try to enforce saturated region, so at least one overdrive at the drain node → too high power supply for this application.

Hence if we want the drain node to be above of at least one overdrive, then we require the PS to be too high, so maybe the design constrain can be relaxed and instead of operating the MOS transistor in the saturated region but in the ohmic region, with a very high Vgs. If the opamp can do this, current will be the one we want anyhow and everything will work, but we have to be sure that the opamp can do this and that the MOSFET can withstand such high voltages → as a rule of thumb let's try to operate with the transistor in saturation, the opamp with a proper value and also the PS.

There are some application where the load is connected to ground. Hence we cannot connect it directly to the power supply, because one end of the load is to ground. We simply use a p-MOS transistor and invert everything. When Vin = 5V, across the resistor I have 0V and 0 current. Whereas if Vin = 0, the current is the maximum.

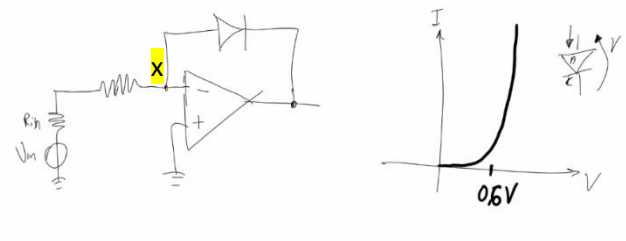


So if Vin varies between 0 and +5V, iout changes as on the left (condition of n-MOS), whereas in the other case it is the right plot.



LOGARITMIC CONVERTER

Let's consider an opamp with a signal applied, e.g. entering in the – terminal and placing a non-linear component in feedback, e.g. a diode. In forward bias, the diode has the displayed characteristic, with a current of more or less mA in the forward region and in the order of uA in the inverse region.
When Vin is high, the voltage at the – termina increases, so the output start to becomes negative and hence the voltage across the diode rises in forward direction. Let's imagine to apply 1V as Vin and use a 1k resistance. The current is 1mA, will flow through the diode and since it is on, the output voltage will be 0.6V. If we now apply Vin = 5V, the current will be 5mA, so Vout will be like 0.63, because the voltage increases a little bit across the diode.



So the output stays almost constant. If we consider the equation of the diode, the current has an exponential trend depending on V thermal (Vth) where Vth = kT/q (25mV at RT) → **Id = Is*[e^(Vd/Vth) – 1]**. So Vth compared to Vd causes an exponential increase of Id in reverse bias. In forward bias, we can forget of the -1.

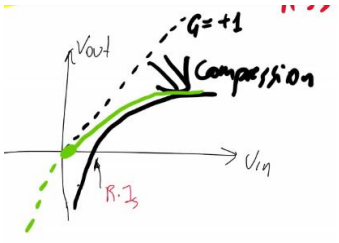
Small signal analysis

Let's consider a small variation around a central point and see what happens in output.
If we apply something positive at the input, the – terminal of the opamp goes positive, but then the output goes negative → the diode turns on and the voltage at node – is pushed back to a negative value. So the feedback is negative and node x acts as a virtual ground.
Hence the current that enters the feedback circuit is Vin/R, and we are forcing the opamp to drive the current and let it flow through the diode. Now let's compute the dependency of Vd over Vin.

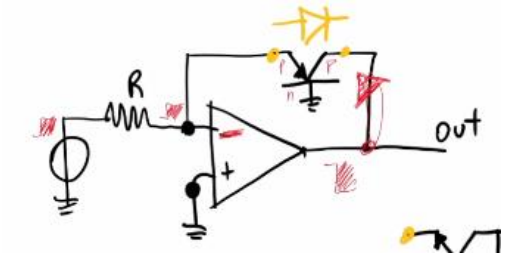
$$\frac{V_{in}}{R} = I_D = I_S \cdot e^{V_D/(kT/q)}$$

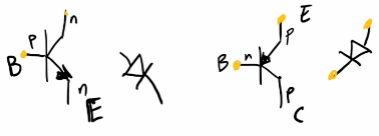
$$V_{out}(t) = \frac{kT}{q} \ln \frac{V_{in}}{R \cdot I_S}$$

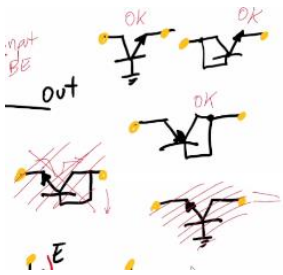
The relationship between Vin and Vout is no more an amplification but it is a logarithmic scaling. So the stage behaves like a logarithmic amplifier, meaning that a compression is performed over the input signal.
However, if I forget the – 1 in the equation of the diode, the logarithm is equal to 0 when Vin = I*Rs. This is not true because the real curve of a logarithmic amplifier never goes towards negative values of Vin, otherwise the current would flow in the opposite direction, but this is not possible. So the system works if Vin is positive (green curve).



Similarly, we could use a BJT transistor (eventually with base and collector shortened). This transistor must be inserted into the circuit in a proper way. We want to apply a voltage and convert it into a current, and the transistor in feedback must go as the feedback acts. For instance, a transistor connected as below cannot work, because we have a pnp transistor where the collector is brought to a negative value. The best thing is to apply something on the transistor that acts on the base-emitter junction, because it's this junction for a npn that is a lovely forward bias, while in a pnp transistor the other way.



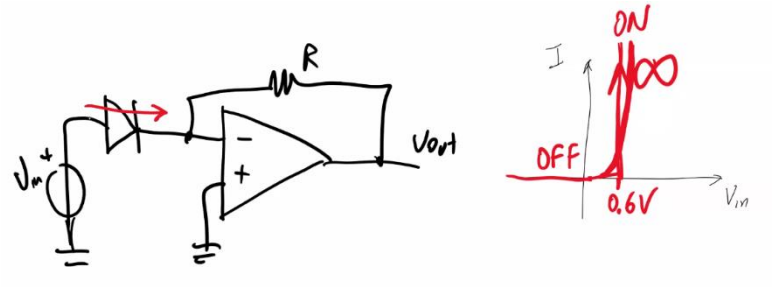


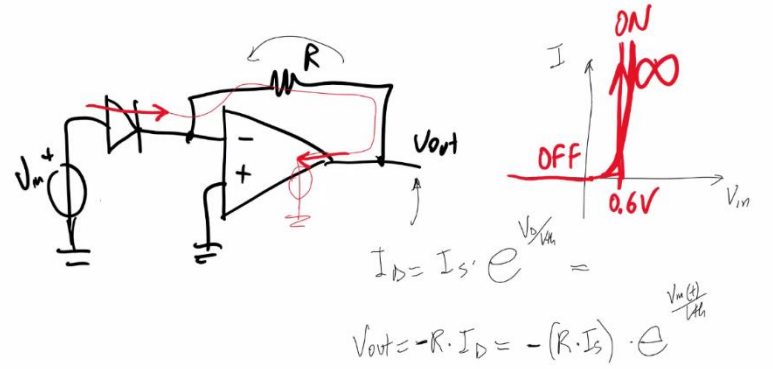


So the output of the opamp must act on the input of the transistor, that is the base emitter junction.

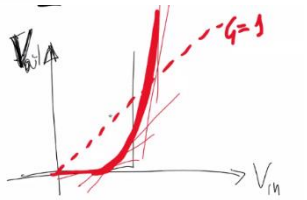
EXPONENTIAL AMPLIFIER

In this configuration, if Vin is negative, the current is 0, whereas for positive Vin the current is infinite. As soon as the diode turns on, we have the maximum current that can flow through the diode. Still, Vin has to have reasonable values, so e.g. 2V is too much, it must be reasonably close to 0.6V otherwise the diode breaks or the opamp isn't capable of providing an output.

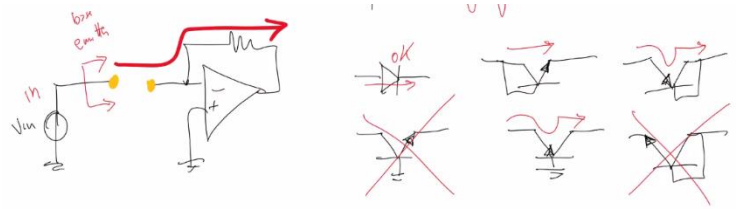




The final result is an exponential processing of the input signal. It is not an amplifier with a gain of 1, but something that has a higher and higher gain as Vin goes close to 0.6V.
With this stage we are not compressing the dynamics as in the previous case, we are expanding it.



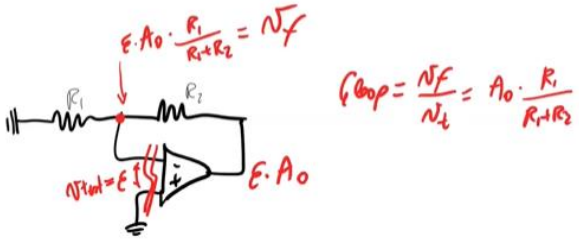
As before, the Vin can be connected to the opamp also with another component, e.g. a BJT transistor. Since we want to move the pin, the input should be attached to the terminal of the transistor that is not the collector, but the base or the emitter of the BJT.



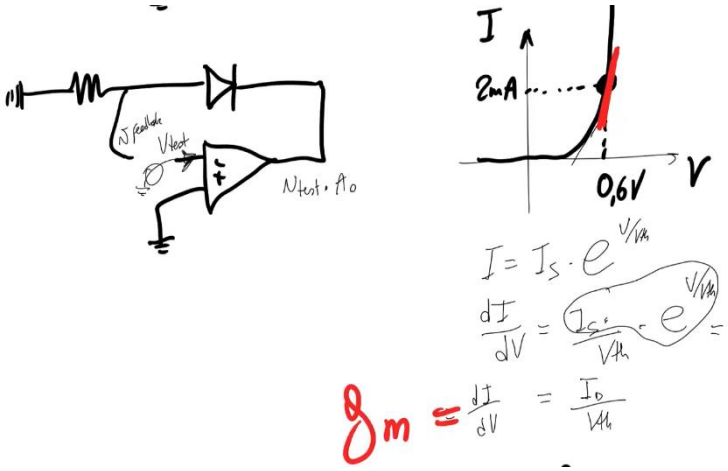
If the Vin is negative, the diode should be reverted and consequently the configurations of the transistors.

Gloop computation in the Logarithmic converter

When i want to study the Gloop we have to turn off the source, wherever it is and we cut in the loop, applying then an epsilon to the circuit to then see what is the signal that returns back to the input. At the output, in a general resistive case, we have epsilon*A0 and eventually at the – terminal we have epsilon*A0 multiplier by the resistive partition. Hence then Gloop is v_f/v_test.

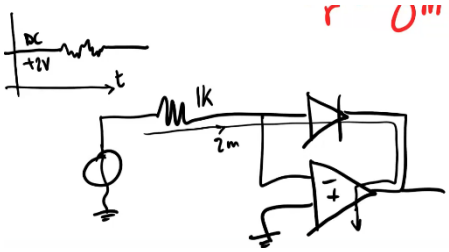


Now, if we have the diode we have to study a completely different circuit, but with the same procedure. Unfortunately, we have a non-linear component. So in output of the opamp we will have v_test*A0, but then we apply a voltage to a series of a resistance and a diode. If we consider small variations, we should substitute the diode with its proper small signal equivalent. We know the i/o curve of a diode, and if we fix a given position on the curve, e.g. at 0.6V, we can compute the slope of the curve, so how the diode behaves around that position, calculating the derivative.

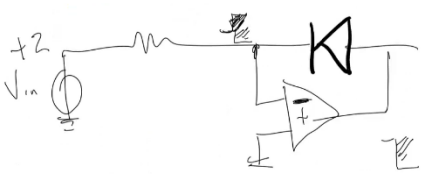


dI/dV is defined as the transconductance (gm) of a diode, that is the inverse of the resistivity of a diode and related to the I-V curve slope. Hence around the position we can consider the diode to act as an impedance whose value is 1/gm, that with 2mA and Vth = 25mV at RT is 1/gm = 12.5 ohm.
Now that we have substituted the diode with a resistance, we can calculate the Gloop with the resistive divider.

This is done assuming a signal in input that has a DC value with some high frequency signal superimposed to it. If for instance the DC value is 2V, it means that we have 2mA flowing through the diode in DC, so I know in which position of the i/o diode curve I am.
Then on the top of the DC value I superimpose my AC signal.

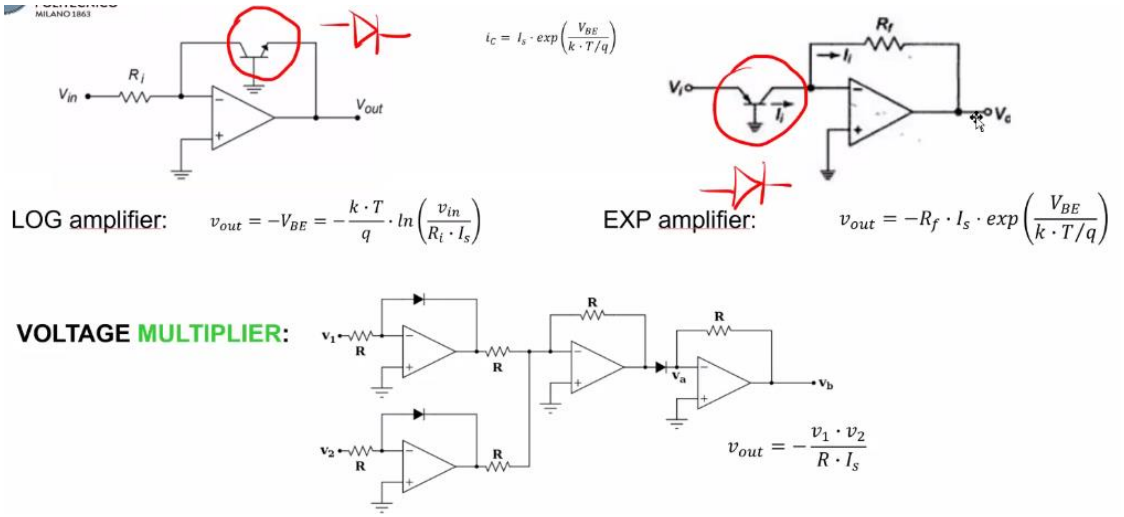


Inverted Diode

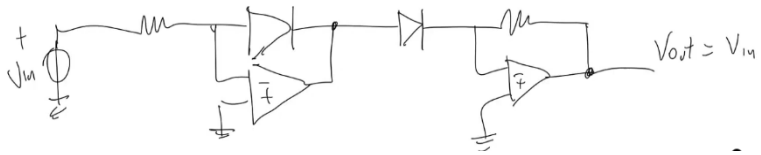


If I invert the positioning of the diode, if in input we have 2V, we try to increase the – terminal value, and in principle the feedback should go negative. If there was a resistor in place of the diode, we have a negative feedback, but since we have a diode, the diode cannot work because it's reversed biased, so no current flows and there in also no current in the input resistance. Hence the whole Vin is transferred at the – input of the opamp, and at the + terminal we have GND. So the feedback in this case is off, not broken but off and the output will saturate to the positive power supply.
Conversely, if a negative signal is applied in input, the feedback is on again.

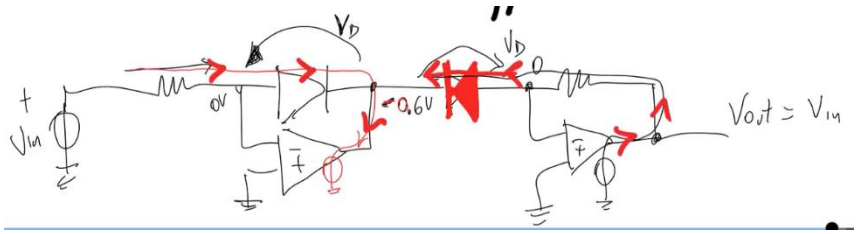
Recap – Exponential and Logarithmic Converters



To have an output that is theoretically equal to the input it would be sufficient to put an exponential stage followed by a logarithmic stage.

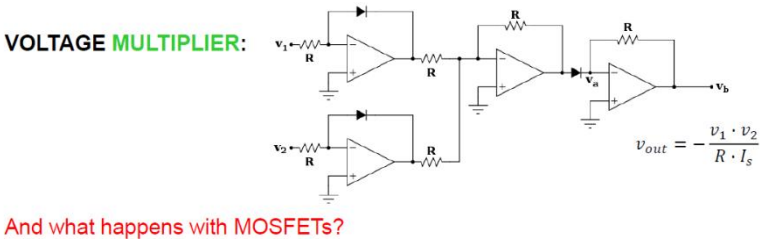


But in this configuration above the system won't work because with an input positive signal, the output goes negative, and hence the second diode in that direction won't be on → need to invert it.



We can see that into the first opamp we have two currents, one from its resistance, the other one coming from the other diode. Eventually, if the two resistances are the same, Vout = Vin.

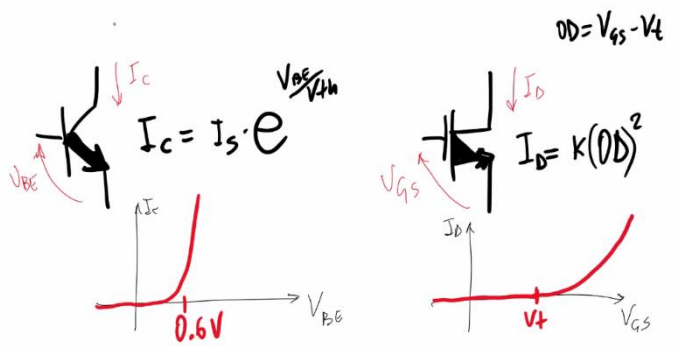
VOLTAGE MULTIPLIER



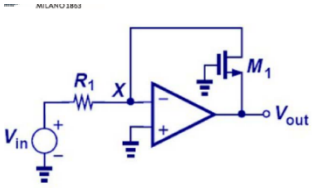
It's an analog processing to multiply two signals.
We apply the V1 and we compute its logarithmic, and the same for V2. Then the outputs are summed and the intermediate sum is exponentially amplified.
So the Vout is the exponential of something and that something is the sum of two logarithms. So we have the multiplication between the two input signals exp(ln(1) + ln(2)) = 1*2.

SQRT AND POWER2 CONVERTERS

We use MOSFETs instead of BJTs.



The Vgs vs Id relationship is quadratic in the MOSFET. In the BJT I like to keep the base-emitter on, while here I like to have a charge channel between gate and source. So the BJT curve is proportional to e^Vbe, while the MOSFET to Vgs^2.



$$i_d = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2$$

SQRT amplifier:

$$v_{out} = -V_{TH} - \sqrt{\frac{v_{in}}{\frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot R_{in}}}$$

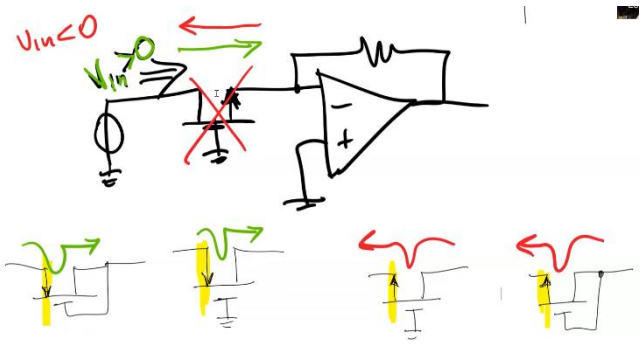
POWER2 amplifier:

$$v_{out} = -R_f \cdot \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (v_{in} - V_{TH})^2$$

If I apply a positive input, node x tries to go positive, so the output is negative, but if so, since the gate of the MOSFET is grounded, the source goes negative and Vgs>0 and channel forms, so a current flows top bottom. If so, node x tends to decrease → negative feedback.
Hence the current Vin/R1 must flow in the MOSFET thanks to the feedback action due to the opamp on the source of the MOSFET, we aren't "pushing" anything from the drain.

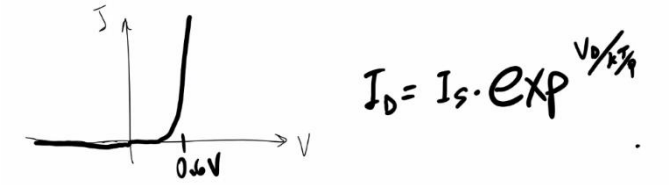
Of course, the sqrt operation is not just on Vin, but on Vin divided by something that is a voltage. Again, the positioning of the MOSFET is crucial, also depending on the sign of the input signal.

Furthermore, if the transistor is in the input branch we have a power2 amplifier. But in this case is Vin that wants to drive the transistor, so we cannot put Vin attached to the drain of the MOSFET, but to the source. Again, the choice of the MOSFET depends on the sign of Vin. If Vin is positive, current must flow from left to right, so the first two configurations are fine, vice versa if Vin < 0.

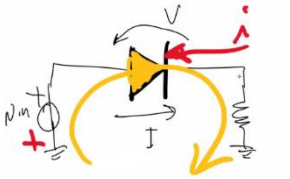


SIGNAL RECTIFICATION

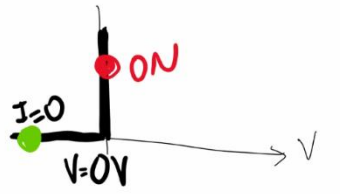
We have a signal and we have a load where we want to apply only the positive (or negative) part of the signal. This action is called rectification; in the time domain we can introduce the diode, which is a rectifying component. Thanks to its i/o characteristic that is an exponential curve that increases around 0.6V and we have a negative reverse current below 0. At RT the thermal voltage is 25 mV.



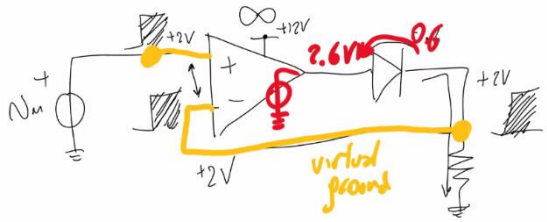
When Vin is positive through the network, trying to impose a clockwise current, this current can pass because it is a forward current, the diode is turned on and it has a 0.6V across it. If the negative voltage is applied, no current will flow.



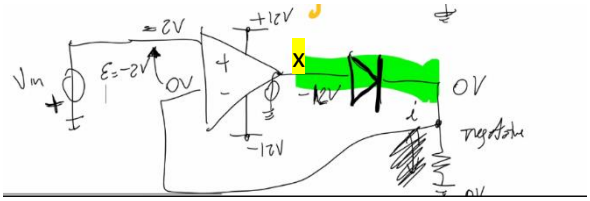
In the exponential amplifier, we considered Vin small enough to have the diode always on and to move around it slightly around the bias point.
Let's simplify the conditions; when the diode is off no current flows, whereas when it is on it has a voltage of 0.6V across it. Only if the signal is sufficiently high it can pass through the diode, otherwise if it is like mV, it won't turn the diode on.



So we need to define a diode that is a shortcircuit when V is positive and that has 0 current in the off condition. To do so, we don't want to use a diode directly, because it requires too much signal to be turned on. So we can use an opamp that will drive the load, to reach a 0 difference in voltage between the input terminals. To have only positive currents in the load, we introduce a diode in the loop (it is driven by the opamp) and in series with the load. It is the superdiode configuration.

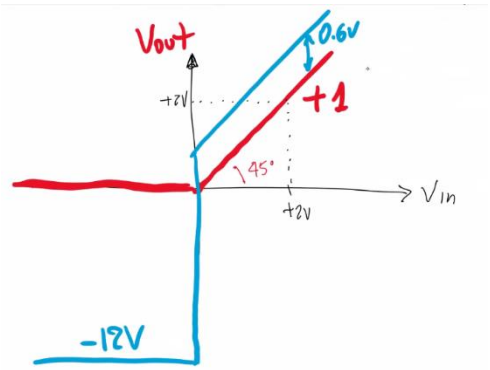


The 0.6V to the diode are provided by the opamp. If Vin is negative, the current cannot flow this time, because the diode won't turn on, the opamp won't drink the current in → the diode behaves like an open circuit, so the voltage on the load will be 0, and the opamp tries to go as low as he can, but since there is no current, in the – terminal we have 0, so the epsilon in input to the opamp is -2V.



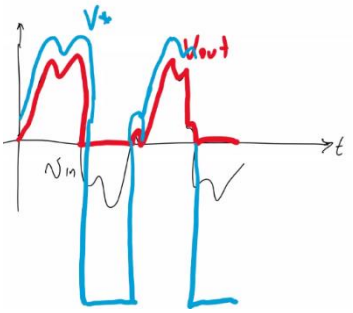
We have no more a negative feedback, because the diode is open and so there is no virtual ground and no loop.

Input-output characteristic

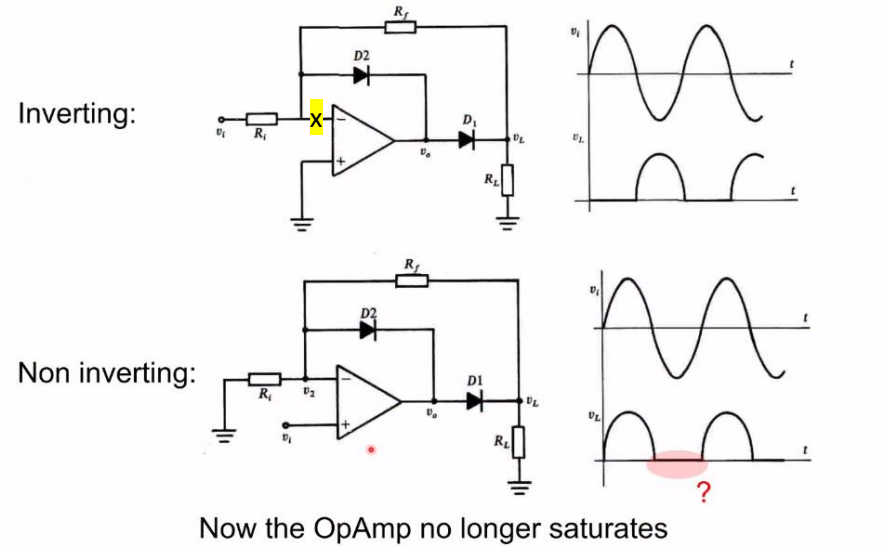


The slope if the diode is on is +1, whereas when Vin is negative, Vout is 0. Let V* be the output of the amplifier at node x. It is the blue curve. With negative input, it saturates to the negative biasing voltage of the amplifier.

In time domain, if I have a certain signal, the signal in output will be the positive Vin, nothing if the signal is negative for Vout. Conversely, V* will be 0.6V above Vout. As soon Vin is negative, V* collapses down to the negative power supply. This is a bad behaviour because I don't see it in output but the amplifier does it. We have to improve this because we don't want to have a saturation behaviour.



IMPROVED VERSION OF THE SUPERDIODE



The two configurations are the same. In the inverting one we enter on the – terminal branch and the other is at ground, while in the non inverting the – terminal branch is grounded and we enter from the + terminal.

Inverting configuration

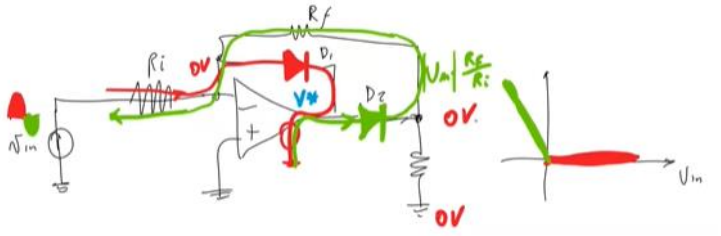
When Vin goes positive, node x (circuit above) goes positive and if so the opamp has a negative output. If so, D2 is in forward bias and becomes a short. Conversely, D1 has negative voltage on the left and ground on the right, so it is reversed biased, so open circuit.

For every positive input signal D2 is on and D1 is off. But if so, the buffer that we have created acts as a voltage source. In reality, a diode is not a perfect switch, because it requires 0.6V across it, so we should model the diode with a battery of 0.6V battery to be taken into account.

When D2 is on, there is something (the created buffer) that brings something back to the input, so we have a negative feedback. So there is Vg at node x. Hence if we apply a Vin, we have VG and we generate a current on Rin that will flow maybe in the Rf resistance. But D1 is open and the series of Rf and Rl has VG and ground on their sides. So no current could theoretically flow through the two resistors.

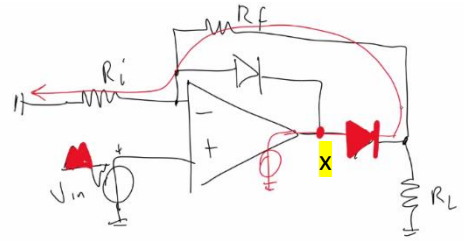
Instead, when Vin < 0, D2 is off and D1 (one side of D1 is at very positive voltage, the other is attached to a resistor, Rl, to ground) is on because the output of the amplifier is very much positive. So we have a current through D1 that goes in Rl and maybe in Rf. If current flows in Rf, the circuit has negative feedback. Hence the VG concept applies and at – terminal we have VG. If so, D2 is off, Rf is in between VG and a positive value → we have a current in Rf. Hence the current Vin/Ri will flow in Rf. But then this current will also flow in Rl. If so, being Vin a negative value, the output is an amplified positive value.

So for positive Vin the overall output of the stage is 0, while for negative Vin the output is the opposite version of Vin and eventually amplified if Rf > Ri (eventually equal if Rf = Ri or attenuated if Rf < Ri).



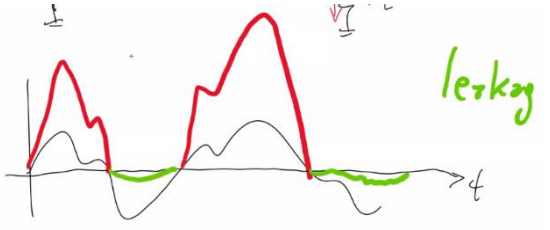
Non inverting configuration

The same diode positioning is in the non-inverting configuration, but now the input is different. If Vin is positive, + is positive. The – is still at 0, because I have applied to + the signal → output of the opamp is very much positive. If Vo goes positive, since – is at 0, D2 is reverse bias and open. Most probably, D1 is instead high positive on the left and on the right side positive. If so, v_1 is high and positive and so if we look at the feedback through Rf we have a negative feedback and VG applies. So we have the – terminal equal to Vin.



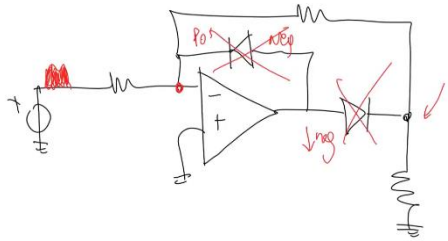
So when the input is positive, the output is positive and it gets amplified by 1 + Rf/Ri, whereas when Vin is negative, Vo is negative, and so D1 is off, it is reverse bias; - terminal was maybe at 0, so D2 is forward bias and it is on, and we have a buffer. So – terminal is now equal to Vin because we have the feedback and nobody touches Rl. But the output is not 0, because Rf and Rl are in series and we are

moving the - terminal copying the + termina input signal because of the feedback. In this case, the output goes to Vin * Rl/(Rl + Rf).
So for positive signals the output gets positive and it is amplified, while for negative signals we have a certain leakage and we get an attenuated version of the input, not 0.



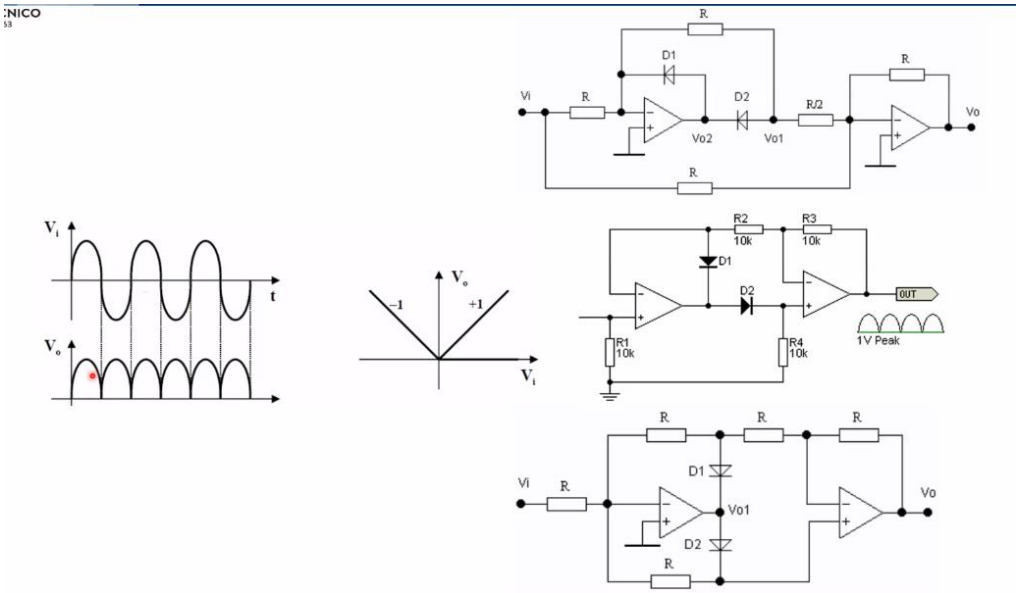
With positive signals at node x, V* = Vout + 0.6 V.
With inverting configuration, V* = -0.6V for positive signals, whereas it is Vout + 0.6V for negative signals.

The first circuit has to be preferred because it kills the signal when not useful, but it is inverting → we might change the orientations of the diodes in order to have not only the negative parts of the signal amplified but also the positive part and the negative suppressed instead. For example, we can invert D2. If input is positive, on the left side of D2 we have a positive signal, and on the right side a negative one. So D2 is off. If the output of the opamp is negative, nobody is touching the right side of D1 and so D1 is off as well. Hence for positive signals the circuit (left) is with both diodes off → there is not a negative feedback, we have a series of Ri, Rf and Rl.

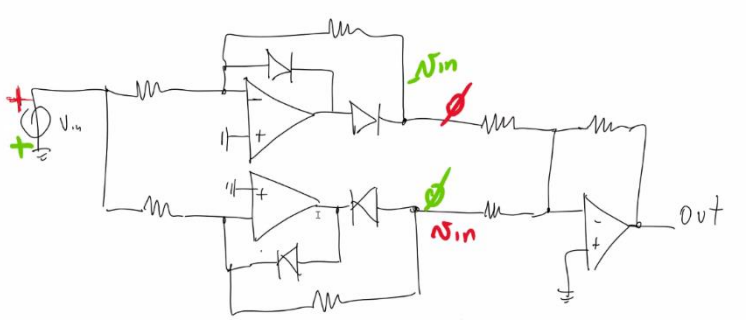


For negative signals, D2 is on and also D1 is on. Hence thanks to virtual ground the – terminal is at 0. This is a bad circuit because the output is slightly 0.

SUPER DOUBLE RECTIFIER



We want both the positive and negative signals to pass through and be inverted. To do so, we take the previous inverting circuit and in the other we invert D1 and D2, then we sum their outputs, we get a double rectifier. With positive Vin the upper circuit gets 0 in output and the bottom one Vin, the opposite if Vin is negative. The gain will be Rf/Ri and hopefully we select the two Rf and Ri of the top and bottom branches to be the same. The problem is that we are using 3 opamps, and a similar solution can be found with just 2 opamps and 2 diodes (configurations in the previous image).

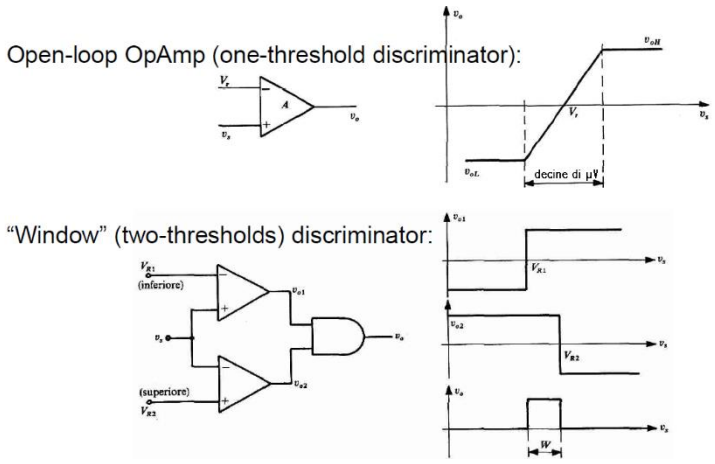


The three configurations differ one from the other but the good thing, for instance, of the second configuration, is that the input is high impedance. In the first circuit the input stage is R/2 (two R parallel to ground. In the third circuit the input impedance is R.

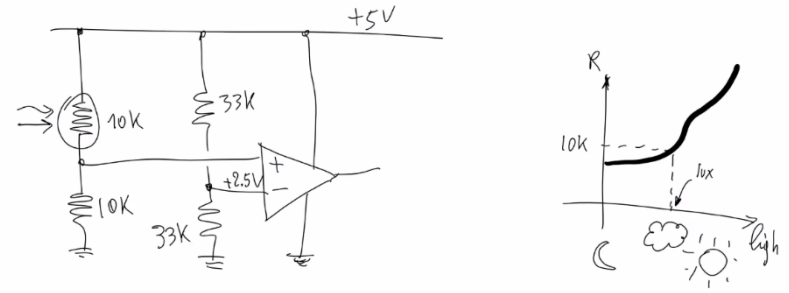
We introduced these configurations with two diodes to avoid saturation towards negative power supply, and now we have 3 circuits where this saturation is not occurring.

COMPARATOR

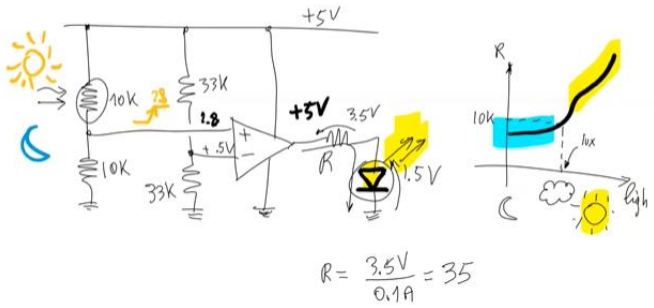
It is an opamp with no feedback at all or a positive feedback. If we don't have a feedback, if the – pin (Vref) is for instance higher than the + terminal, the output will be low to the lower power supply of the opamp. There is of course a transition region but if the gain of the amplifier is huge, the transition region is small, it depends on the swing I want (that is the swing of the PS) divided by the gain of the opamp.



We use a comparator in the case where we have a sensor, e.g. a photoresistor whose value changes with light. The photoresistor value is about 10 kOhm, and we use a comparator to get the output saturated to the positive or negative power supply depending on the reference voltage value on the – terminal. If the threshold is e.g. related to 10 kOhm, I can place a 10 kOhm resistance in series and the crossing point is 2.5V from the voltage partition, obtained with a resistive partition with the 33k resistances.
If the sensor's resistance increases, the voltage will decrease compared to 2.5V, so the – terminal will win and the output will saturate at 0.



If we want the LED to turn off when the weather is sunny, so the sensor's resistance increases, we connect the LED as below. Typically, to be on, the LED requires 1.5V across it in forward bias.

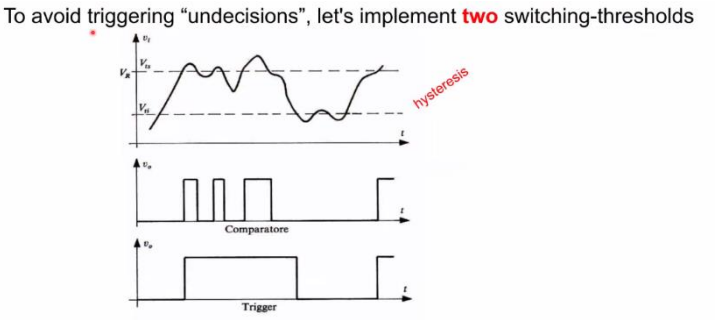


To invert the behaviour of the circuit, we simply can change the position of the input terminals of the opamp, or we could change the positioning of the input sensor, swapping it with the resistor. If in output we have a led with a resistor, we can also invert the biasing of the led.
There is a problem, that we could have minor fluctuations on the input signal that cause the output to display some undesired commutations because the threshold is crossed every other time.

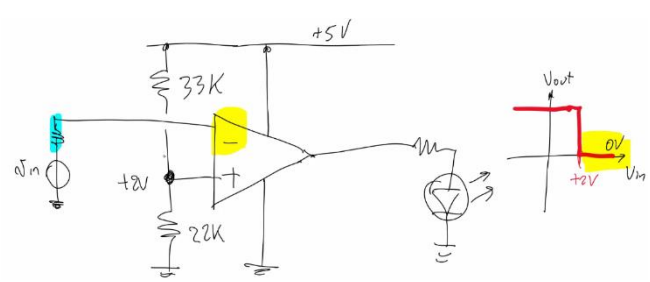
In order to solve this problem, we can introduce hysteresis.

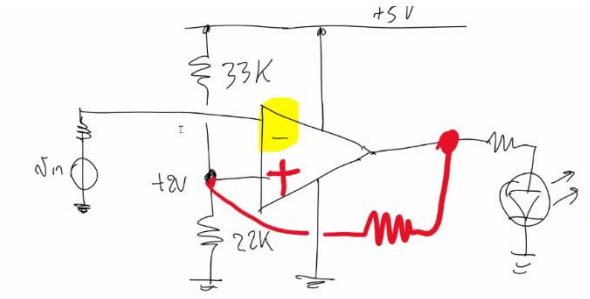
SCHMITT TRIGGER

We implement two different thresholds and we use one or the other as we like. Only when the signal goes higher the higher threshold the output commutes. Conversely, it goes down only when the input is smaller than the lower threshold.

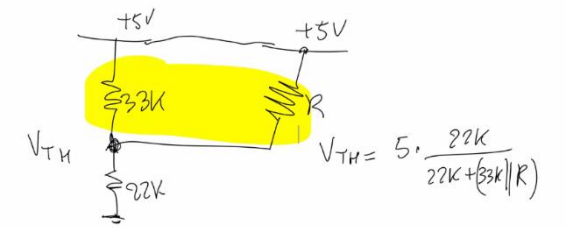


This can be achieved very easily. A possible implementation is the one below. Let's consider a Vin fed to the negative input of the opamp, and the threshold applied to the positive input. On the left we have the classical comparator with a threshold of 2V. On the right we have the circuit with two thresholds.
As soon as the opamp is triggered, I want to change the threshold → I need to introduce a positive feedback connection. In this case, if Vin is very small, e.g. 0.3V, the threshold wins and the output saturates to +5V.





Now the threshold voltage is no longer 2V, but higher and Vth-high is given by the following network (right case).

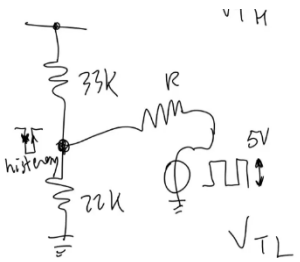


Vth_high is as in the formula in the image. When Vin is very high, e.g. +4V, the minus wins and the output goes negative and hence to 0 (negative power supply). So, thanks to the feedback due to the resistor R, the threshold of the stage reduces, because now the R resistance is not connected to +3V but to 0V (thanks to the opamp) in output. We get that the equivalent circuit is the following now.

Now we have a brand new threshold, Vtl. We can also quantify the hysteresis that is **delta_H = Vth – Vtl**.

However, the hysteresis can be computed also in another way. In the network, we have the series of 33k and 22k and in the middle we attach a resistor R that is driven by the opamp, which can be either 0 or 5V. So the opamp swings its output of 5V.

The corresponding swing that we have in the node where the R is attached is the hysteresis.

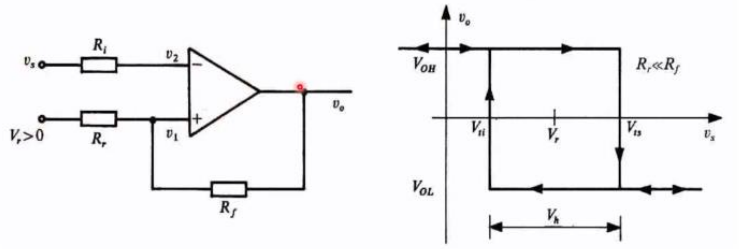


$$\Delta_H = V_{TH} - V_{TL} = 5 \cdot \frac{22K \parallel 33K}{(22K \parallel 33K) + R}$$

<u>The higher the R the smaller the difference between the two thresholds and hence also the hysteresis. For R that tends to inf, hysteresis is 0</u>.

Inverting Schmitt trigger

When the OpAmp output swings, the threshold changes to opposite direction, therefore let's exploit POSITIVE feedback !



Usually let's choose $R_f >> R_r$ so ...

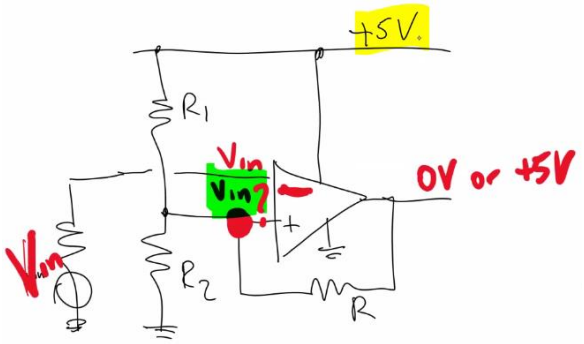
Upper threshold: $V_{ts} \cong V_r + \frac{V_h}{2}$

Lower threshold: $V_{tl} \cong V_r - \frac{V_h}{2}$

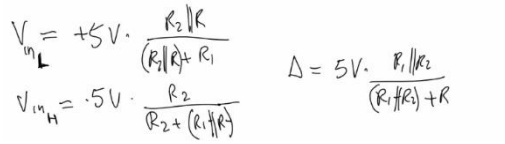
Hysteresis: $V_h \cong (V_{OH} - V_{OL}) \frac{R_r}{R_f}$

Let's compute the thresholds. We have a positive feedback configuration. Since the feedback is positive, the output is either 5V or 0V (limits of the power supply). Now let's compute the value of Vin that causes Vout to commute, hence the threshold. If the signal enters at the minus, it is an inverting Schmitt trigger. If we cross Vtl, the output goes high, while if we cross Vth, the output goes low.

To compute the threshold in the case of Vtl, let's consider to have the same Vin at the negative and positive inputs of the amplifier (it is not an effect of the feedback, but we are making the assumption because we want to find the threshold).
The voltage at red node is Vtl = 5V * (R2||R)/(R2||R + R1).

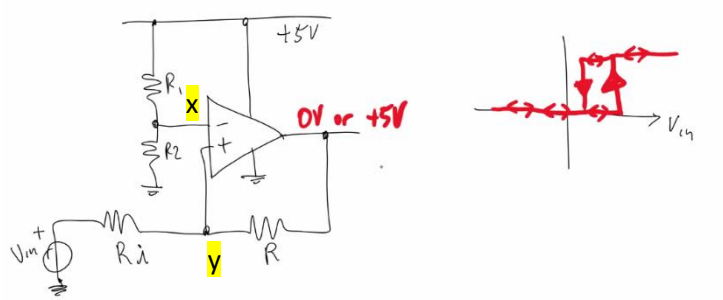


Conversely, when the input is high, the high threshold is:
Vth = 5V * R2/(R2 + R1||R)



Non-inverting Schmitt trigger

Let's compute the i/o characterstic. When Vin increases, the + terminal goes positive, so the output will go positive soon or later. The output can only have two values, 0V or 5V.

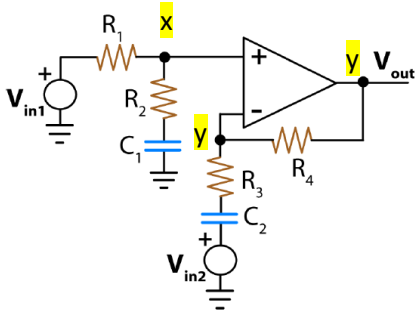


The computation of the threshild is not so easy. Voltage at node x is 5V * R2/(R1+R2). What is the value of Vin to have Vy slighlty higher than Vx? If so, the output of the opamp commutes.
In order to compute the input threshold, let's consider when Vy = Vx. In this configuration, the output is either 0V or 5V. The current that flows in Ri must also flow in R, because no current enters in the + terminal. This current equation over the resistance R1 and R is:

$$\frac{V_{in} - V_x}{R_i} = \frac{V_x - V_{out}}{R}$$

With Vout equal either to 0V or 5V. For Vout = 0V we get Vtl, for Vout = 5V we get Vth.

Example 1



R_1=20kΩ R_4=10kΩ R_2=30kΩ C_1=200nF R_3=1kΩ C_2=1nF
A_0=10^6 I_B=10nA f_0=10Hz

a) Plot the ideal gain $V_{OUT}(f)/V_{IN1}(f)$

b) Plot the ideal gain $V_{OUT}(f)/V_{IN2}(f)$

Request a)

We have to consider the frequecny domain. Vin2 is grounded. To compute the gain, we have an RC towards ground (x) that causes a zero for sure, and then we have an amplifier stage with gain 1 + R4/R3. In DC, all the capacitors are open, so Vin is dierctly attached to the + terminal of the amplifier. Then the amplifier is a buffer because there si no resistor to ground, R3 is not to ground. So the DC gain is +1.

The RC net at node x causes a zero that is given by:

$$zero_1 = \frac{1}{2\pi \cdot 30k \cdot 200n} = 26Hz$$

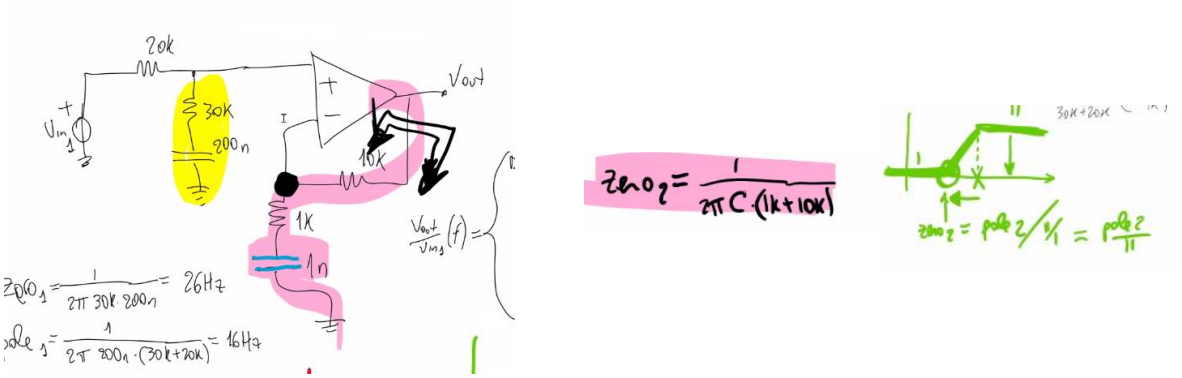
$$pole_1 = \frac{1}{2\pi \cdot 200n \cdot (30k+20k)} = 16Hz$$

Then there is also a pole given by the total resistance that is the R1 in series with R2.
Also the other capacitor has a pole, where the total resistance is given just by R3, because the impedance of nodes y is 0.

$$pole_2 = \frac{1}{2\pi \cdot 1n \cdot 1k} = 160kHz$$

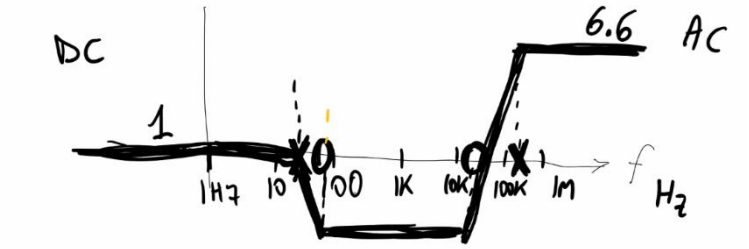
But does C2 introduces a zero? If the capacitor is open, the gain is 1, whereas if it is shortened, the gain of the amplifier stage is something. Hence it has for sure a zero to switch from a small gain to the higher gain.
To compute the gain we use the GBWP. Zero = pole/gain_high.

From an empirical point of view, the zero of this stage can be computed as the C2 multiplied by the series of resistances looked from the output, because if I look from the capacitor they are not in series, but if I look at the feedback from the output, they are.



The HF gain is R2/(R2 + R1)*(1 + R2/R1).

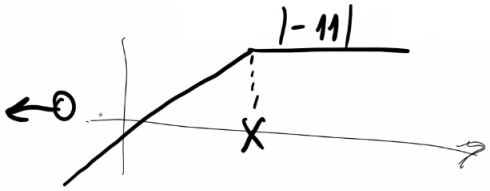
Bode plot:



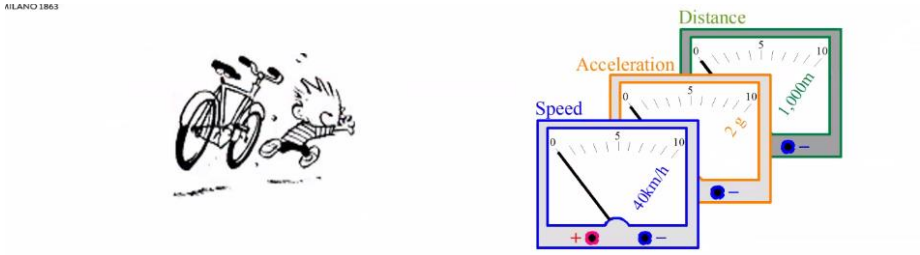
So in the circuit at middle frequencies we have attenuation. It is true because if we look at the circuit, in DC both the capacitors are open, but then at middle frequencies the attenaution is caused by C1 that have an RC circuit that enters into action before the other capacitor → attenuation of C1 is before amplification due to C2. This sizing of the circuit is not ideal, because we have attenaution in between, but we wanto to achieve amplificaiton. If it was in the opposite case, with C2 acting before C1, I would have had a gain of 11 in the middle frequencies, and 6,6 gain at HF.

Request b)
Now Vin1 is in shortcircuit. We are in an inverting configuration, I don't touch the portion of the circuit that has the RC branch.
In DC, G = 0, whereas at HF, G = -R4/R3. We have a zero at 0Hz and a pole at p ? 1/(2*pi*R3C2) = 160 kHz.

Bode diagram:



Example 2



Calvin measures his bike's speed S, by employing Hobbes' speedometer ($v_s = S \cdot 50^{mV}/_{m/s} + 200mV$) and 10V FSR voltmeters

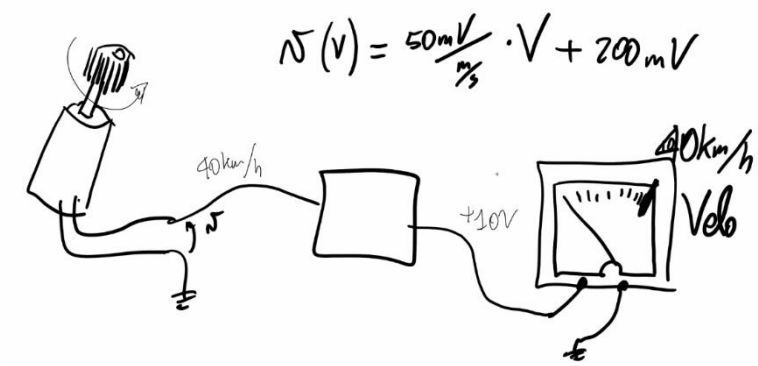
a) Design a circuit for displaying the speed, up to 40km/h, with 3s smoothing

b) Display acceleration/deceleration, up to 2 g ($g = 9.81 m/s^2$)

c) Measure and display the ride distance, up to 1km

Request a)

So the dynamo gives a constant voltage and I want to see it on my voltmeter and see the speed. The white box is the circuit I have to design. 200 mV is a constant offset that I have.



So I need an amplifier with a 18 gain. I choose an inverting amplifier, i colud use a 17 kOhm and 1 kOhm resistances. But if the input is 0, the output is not 0 due to the DC offset of 200 mV. How can I subtract it?

Let's go to speed = 0, I will have 200 mV in input that corresponds to 3.6V output in DC.

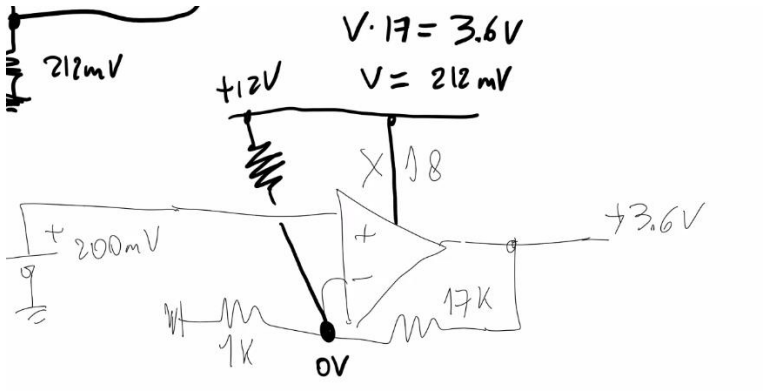
With 200 mV in input I want the output of the amplification stage to be 0V. To remove the 3.6V due to the 200 mV, I could take node x and put it to a specific voltage, not grounded, so that its value per 17 is equal to 3.6V → 212 mV.

$N(40km/h) = 50mV \cdot \frac{40 \cdot 1000}{3600} + 200mV = 0.555V$

$G = \frac{10V}{0.555V} = 18$

But using a battery to remove the DC value is not useful → I create this voltage from the power supply with a voltage partition. Thanks to this, Vout is 0 when there is no speed.

Another possibility is that we have the amplifying stage with two resistors of 17 and 1 kOhm because we want a gain of 18, but if we have the 200 mV, to remove the 3.6V in output, I could take the PS and use it to inject current at the – terminal.



The current flows in the R resistance, so I choose this resistance to have a voltage drop across it of 3.6V. In the case of the image:

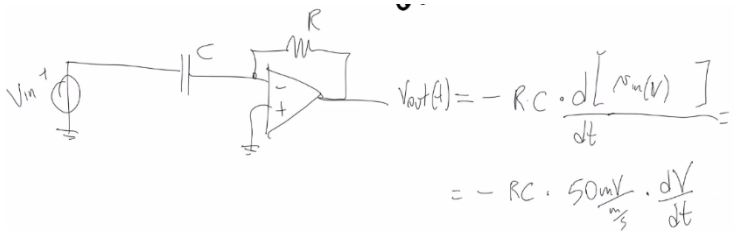
$$V_{out} = -3.6V = +12 \cdot \left(-\frac{17k}{R}\right)$$

$$R = 12V \cdot \frac{17k}{3.6} = 56.66\ \Omega$$

In this way the output is zero for a 200 mV offset input signal.

Request b)

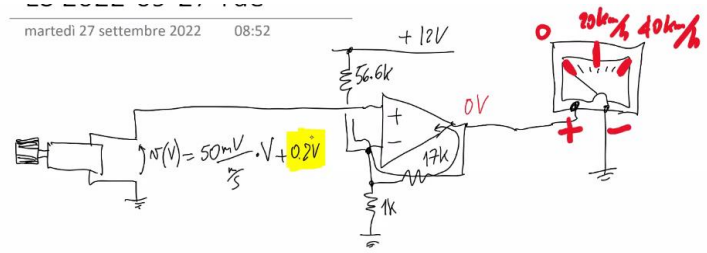
The dynamo produces a velocity, and we know that the acceleration is the derivative of the speed. So to perform the derivation we use the derivator circuit. So if I consider Vin and I derivate it, I get the information related to the acceleration.



Let's now size the components. If a = 2g = 2*9.81, I want Vout = 10V.
Now we substitute back the values 10V = -RC * 50mV*2*9.81m/s → RC = 10.2 s

Hence the circuit correctly converts the dynamo in a voltage if we choose RC = 10s (we select the value of R and C as we want).

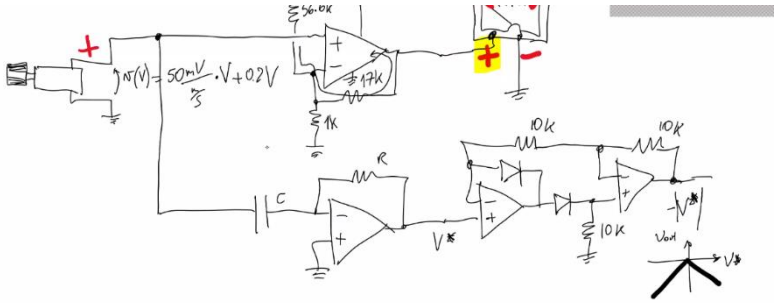
Let's now draw the full schematic of the first and second stages of requests a and b. The signal form the dynamo is proportional to the velocity; in the first stage we used a non inverting configuration, and the DC offset is killed by injecting a current in virtual ground from the power supply.



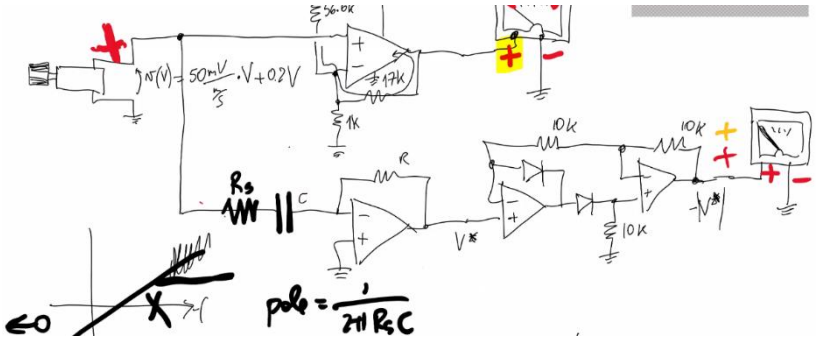
Negative power supply can be ground, while the positive one has to be 12V because we want to reach 10V in output.

As for the second stage, it has to be a derivator, in particular a real derivator. Indeed, Vout is a function or -RC*dVin(t)/dt. When we reach the maximal acceleration of 2g, the output should be equal to 10V FSR. Since we have a minus, this means that when the input is positive, the output is negative → we can simply connect the voltmeter driving the negative input, while the positive one is to ground. For the previous stage it was the opposite.

However, this circuit is capable of measuring only positive accelerations, not also negative, because when we have no acceleration the output is 0, and then if we have a deceleration we cannot go below 0. So we need a double rectifier.
The double rectifier circuit is the one at the end of the bottom branch. Negative signals and negative ones are amplified by 1 and the output is the inverted version of V*.



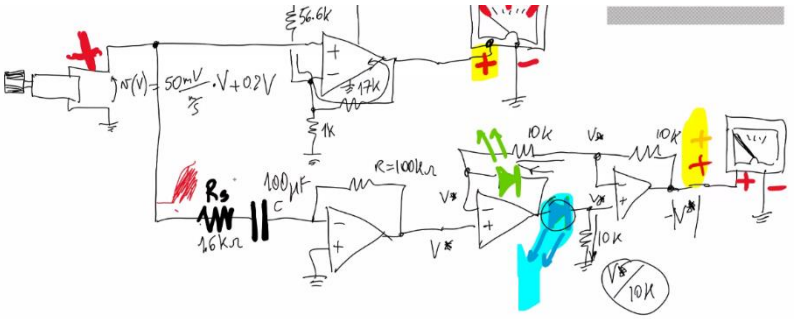
Given V*, the output is -|V*|. Since when the acceleration is negative and V* is negative, the output is positive, while in the opposite case it is negative → no need of inversion in the voltmeter.
The real derivator is needed because if we have noise at HF, the output is incredibly high (saturated), so we need a resistor to prevent this problem and have a finite gain at HF.



We can choose the pole depending on the position when I don't want anymore the derivative action to be done, e.g. every time I simply press the pedal I don't want the derivative, whereas I want it when I'm accelerating over some seconds → I can operate for example up to 1Hz, so I set the pole at 1Hz. If C = 100 uF, Rs = 16 kOhm.

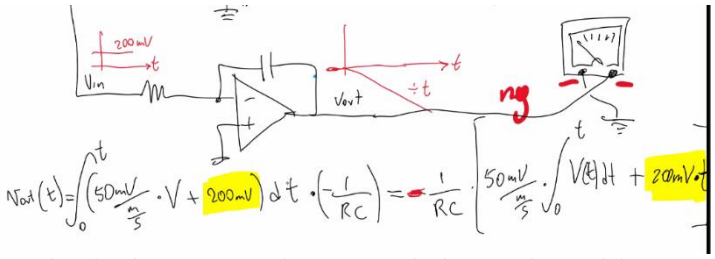
We need a double rectifier to rectify the signal because the input signal is either positive or negative, so instead of two diodes we can use two LEDs in place of the diode, in this way the two activate alternatively when we have acceleration or deceleration. The LED activates very bright or not?
If the signal increases at the input, V* is negative, so the green diode turns on because the output goes negative. The current flowing through the diode is V*/10 kOhm because in the upper branch we don't have current → if I want the LED to be much brighter I need to reduce the resistor to provide more current.

As for the LED in blue, I have to reduce the resistors in the upper branch, because they regulate the current.



The opamp must be biased between -12 and 12 V (the one of the derivator), whereas the one of the rectifier one (the first) between +- 12V and the other between +12V. Since it is all a mess, it is better to bias them all between +- 12 V.

Request c)
The distance is the integral of velocity, so I need an integrator stage.



If I have a constant input in the integrator, the output is increasing with a ramp, it is proportional to minus time (because it is inverting). Moreover, the output is negative, so we need to plug it into the – input of the voltmeter.

So we need to get rid of the DC signal. Moreover, I want to get 10V when I've done 1km.

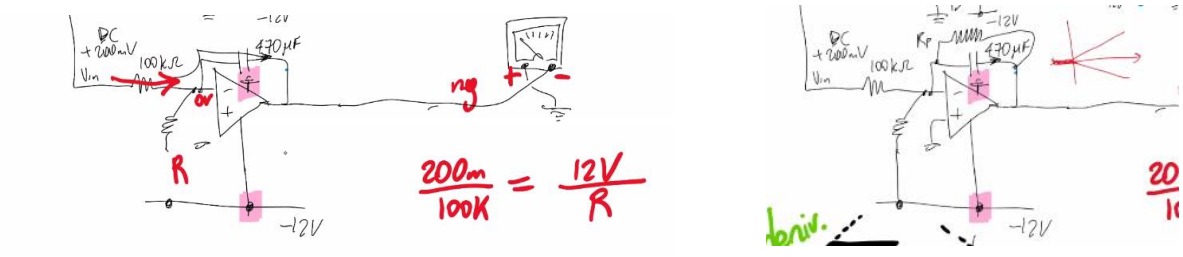
$$10V = V_{out}|_{max} = -\frac{1}{RC} \cdot \frac{50mV}{m/s} \cdot 10\,km \cdot 1000 =$$

$$RC = \frac{50mV \cdot 10 \cdot 1K \cdot m}{m/s \cdot 10V} = \frac{500}{10} = 50s$$

Let's choose a very big capacitor, e.g. 40 uF and a 100 kOhm resistance. But if we do so, the problem is that after 1km the voltage displayed is not 10V but a little bit small → so we sell the device at lower price, while if I want precision I size the components correctly.

Now I want to avoid that if the input is constant, the output is not a ramp, but a constant 0. So I need to remove the current generated by the constant Vin, so I add a resistor. – terminal is at 0V, so the resistor should be biased with a negative voltage. Also the opamp should be biased with negative values and the positive ones with ground.

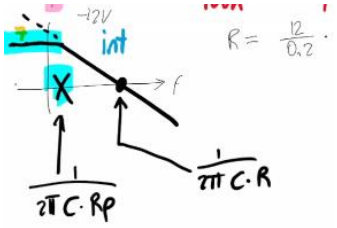
To choose the value of R, we know that the input resistor has 200mV applied to it, whereas the R experiences 12V.



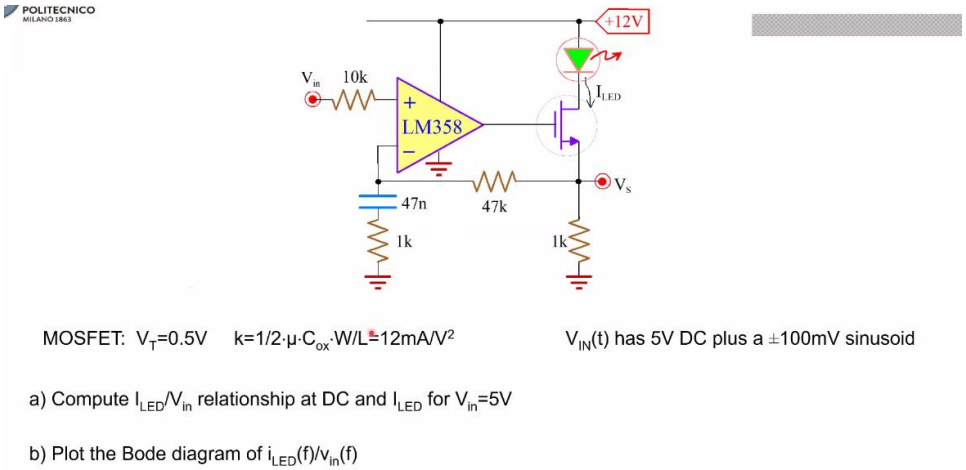
R = 6.4 MOhm. If I choose the most close value to the value I need, there still could be mismatches due to tolerances → there will be a current moving in the fieedback and if there is a DC mismatch, the output will ramp up or it will ramp down. So the voltmeter in output still will be increasing.

So the integrator, to decrease the bad gain in DC, adds a pole to limit the DC gain → we add a parallel resistance in feedback, Rp. Thanks to Rp the integrator has now a pole that saturates the gain. Thus, even if we have a DC signal we don't have amplification. So the resistor R steals away the DC component, whereas the feedback resistance decreases the gain. We need both the two configurations to limit the DC effect.

So for the last branch of the circuit we have the following Bode Plot.



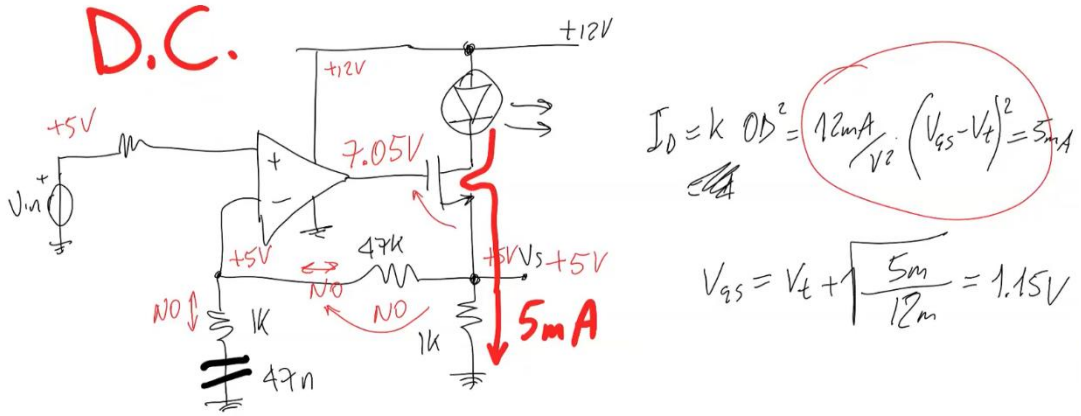
Example 3



Resolution

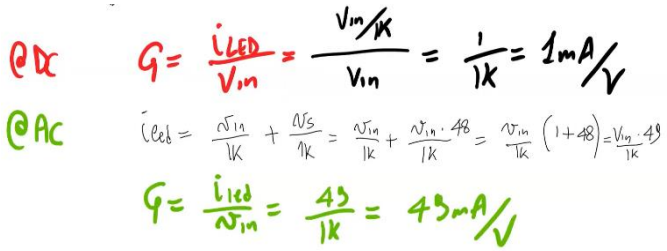
If I apply positive Vin, the output goes positive. Then we have a nMos so if we increase the gate we will create a channel and a current will flow and pump increasing Vs going to ground towards the 1k resistance and the 47k and 1k series and voltage at – terminal increases.
So the feedback is negative. If Gloop is very high, then the virtual ground concept applies, so the Vin applied on the + terminal will be also applied to the – terminal.

At DC, we find Vin also on the – terminal, and the C is open, so we find Vs = Vin. If we have this, we must have a current on the 1k resistor (5mA), that must be provided by the transistor and hence must be drank from the drain. Since a current flows through the LED, the LED turns on, not because I'm applying a voltage of 12V but because I'm applying a current through the transistor that is driven by the opamp.

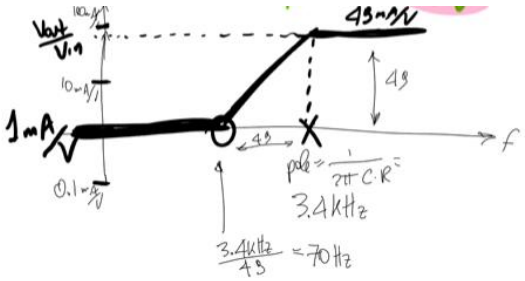


Of course, the opamp will be biased so that it can be provided a proper V*. to compute V*, we know the current through the drain that is k*Vov^2, and I want it to be 5mA. So we can compute Vgs.
In the end Vgs = 1.15V and V* = 5 + 1.15 = 6.15V.
This value of V* is possible because the opamp is biased at 12V so it is able to provide this value of voltage.

Now we can compute the gains at DC and AC. In DC I find Vin at node Vs, because the + terminal Vin is copied on the – terminal and the capacitor is an open, so no current flows. The gain we get is a transconductance gain, because we have a voltage to current converter. The input is a voltage and the output is a current.
In AC, the capacitor is shorted, so we have Vin across the 1k resistance on the left, which generates a current. So we can say that Vs = (Vin/1k)*(1k + 47k). Hence we also know the voltage on the right 1k resistance and we can compute the current through it.

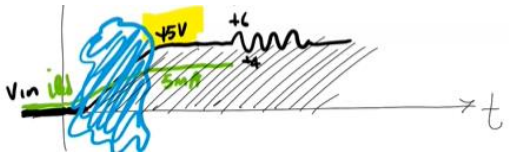


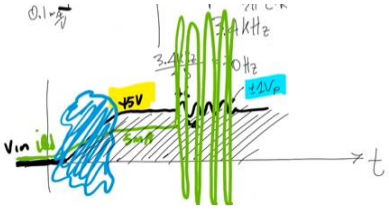
Let's now draw the Bode plot. The gain now has a size, we are not computing Vout/Vin and hence getting an adimensional value. Since the AC gain is higher than the DC gain, we have a zero and a pole.
The Req of the pole is just the1k resistance on the left, because in computing the pole the input signal is 0, and if so also the – terminal is at 0V. The pole results at 3.4 kHz
As for the zero, since I know that the ratio between AC and DC gain is 49, also the ratio between pole and zero must be 49, hence the zero is at 3.4 kHz/49 = 70 Hz.



Let's now see in the time domain what happens.

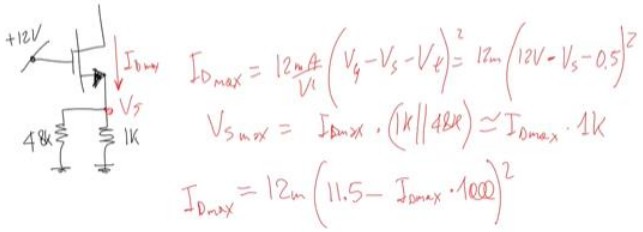
If the input is 0, the output is 0 (0*1 = 0), so no current in the LED (i_led = 0). If now Vin goes to 5V, i_led will go to 5mA. In the time domain, shifting between these two values we have a transient, but we will discuss about it later.
If now we apply a small signal over the DC value, e.g. 1Vpp AC, the gain is 49 V/A, so we should multiply this value for 49.



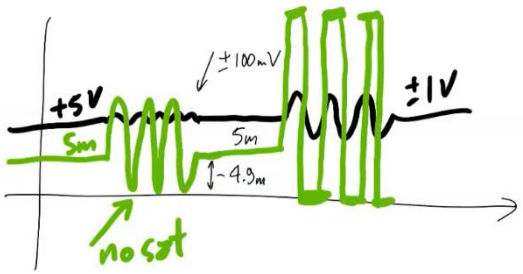


Hence in output we get 49 mA peak to peak superimposed to the 5mA DC.
Theoretically this reasoning is correct, but practically it's wrong. Firstly, we cannot have negative currents, because it should go in the direction down to up in the MSOFET, and to have so, the drain must be negative and the output of the opamp should hence be negative. However, this cannot happen because the PS of the opamp is 12V – 0V. Hence we have a saturation at the bottom, we have no negative portion of i_led. Moreover, the signal cannot reach about 50mA. If the i_led was 49mA + 5mA = 54mA, because the voltage Vs should be 54V. This cannot be feasible again for PS limitations. Hence we have also a upper-side limitation due to the upper part of the PS. The maximum current we can have at the output is when Vg is the maximal possible value, that is 12V.

To compute the maximal current that can flow, I have to write the equation of the network.



Maybe in the end we find that Id_max = 25mA.
We can now redraw the time diagram. The limitation is provided by the opamp that cannot provide higher voltages than the PS and by the k of the MOSFET that requires a high Vgs to provide a high current.

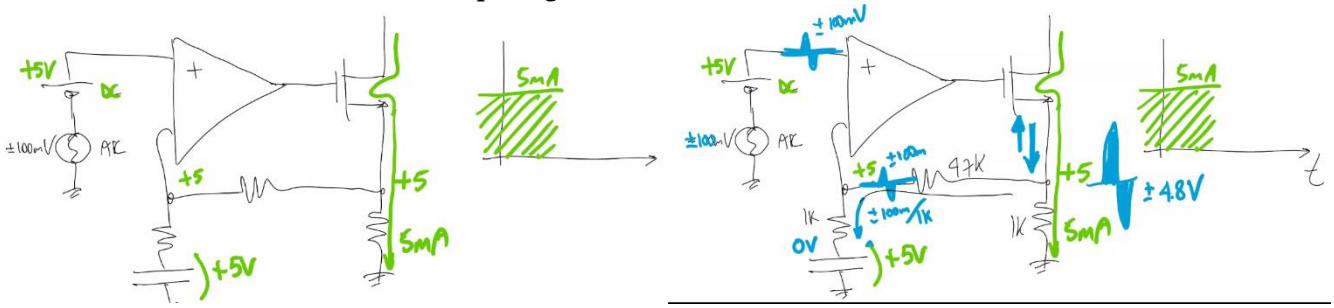


So this circuit is a voltage to current converter with a low gain at DC and a high gain in HF.

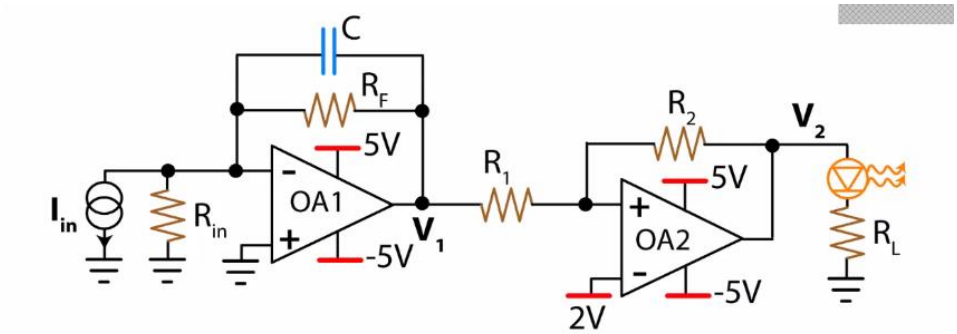
Let's revise the concepts of AC and DC. Let's consider the circuit and an AC signal in input with a DC value of 5V. At DC, no current across the capacitor so it charges to 5V DC, and we have 5mA DC current through the transistor.
Now let's consider AC that is on top of the DC (+-100mV), it is not around zero, but around the DC value. The capacitor is now a shortcircuit because it cannot charge or discharge, so it cannot change its voltage. Which voltage? The voltage it has in DC that is 5V. So the capacitor won't change and it will be like 0V in terms of DC (but in reality it will be at 5V), so I see a current that flows through the resistors

(1k and 47k). In output we will see a fluctuation higher than in input (48 times higher than the input). This fluctuation causes a current that can be in one of the two possible directions, up or down in the mosfet. This current is like seeing 48k| | 1k at the output node. So for the signal the current can be positive or negative (I'm considering fluctuations at the top of the DC value).

The current can be either positive or negative but on the top of the DC value that was of 5mA, so the real actual signal will be always positive, it is 5mA in DC, it reaches 9.8 mA at top and 0.2 mA at the bottom. If I further increase or decrease the input signal, I reach a saturation.



Example 4



R_{in}=100kΩ R_F=10kΩ C=10μF R_1=2kΩ R_2=20kΩ R_L=1kΩ

a) Compute the effect of I_B=100nA and V_{OS}=3mV of OA1 on V1

b) Plot the static curve V_1 vs. V_2

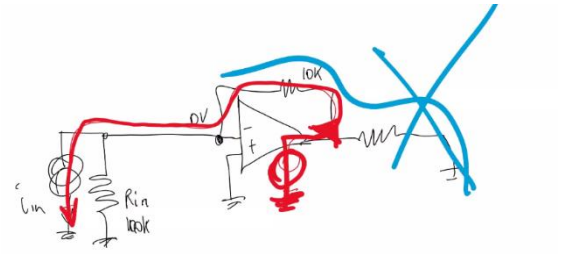
c) Compute the minimum amplitude I_{in} (20Hz sinusoidal) to switch on the LED

Resolution

I always need to check if the loop is negative. If the OA1 – terminal increases, the output decreases, so the left end of Rf decreases with respect to my increasing perturbation → It is a negative feedback and the concept of virtual ground applies. If so, Rin is making no effect because no current will flow through it and Iin will go all to virtual ground.

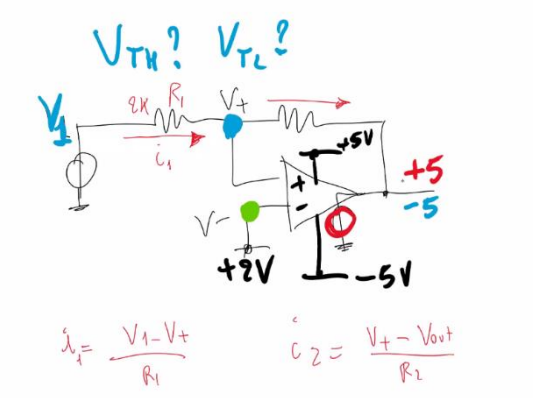
Then I have OA2, and we have a positive feedback, so it is a Schmitt trigger. So the two terminals are one fixed, and the other one moving, whereas in the negative feedback they move together. So the output is either + or – 5V.

The first stage is an integrator, so at DC we have Vout = -Iin*Rf, while at HF Vout = 0, so the gain is 0. Both in DC and AC the current is not coming from the right branch, but the current comes from the output generator of OA1. Then if I have an output resistor also the generator will provide a current, but it is not the Iin current goes as in blue.

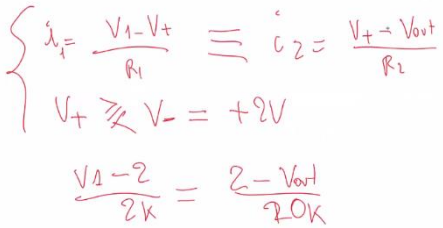


So the Bode plot of this stage is:

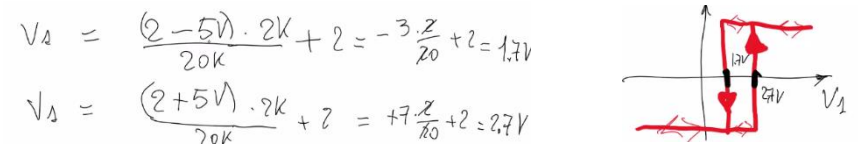
Once we have V1, let's study the second stage. The OA2 is a voltage source, so I don't care about the output because the opamp does what it wants to do plus taking care also of the output.
We have a Schmitt trigger (non-inverting because I enter in the +) biased between +- 5V, so the output will be one of these two values. Now I want to compute the values of Vtl and Vth, that cause the comparator to trigger.
I want the condition when V+ is equal to V- or when i1 = i2.



Then I want V+ = V- = +2V.

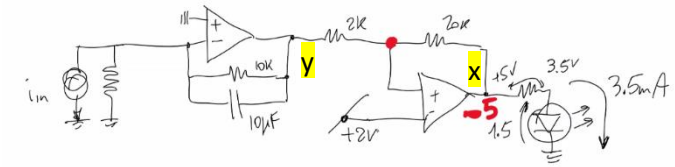


We will find two values of V1 depending on having either + or – 5V in output.



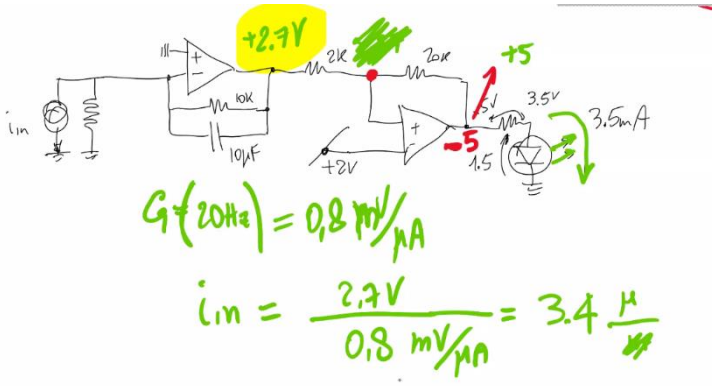
When the output is high, the output LED emits light, otherwise it is off. The one in red is the static curve V1 vs V2.

Request c)



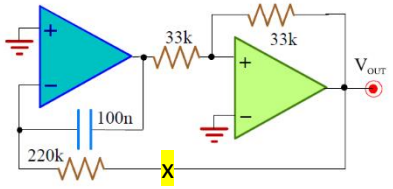
Usually the LED requires 1.5V, so the voltage across the resistor in output will be 3.5V, so the current across the LED is 3.5mA (when on) because Rl = 1k.
When will the LED turn on? It will be in the off condition before, so in the off condition we have -5V at node x and if so, the voltage across the diode will be very negative. I know that the threshold to turn the diode on is the high threshold of 2.7V, so we will experience the commutation between -5 to +5V only when the voltage at node y becomes higher than 2.7V. This happens depending on the gain, which is frequency dependent. The pole is at 1.6Hz (C and Rf).

We want the minimum amplitude at 20Hz, with the pole at 1.6Hz. so the gain is given by the GBWP and it is: G(20Hz) = (10mv/uA, that is the DC gain) / (20/1.6) = 0.8 mV/uA.
So in input I will have iin = 2.7V/(0.8mV/uA) = 3.4 mA.



So iin must reach 3.4mA at 20Hz to have the LED on. If the frequency was higher, I would have needed a higher value of iin.

Example 5



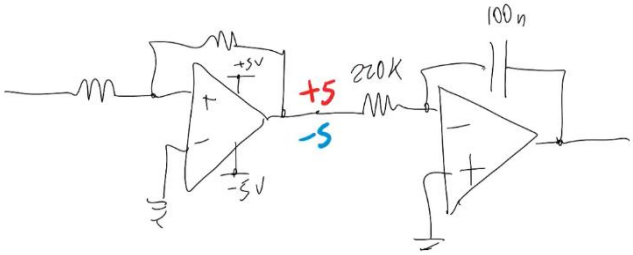
Rail-to-rail OpAmp biased at ±5V. At t=0s, the capacitor is discharged.

a) Plot all waveforms

b) Write the equation of output frequency vs. R and compute the value for R=220kΩ

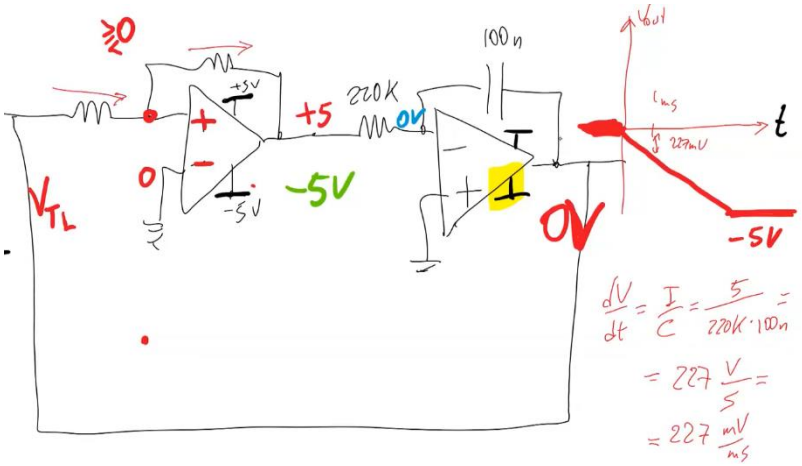
There is no input in this circuit. There is a first opamp, then a negative feedback through a capacitor, so maybe the first opamp is an integrator. In fact, if we apply an input at node x, the first opamp behaves as an integrator. Then the second stage is not an amplifier because the feedback is positive, so its is a Schmitt trigger, whose output will be saturated to the lowest or highest power supply. Moreover, it's a rail to rail opamp, meaning that the microelectronics of the opamp is in a way that the lower and higher PS can be reached. Furthermore, since the first stage is an integrator, at – terminal of the opamp we have VG; this integrator is fed with either + or – 5V from the trigger.

Let's redraw the circuit.

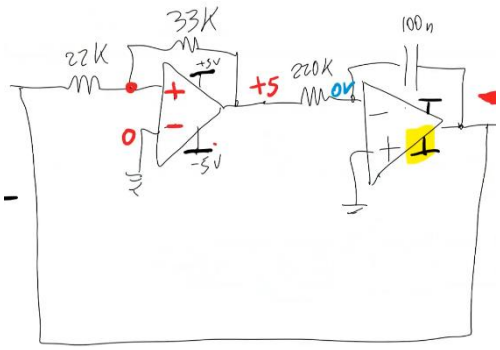


Let's start with the input +5V, I don't know which is the output of the trigger at the beginning. If we have +5V, since the voltage at the – terminal of the integrator is 0V due to the negative feedback, it means that when we have +5V there will be a constant current provided by the opamp equal to +5/220k. This current must flow through the capacitor because the opamp drinks it because of the negative feedback. I can then write the equation into the capacitor I = C*dV/dt, and dV/dt will be a constant number, so the V will be a ramp. Vout will start from 0 at the power on and then will go down after it with a slope I/C = 5/(220k*100n) = 227 V/s.
The output goes negative and eventually it will saturate to the power supply, but no, because the output feeds back the input of the circuit. The first trigger in fact monitors the output voltage because it's a Schmitt trigger, so we need to compute at which voltage the commutation occurs.
The trigger will have 5V in output and the input will go so negative that the + terminal of the trigger, which at the beginning was positive (we started with 0V at the output of the integrator), will reach a value less than 0 and if so the epsilon in input to the trigger will go negative and soon or later I will trigger the commutation of the output to -5V. let's compute the threshold value at which the trigger commutates.

The threshold can be easily found by assuming the + terminal equal to 0 and the output is +5V. this happens if Vtl causes a current in the feedback resistance that is equal to the current in the input resistance.



We write the balance of currents (Vtl-0)/33k = (0-5)/33k → Vtl = -5V.
Hence this circuit causes a ramp at the output and when the ramp reaches -5V the + terminal of the Schmitt trigger reaches 0 and the trigger probably commutes. However, I'm not sure this commutation happens because it is not so likely that the output reaches -5V, so I can slightly modify the circuit, for example by changing the resistors.

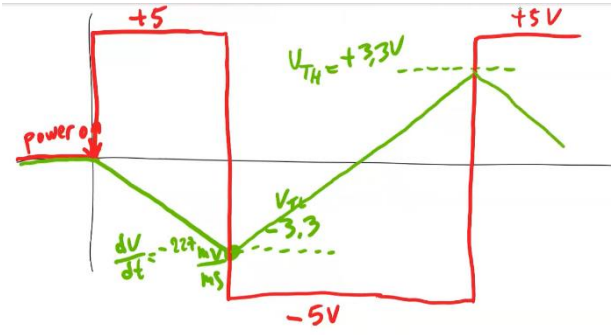


In this case Vtl = -3.3V.

When the output ramp reaches -3.3V, then the trigger commutes and the output moves suddenly from +5V to -5V. If so, we will have a current in the opposite direction chargin the capacitor and so we will have an increasing

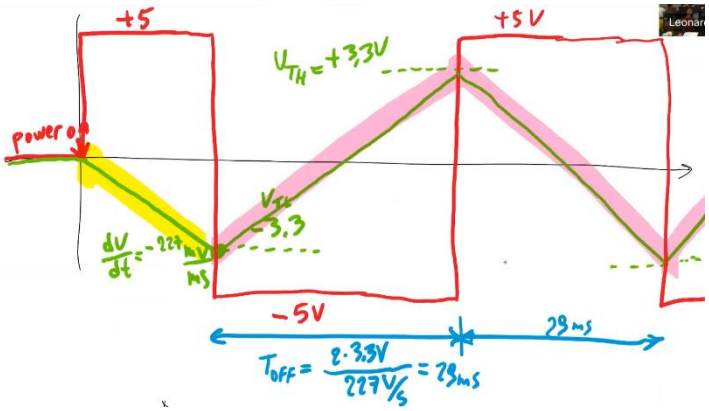
ramp. The current is still with the same slope, but with the opposite sign. We will move from -3.3V up to eventually +5V. But before the saturation at the output we have the commutation of the trigger. As soon the + terminal of the trigger goes slightly above zero we commutate the output.

To sum up, when we start we don't know where the capacitor will be after power on. Let's suppose the output of the trigger is positive. If so, then the integrator starts integrating charge, and I will have a decreasing slope on the capacitor. Eventually, when I will reach Vtl, we have commutation, and from +5V the trigger saturates to -5V. If so, the current reverts and the capacitor charges. Soon or later we will reach the new threshold Vth, that will be equal to +3.3V (because of symmetry).



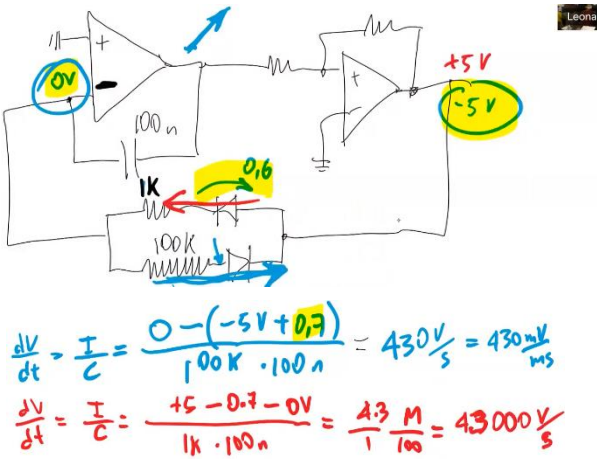
I can conclude that we have a bistable oscillator that keeps running.

NB: only the first transition is short because the signal moves from 0V to -3.3V, the other ones are larger. We can compute the period of oscillation, and since the two commutations are symmetrical, I can compute the time from dV/dt = I/C → t = dV/(I/C).

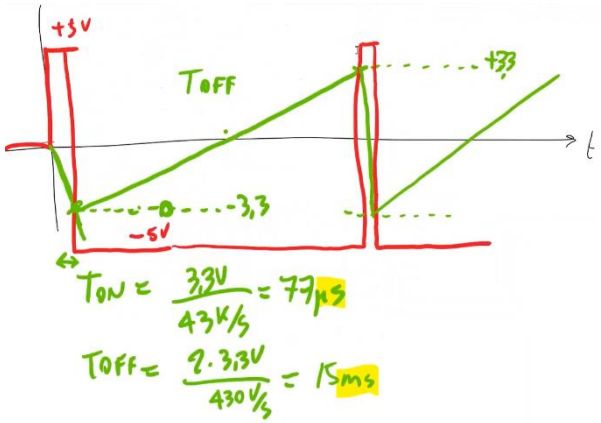


I can conclude that T = Ton + Toff = 58 ms.
Hence this circuit oscillates with a f = 1/T = 1/58 = 17 Hz.

If we exit at the output node after the trigger we have a square wave oscillator, while if we exit after the integrator we have a triangular wave.
However, sometimes we may want to change the duration of Ton with respect to Toff and this can be done in different ways. Knowing that the slope depends on I and C, I then depends on R, to change the duty cycle DC = Tou/Toff we can design a new circuit.
We need for sure an integrator and then a Schmitt trigger. To change the duty cycle I can change the charging current I with two resistors. One for charging and the other for discharging.
When the output will be +5V or -5V the current will flow in one path or the other. We have always to remember that we have still a voltage drop on the diode (e.g. 0.7V).

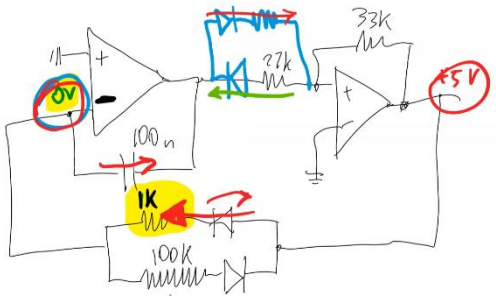


We can see that the transition with +5V has a much steeper slope simply because the resistor associated to that path is much smaller, 1 kOhm. So what happens is the following (thresholds remain the same because they are related to the resistive network of the trigger).

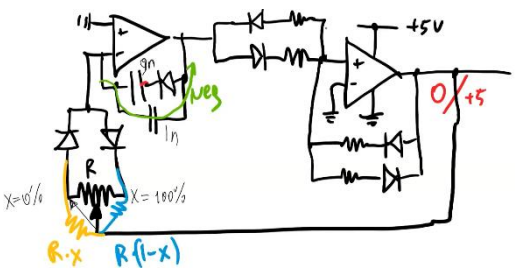


DC = 77us/15ms = 0.5%

Another way to change the circuit is by changing also the thresholds. So let's introduce another branch with two different resistors and we use one resistor or the other thanks to diodes. When the output is positive the current will flow in the upper branch, if negative in the lower branch. Thus the two thresholds will differ.



In general term, we may have a circuit that has thresholds in two different points to have a different duty cycle. We can also change Vout,high and Vout,low by changing the biasing of the trigger.
We could also use two capacitors in parallel (always with diodes). Eventually, we can also have variable resistances that changes according to a potentiometer.

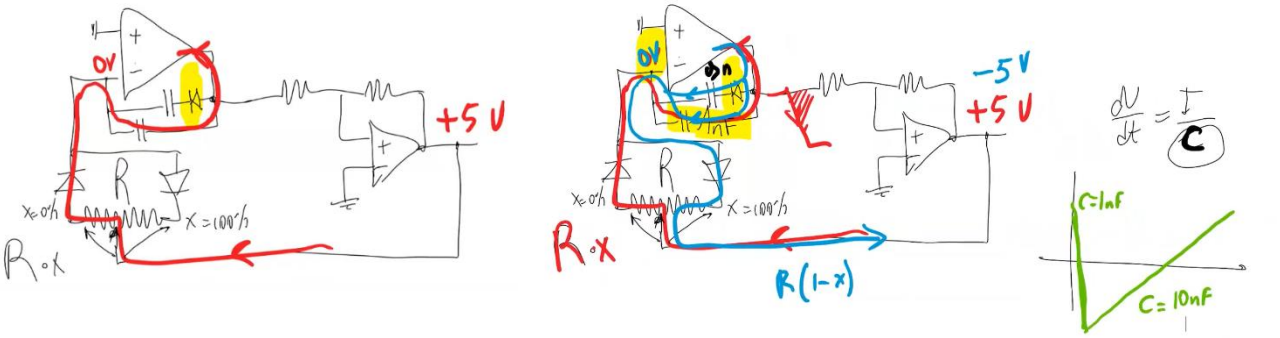


In this case, since the sum of the two resistors doesn't change, the oscillation frequency of the circuit is constant

(apart from the first short commutation). But we can change the duty cycle by acting on the potentiometer.

NB: if the input resistance of the trigger was higher than the feedback resistance, the circuit would never oscillate, because to bring the + terminal below 0 it would require a voltage in input smaller than the PS limit (same reasoning for the Vth).

In the previous circuit, the capacitors are placed like that because, when the Schmitt trigger commutes, e.g. to +5V, we have a current in the potentiometer. Once the current flow in that direction, it means that it should also flow in the capacitor, and only through the bottom capacitor because the diode stops the current. Vice versa, when the output commutes to -5V, current should follow the blue path in the opposite direction and through the parallel of the two capacitors, and the slope is slower, and so we can change the DC.



FREQUENCY COMPENSATION

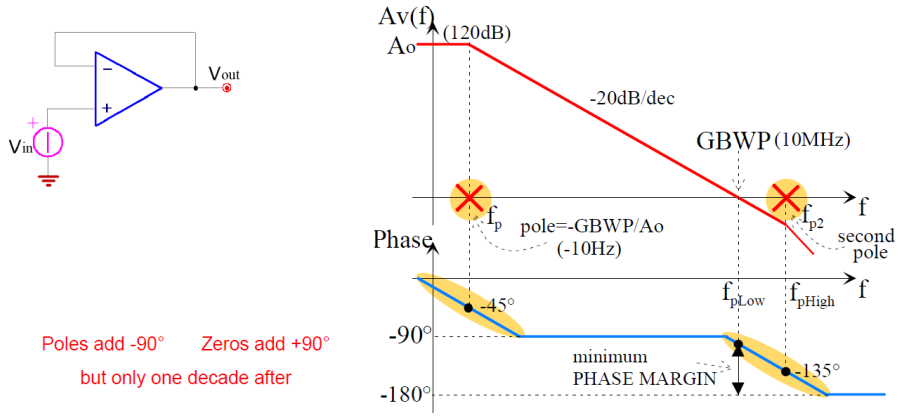
We want to understand if the negative feedback is strong enough.

INSTABILITY

Let's consider a loop we set to drive a car. We want to drive at a specific velocity, we look at the speedmeter, we know the speed we want to reach. The difference between what we want to reach and the actual speed determines our action on the pedal. The pedal drives the engine, the engine drives the traction wheel and the speedmeter is connected to the wheel. If there is no delay, we reach the target speed soon. But if there is a delay in our system, we are in an unstable condition.

OPEN LOOP FREQUENCY RESPONSE

The OpAmp introduces amplification, but also phase shift!



We have an opamp connected as a buffer. So far we considered the opamp as an amplifier with a very high gain in DC (120 dB), then a pole and then there will be a frequency where the gain of the opamp reduces to 1, the GBWP.

Let's imagine the connect the opamp that we have in OL. The opamp has a gain A(s) that is DC gain and when the frequency increases to much there is a pole inside the opamp that causes the gain to drop. Every time we have a pole, the pole causes a phase shift between the input and the output. At the frequency corresponding to the pole the accumulated phase shift is 45°. 1 decade before the pole there is no phase shift, one decade after it there is a full phase shift of 90°.

So if we look at the OL opamp, we have the phase shift in the upper right plot. A phase shift of -180° is very bad, because if we take the opamp and we want to introduce feedback, we may have a negative feedback in DC, but if we increase the frequency we run the risk that the output has a shift. If the shift is remarkable, to the input of the feedback circuit it returns the signal with a shift, so the epsilon is no longer 0 at the input terminals of the opamp and the feedback could become positive.

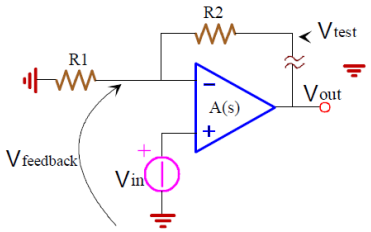
NON-INVERTING STAGE

We want to compute the Gloop, so the amplification that a signal epsilon at the input of the opamp experiences before coming back to the input. Given a feedback circuit, we should cut the loop, excite it with a V_test and compute the ration of the voltage that returns over V_test.

Once we have the circuit as below, in principle we could cut wherever we wish. However, the good position where to cut the loop is just after a voltage source, e.g. after the output of the opamp.

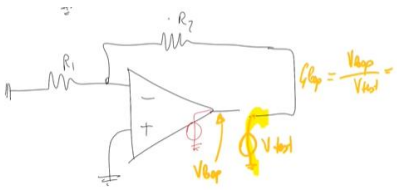
In the case of the non-inverting stage, we have V_test, then the partition R2 and R1, so we eventually have the V+ voltage. Once we have it we multiply by the gain of the amplifier A0 and we have reached Vloop. So we have Gloop.

G_{loop} must be assessed, in order to check **quality** and **stability** of feedback



$$\beta = \frac{v_{feedback}}{v_{test}} = \frac{R_1}{R_1 + R_2}$$

$$Gloop = -A(s) \cdot \frac{R_1}{R_1 + R_2} = -A \cdot \beta$$



Real closed-loop gain:

$$G_{NI} = \frac{A}{1 - A \cdot \beta} = \begin{cases} 1/\beta \quad \text{if} \quad 1/\beta << A \quad (\text{i.e. if } G_{loop} >> 1) \\ A \quad \text{if} \quad 1/\beta >> A \quad (\text{i.e. if } G_{loop} << 1) \end{cases}$$

:Ideal gain

We could have decided to cut in other positions, like below.

This position is ok because we cut in a position where the impedance is infinite on the right, so we can leave it unconnected. Moreover, now I can drive the high impedance on the right with a V_test and see what returns. Gloop is negative because we pass through the – terminal of the opamp.

The real gain of the non-inverting stage is the ideal gain divided by 1 – Gloop. If Gloop >> 1 we are left with Gloop and what remains is 1/beta. If Gloop is very high, the ideal gain is 1/beta, which means that A >> 1/beta. If we have a Gloop poor, so much smaller than 1, than the gain is A, so it acts as if it was no feedback.

Gloop ASSESSMENT

In order to assess **G_{loop}**, to check feedback **quality** and stage **stability**, and to be able to draw the **real closed-loop frequency response** and not just the ideal one, do follow these hints:

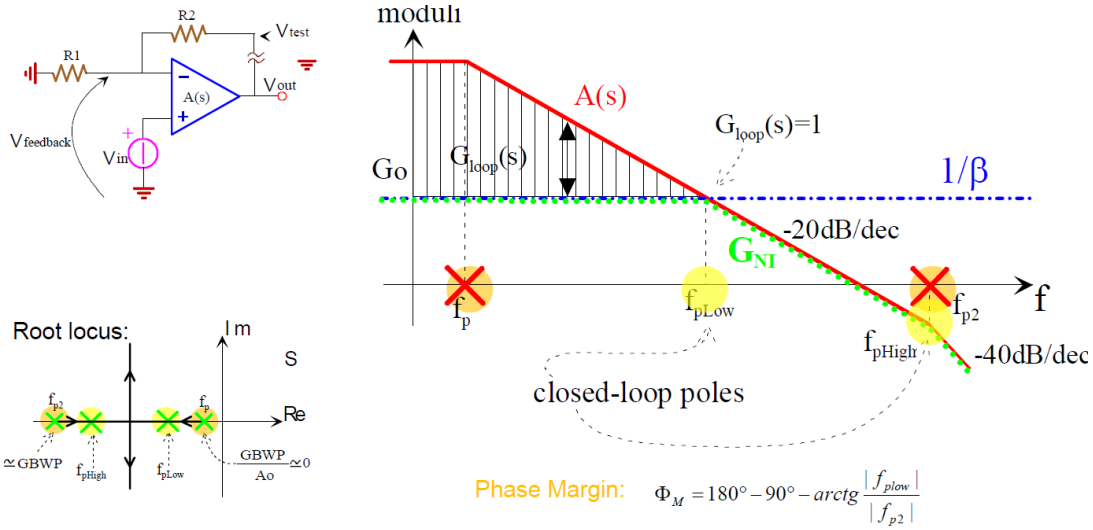
1. Draw A(s) — i.e. the frequency response of the OpAmp
2. Compute β(s) — NO Laplace, just asymptotic analysis at 0, medium and ∞ freq.s
   then draw 1/β(s) — remember: poles of β(s) lift up, zeroes of β(s) put down 1/β(s)
3. The split between A(s) and 1/β(s) is **$G_{loop}(s)$**;
   the larger the split, i.e. G_{loop}, the better the feedback **quality**
   instead, at frequency f*, where A(s) intercepts 1/β(s), $G_{loop}(f^*)=1$
3. Draw the expected ideal gain — don't care about A(s) nor 1/β(s)
4. The **real closed-loop frequency response**
   follows the ideal one when "*there is G_{loop}*", i.e. $G_{loop}(s)>1$,
   instead, beyond f* there is no more feedback, hence the
   real gain rolls off from the ideal trend, experiencing all following
   poles and zeros of A(s); "*the real gain dies as A(s) is dying*"
5. Stage **stability** depends on the "closure angle" between A(s) and 1/β(s) around f*
   stable when 20dB/dec before and after f*
   marginally stable when 20dB/dec before and 40dB/dec after f* or viceversa
   unstable when 40dB/dec or higher before and after f*

In our circuits we usually have two blocks; one is an amplifier A(s), and then a network beta(s). A(s) is defined by the datasheet, and we design beta, that is the attenuation and can also be frequency dependent.

We don't compute Gloop as A(s)*beta(s), because I want to assess if Gloop >> 1, so if A(s)*beta(s) >>1, and this happens if A(s) >> 1/beta(s). So we draw the Bode diagram of A(s), the one of 1/beta(s) and we check this condition. So we don't compute Laplace transforms but we simply plot the two.

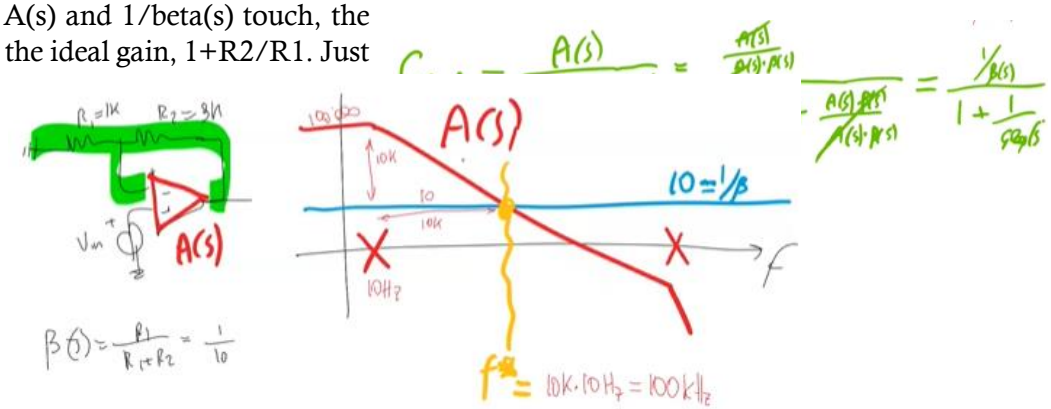
When A(s) = 1/beta(s), there Gloop = 1. Before it, Gloop is good, after it Gloop collapses.

Non-inverting stage computations



The ideal gain is 1+R2/R1. Let's now study the loop by cutting at the output and let's drive it with a voltage source and compute beta(s). We remove A(s) and see what returns at the input of the A(s) block, and then we draw the inverse of beta(s).
Since we are in a log-log plot, the ratio between A(s) and 1/beta(s), so A(s)*beta(s), is the distance between the two curves.

Up until f*, where A(s) and 1/beta(s) touch, the real gain is equal to the ideal gain, 1+R2/R1. Just



by chance it is equal to 1/beta(0) in this non-inverting configuration. Then after f*, Gloop dies and the real gain proceeds like the A(s) because the real gain dies by Gloop. In the formula we have the + at the denominator, but in reality it's a minus and Gloop is negative.

Ideally, there are no poles and so the gain is 10, but actually there are poles and they are inside the opamp, so after f* Gloop dies. Hence after f* I don't proceed like the ideal gain, but it depends on A(s) also. So since 1/beta is not dying and A(s) is, the maximum death rate is the one of A(s). The frequency at which the gain dies is at f*, and it's not dying due to the poles, but because the overall Gloop dies.

Stability of the circuit

I want to know if the stage is stable or not. In fact, if Gloop gets a too high phase shift, then the feedback is no longer negative and becomes positive. This happens if the phase shift accumulated is 180°. Hence I want to check the phase margin that is left before reaching 180° of phase shift.
Every time in the path there is a pole, the pole will introduce the full phase shift of -90° one decade after its frequency.

To do so, it's sufficient to look at the Bode plot of the modulus. If the closing angle between 1/beta(s) and A(s) is 20-20 dB/dec, it means that before the frequency of f* one pole acted in the loop, and that pole is the pole of the opamp in this case. But only one pole has happened and introduced 90° of phase shift. So the phase margin is still of 90°, so good.
Of course, then there will be another pole, but its phase shift will be limited ad HF. So we need to compute the phase margin with the formula of the image above.

Conversely, if the crossing is e.g. 20-40 dB/dec, this is a bad situation because the phase margin is maybe not high enough, the system is marginally stable (phm = 45°).

Finally, if the crossing is 40-40 dB/dec, the system is not stable anymore, the phase margin could be different from 0° but very close to it, and consequently very close to instability.

STABILITY ASSESSMENT

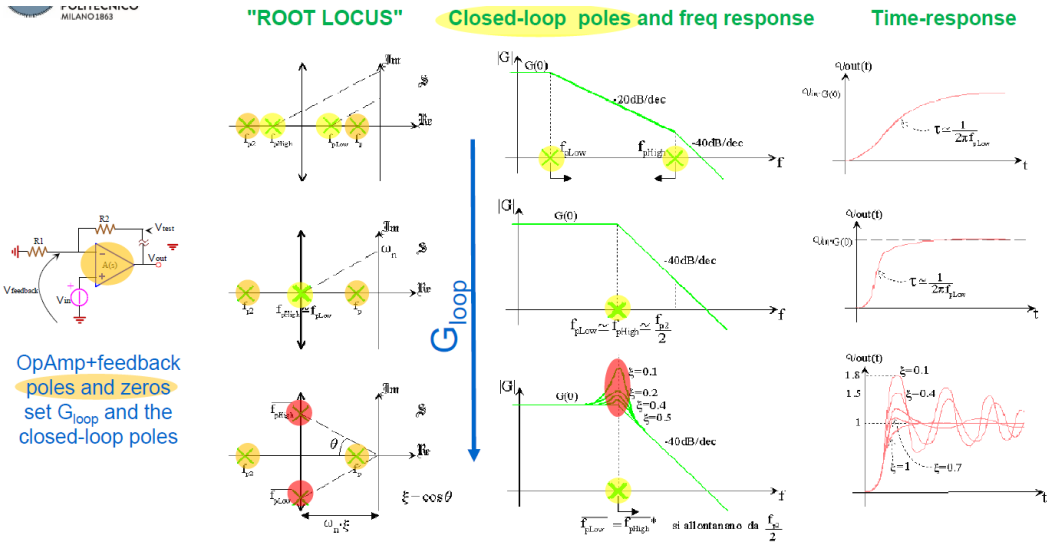
The reason why "closure angle" matters on **stability** is straighforward:

1. Given that the split between A(s) and 1/β(s) is **$G_{loop}(s)$**
2. The "closure angle" at f* measures the slope of $G_{loop}(s)$ versus frequency at f*
3. If before f* the G_{loop} experienced...

1 pole,	-20dB/dec slope
2 poles,	-40dB/dec
n poles,	n x -20dB/dec
2 poles 1 zero,	-20dB/dec
p poles z zeroes,	(p-z) x -20dB/dec
1 pole before f* and 1 pole exactly at f*,	-20dB/dec before f* but -40dB/dec after f*

4. Each pole (zero) adds a phase shift of -90° (+90°) to the feedback signal after one decade from it; instead the pole (zero) adds just -45° (+45°) if the f* is coincident with the pole (zero) itself
5. Therefore, by measuring the "closure angle" it is possible to infer the difference (p-z) and eventually the overall phase shift accumulated along the feedback path
6. The stage is **stable** if the feedback stays negative and does not accumulate -180° phase shift, in which case it turns to be positive and the stage is **unstable**

CLOSED LOOP FREQUENCY RESPONSE

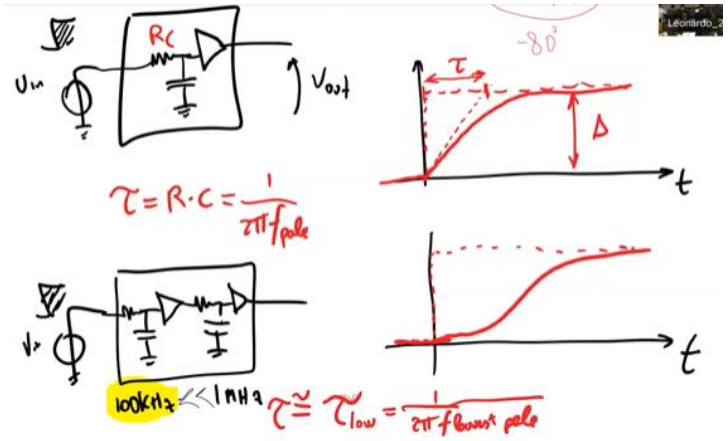


If we start from an opamp whose poles are the orange one, the root locus describes where the CL poles will appear. With the opamp open, I see just the orange poles. If instead I introduce feedback, then depending on the strength of Gloop, if Gloop is more or less 0, the poles of the loop will be equal to the poles of the OL opamp. If instead Gloop is different from 0 and gets stronger, the poles of the circuit move apart from the original poles and they move one close to the other (it happens also in the non-inverting configuration, see previous image).

If instead there are two poles and the Gloop is so strong that the crossing happens in a bad way, if we look at the root locus it means that we started with an opamp whose pole were separated but Gloop is so strong that the pole move across the imaginary axis. If the crossing is bad, 40-40, the system has to c.c. poles. They are so bad that the Bode diagram will see those two poles because it will collapse with -40 dB/dec and we have also peaking due to the instability of the circuit.

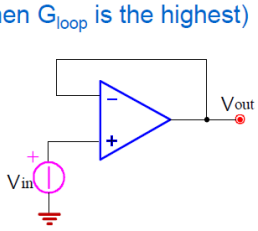
In the time domain, having a peaking it means that the output will have some resonance and it will settle after a certain time.

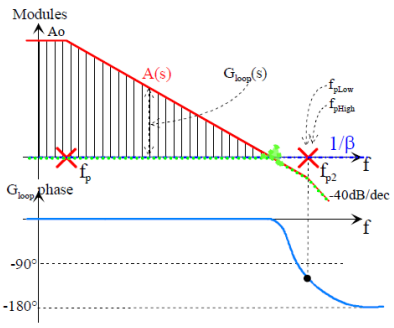
NB: every time we have a circuit that can be described with one pole, if we enter with a step at the input the output won't be an ideal step but an exponential curve. If we take the tau of this exponential, it is indeed the time constant of the pole.
If instead the circuit has two poles, if I enter with a step, the response won't be an ideal step but it will start flat and then have two exponentials curve. If one pole is much lower than the other, than the tau that dominates the transient is the one of the lowest pole and we are similar to a one pole response.



BUFFER CONFIGURATION

… the **BUFFER** (when G_{loop} is the highest)



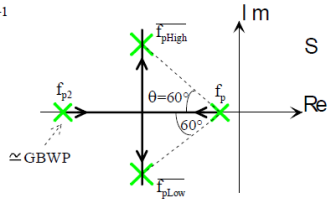


$$Gloop(s) = -\beta \cdot \frac{A_0}{\left(1+\frac{s}{GBWP}\right)\cdot\left(1+\frac{sA_0}{GBWP}\right)} = +1$$

When f_{p2}=GBWP...

$$s_{1,2} = -\frac{GBWP}{2} \pm j \cdot GBWP \cdot \frac{\sqrt{3}}{2}$$

... Phase Margin=45°



$$\xi = 0.5 = \cos\theta = \cos 60°$$

Margine di Fase = 45°≠ θ

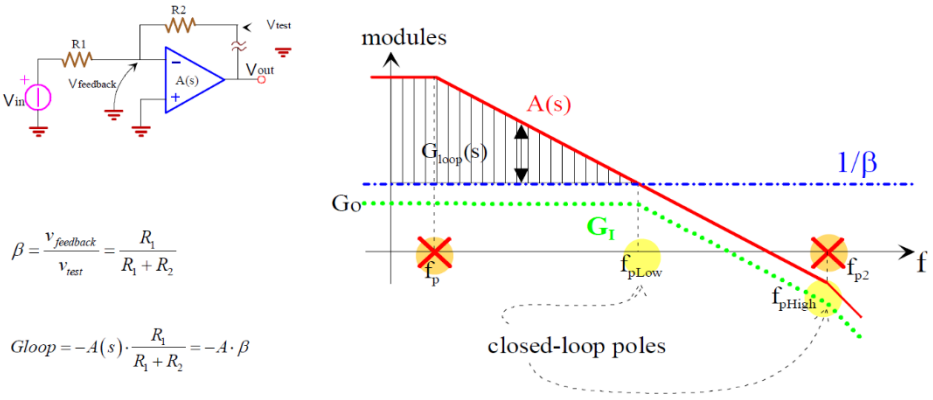
Some amplifiers are sold in order to be operated as a buffer, so the output is connected to the – terminal. The gain is 1. The opamp itself has an A(s), and the beta is 1 (as well as the 1/beta). If the second pole of the opamp is before GBWP, the system is instable because the crossing between A(s) and 1/beta(s) is with 40-40.

To avoid this problem, some manufacturer designed the amplifier with the first pole move backward to lower frequencies so that the second pole happens after the GBWP and the closure angle is 20-20. If there is no second pole it's even better, we are perfectly stable. If the second pole gets closer to GBWP we are still stable, so we reduce the frequency of the dominant pole so that the second pole is just after the GBWP, not far away from it or we reduce too much A0.

An opamp that has the second pole after the GBWP is called compensated opamp. If the second pole is before the GBWP the opamp is uncompensated.

If the manufacturer places the second pole just after the GBWP, the closure angle is no more 20-20, but it's almost 20-40. But if this is the case, it's still reasonable because the phase margin is still 45°, which means that the two poles are c.c. but the angle in the root locus is just 60°, so they are not that much c.c., we have just a small peaking, not perceptible. Bad c.c. that cause peaking have an angle in the root locus higher than 60°.

INVERTING STAGE

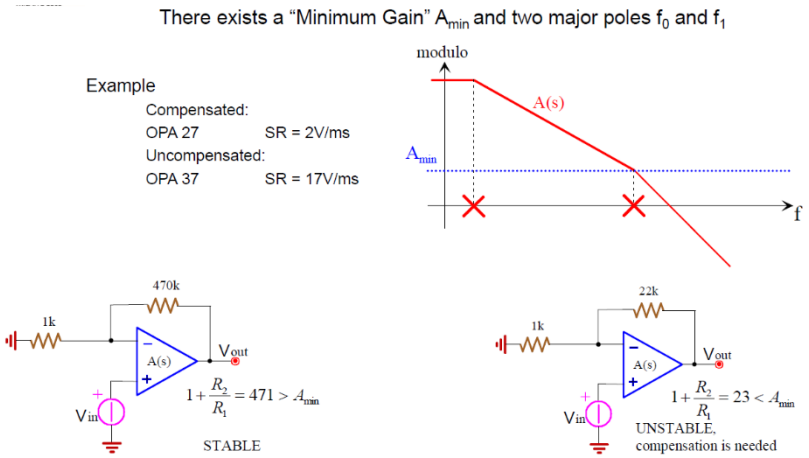


It differs from the non-inverting stage. When we switch off Vin, however, the two circuits are the same, so the Gloop of the non-inverting and inverting stages is the same. We can see that the opamp is compensated. We cut it after the opamp, we put a voltage source and check Vf to get 1/beta.
So we plot the ideal gain that is -R2/R1, that IS NOT 1/beta. Than the real gain is the ideal one up until f*, when A(s) touches 1/beta(s), and then we will dye as Gloop dies. Hence f* is the pole of the close loop gain.

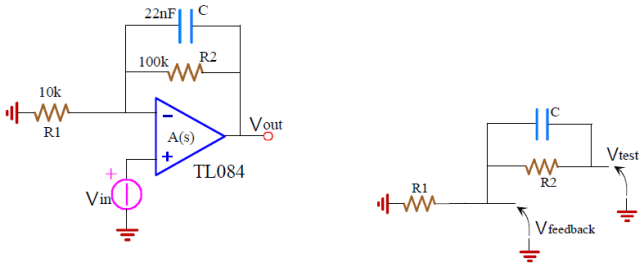
Uncompensated opamp

Sometimes datasheets quote the first pole, A0 and the height Amin. Then we can compute the second pole because A0/Amin = f2/f1. They give us the height Amin because if we connect a feedback to that opamp, e.g. in the non-inverting configuration, if we have an ideal feedback smaller than Amin we are unstable, because we are in the portion of the Bode plot where the crossing will result in 40-40.

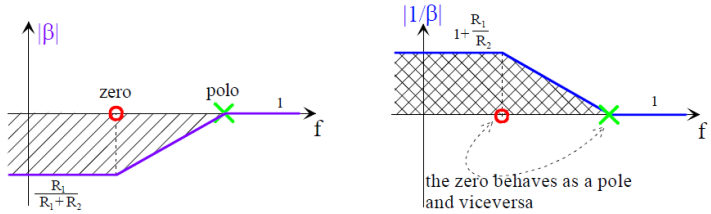
So when we use an uncompensated opamp we run the risk that, with the same configuration but with different values of resistances, one configuration is stable, the other not.

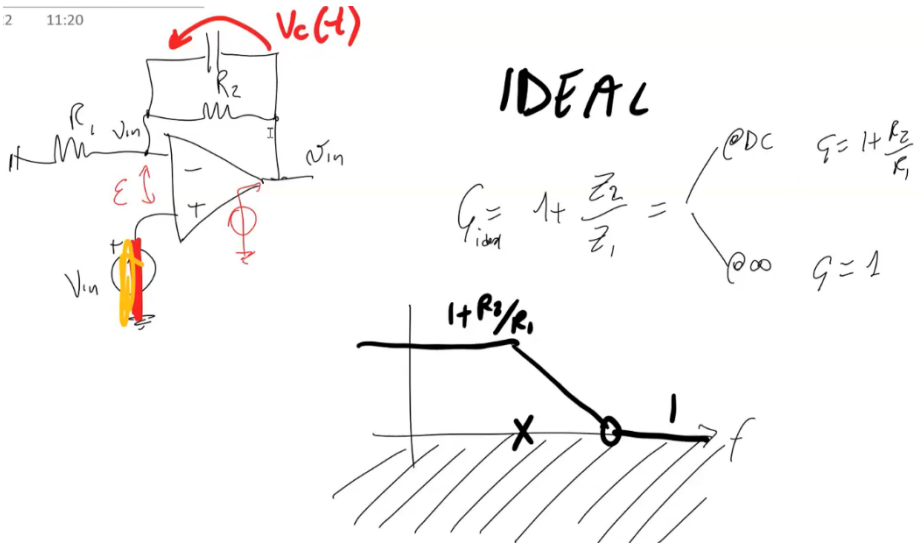


ROLE OF FEEDBACK CAPACITANCE



The 1/β(s) has a low-pass shape, which increases G_{loop} at high frequencies

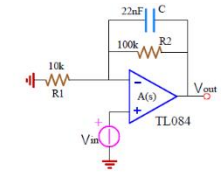




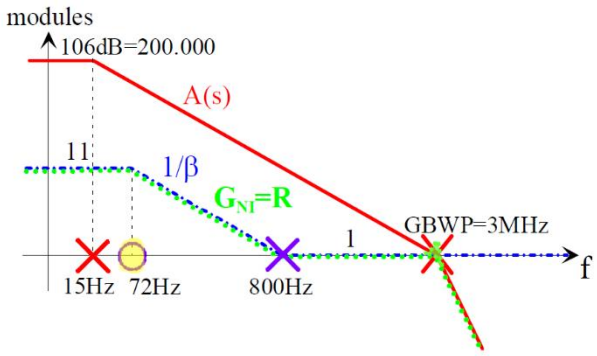
To compute the pole, we consider the input generator off, so we have 0V on both the – and + terminals of the opamp (epsilon is 0). So no current through R1. So the current recirculates on the C – R2 network → the pole will depend just on Req = R2.

Then we can compute the zero in a graphical way. It will be at the frequency of the pole multiplied by the distance between the DC and AC gains.

But we are studying the ideal gain. So we define an A(s) and beta(s). We excite with a voltage source the beta(s) in the same position where the opamp touches the feedback network and then we measure the signal that returns at the input of the opamp. We consider two regimes, DC with C open and AC with C shorted.



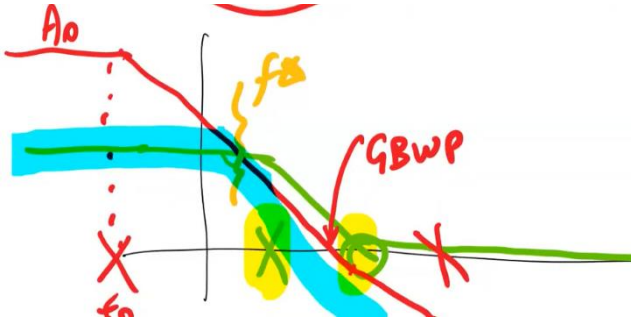
The closed-loop pole is equal to the open-loop zero of the feedback network β(s) !



To study the real behaviour of the circuit. We remove the opamp and the plot A(s), and the second pole is set coincident with the GBWP. So the crossing may be 20-40, the phase margin is 45° and the poles are c.c. with an angle of 60° in the root locus. Hence we will have a negligible overshooting in the time domain.

Then we compute beta and we plot 1/beta and we see that the crossing is so that the phase margin is 45° and poles are c.c. with 60° in the root locus. The response is ok, I like it.
To compute the ideal gain of the circuit we have to compute the ideal gain, that in this case overlaps with 1/beta, but in this case it happens just by chance.

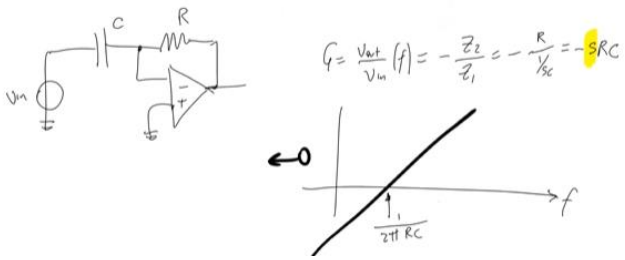
The real gain is equal to the ideal one up to Gloop, and then it dies like A(s). Hence the real circuit has two poles, not just one. The pole of the CL gain is coincident with the zero of the beta, that was 1/(2*pi*C*R2).
The circuit is well designed if the pole of the beta network happens before we reach GBWP. If we make the wrong sizing and e.g. the pole of the opamp is at too low frequency, we may have another configuration in which the pole of the circuit is due to f* and not due to the feedback network.



DERIVATOR STAGE

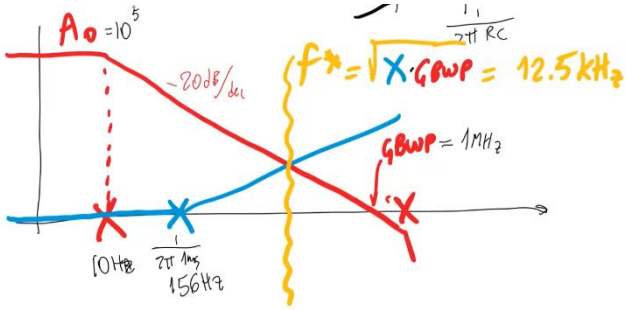
Ideal derivator

The gain of the stage in the frequency domain is -sRC. Hence we have a zero in the origin in the Laplace domain, and so a derivation in the time domain. This is the ideal gain of the stage. The frequency at which the gain is 1 is 1/(2*pi*R*C).



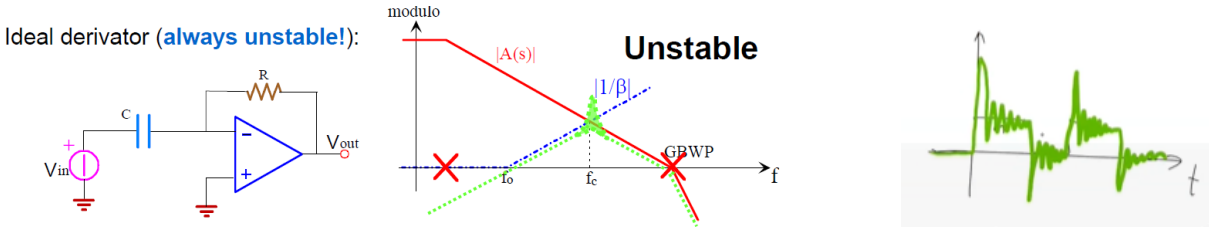
Let's now study the circuit with the Gloop considerations.
We have an opamp that will be our A(s), and then we remove it and compute the beta(s). Let's suppose the opamp is compensated.
To compute beta(s), we should shortcircuit the voltage generators and open the current generators, cut at the output and excite with a v_test and I want the v_feedback that returns to the opamp. Of course, beta(s) is frequency dependent. At DC, beta is 1 (C open), while in AC beta dies.

Now we plot A(s) and 1/beta(s).



f* is the frequency above which the loop is not working in the proper condition because with increasing frequency the opamp amplifies less (A(s) decreases) and beta is attenuating more and more, so A*beta decreases.
The crossing in this derivator is at 40-40, so we have two c.c. poles, the residual phase margin is almost zero and the circuit is unstable → in the Bode diagram we will have a strong peaking.

As for the real gain, firstly we need to plot the ideal gain, that is a straight line with a zero in the origin. To put it in the plot, we know at which frequency the ideal gain crosses the 0 dB axis. This frequency is the pole of the beta network in this case.
As for the real gain, we follow the ideal one up to f* and then we decrease as Gloop dies, so with -20 dB/dec (because A(s) dies by -20 dB/dec and 1/beta increases +20 dB/dec, and so compared to the previous trend we need to go down by -40). The real gain is the green one, it is a derivator up to f* and then it dies.



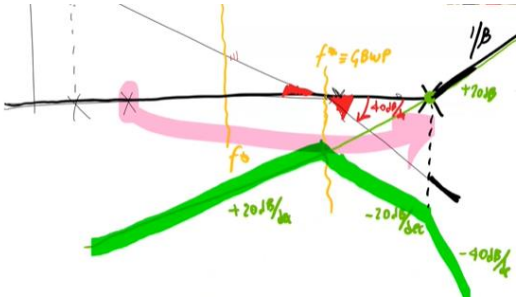
Moreover, since the closure angle is bad, the poles are c.c. and the circuit is unstable, so I also display a strong peak in correspondence of the f*.
The part up to f* is good, the one after is bad, because if A(s) changes (e.g. due to time, temperature ecc.), the falling side of the Bode diagram can change. And example of the response we may have is in the plot on the right.

Real derivator

We add a resistor in series with the capacitor. Now we have again a zero in the origin but we have a pole that causes a saturation at HF.

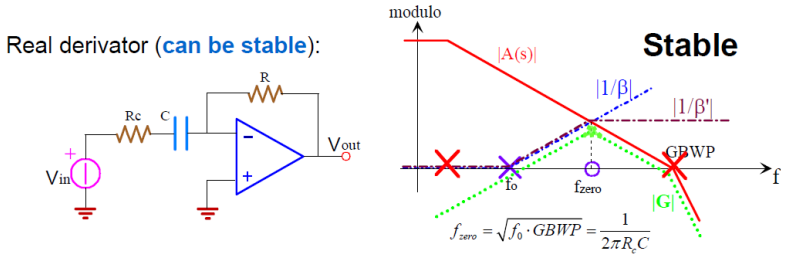
Previous issue possible solution

To improve the crossing of A(s) and 1/beta(s) like 40-40, what we can do is to move far away the pole of the beta network so that the crossing is 20-20, and the f* will be at GBWP. But the problem is that in this case, still in the case of the ideal derivator, the real gain will go from +20 to -20 as slope. After the pole of the beta then, it will die with -40 dB/dec. Technically we removed the instability condition, but the behaviour is bad, because I expect the derivating action from 0Hz up to the pole of the beta network but it is not like that, because it is up to the GBWP of the opamp.

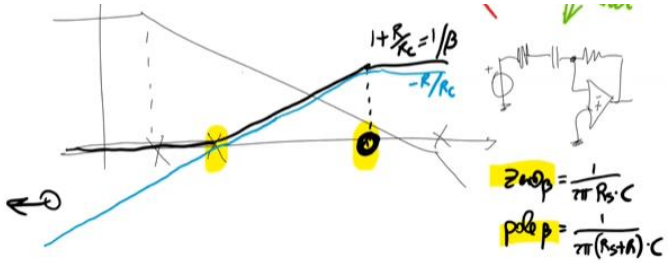


Another possibility is to change the opamp so that the GBWP is aligned wit the pole of the beta network. But decreasing the performance of A(s) to gain stability is not a smart move.

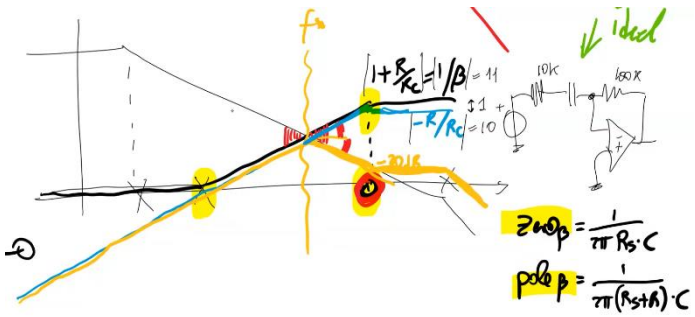
So in reality what we should do is to keep the same opamp, and in the beta network we add a zero thanks to the resistor in series with the capacitor, because we are causing 1/beta to saturate. The zero is given by C*Rc.



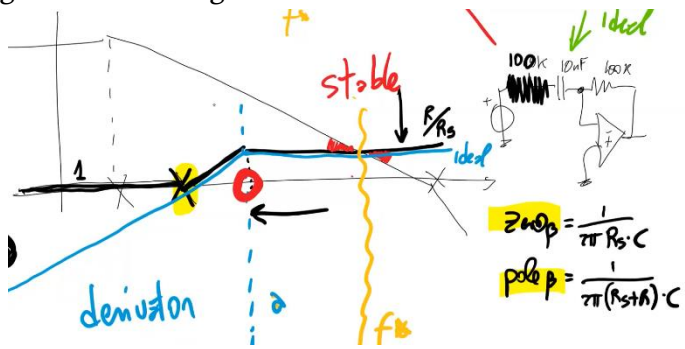
Now the ideal gain of the stage is different with respect to the case of the ideal derivator, because we have a saturation of the gain at infinite frequency (blue line). In this case the curve of 1/beta and ideal gain differ by 1.



We introduced the resistor Rc, but if the zero happens at the wrong frequency, the circuit is still unstable, like in the case above, because the crossing is still 40-40. The orange one will be the real bode diagram.

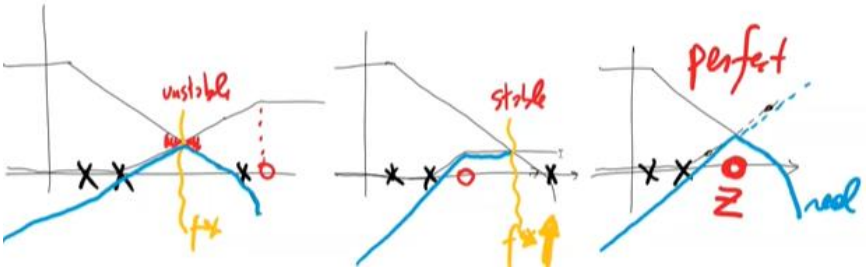


Conversely, if the zero of the beta is placed at too low frequencies it happens as below. Now the stage is stable because C is so big that the crossing is 20-20.



If now we increase the resistor value, for instance from 10k to 100k, it happens that the zero moves to lower frequencies and also the saturation value decreases and the one above is our new Bode plot. It is stable and the circuit is now a derivator but up to a certain frequency that is the zero of the beta, and then it behaves as an amplifier up to the f*. The problem is that this stage is bad because it is not a derivator across the full stage, even if it's stable.

We need to find a compromise; let's put the resistor such that the zero and f* have the same frequency.



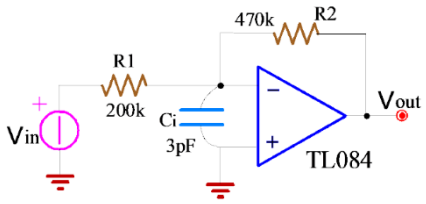
The real behaviour is a derivator up to where the circuit collapses, so it doesn't show an amplifier region. Now the derivating gain collapses at the highest frequency. Given the value of the GBWP, we set the zero at the following frequency.

$$z_{ero} = \frac{1}{2\pi R_c \cdot C} = \sqrt{\frac{1}{2\pi R \cdot C} \cdot GBWP}$$

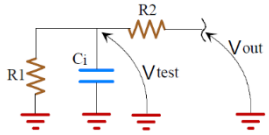
Moreover, now the stage is also stable because the crossing is 40-20, so we have a marginally stable circuit.

EFFECT OF AN INPUT CAPACITOR

This capacitor can be the parasitic capacitance of the opamp or something that we have in the circuit.
If the circuit was ideal, in Ci there should be no current because VG concept applies and Ci is between VG and ground.

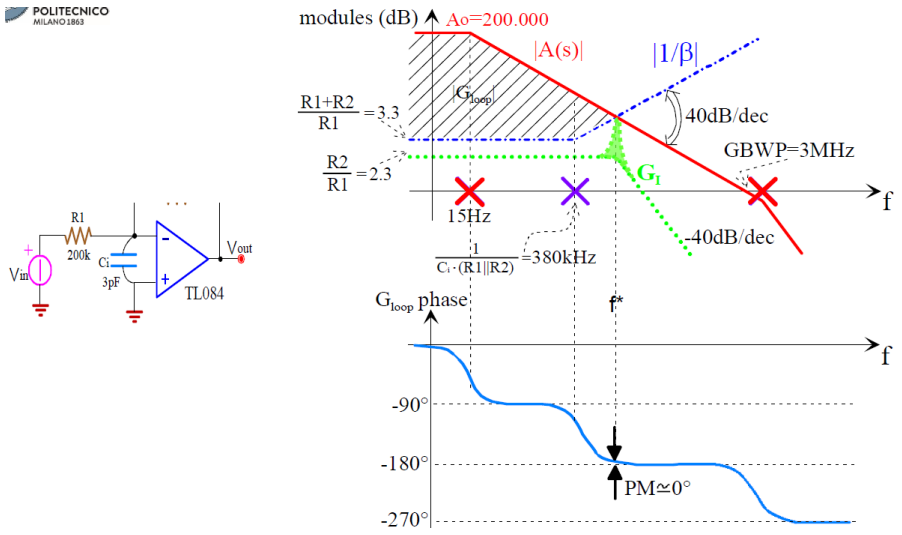


It threatens stability because it alters the feedback β



$$\beta = \frac{R_1}{R_1 + R_2} \cdot \frac{1}{\left[1 + \frac{s}{\left(C_i \cdot (R_1 \| R_2)\right)}\right]}$$

Hence the capacitor is like having no role, like if it was not connected there. The ideal gain of the previous circuit is 1+R2/R1, constant in frequency.
Let's now study the real gain, with A(s) and 1/beta(s) (in the image v_test is v_feedback and v_out should be v_test).

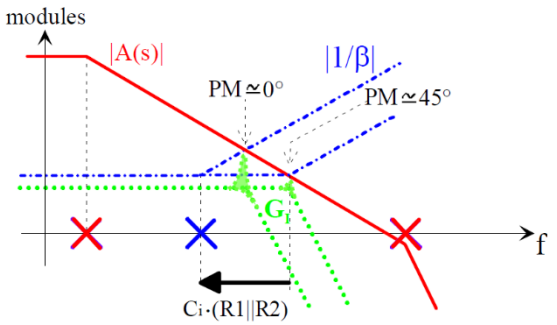


We see that the real gain, instead of proceeding flat, dies with -40 dB/dec. Moreover, the circuit is unstable and if the second pole happens 1 decade before f*, the phase margin is almost 0. So it is not correct to add a capacitor just at the input of the amplifier, because it causes instability.

However, we can bring the circuit back to stability either by removing the capacitor (so that 1/beta is constant for any frequency) or we can perform components' sizing, because maybe the capacitor cannot be removed and it's simply there because of parasitisms. If Ci gets smaller, the pole of the beta moves to higher frequencies and maybe the crossing returns good. Thus we achieve a 20-40 crossing and achieve marginal stability.

Re-catch stability: no additional component, but simple value adjustments

By changing C_{in}
or the value of $R_1//R_2$
(i.e. the input pole)



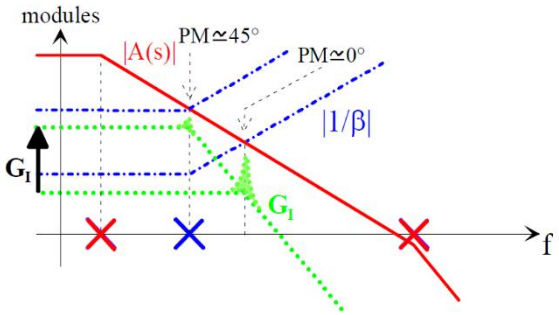
If Ci cannot be changed, we can still act on the resistive network. R1 and R2 are needed to set the gain, but we can change them by scaling them of the same factor so that the gain remains the same, because it depends on their ration, while the pole, that depends on their parallel, moves at higher frequencies. Now we are changing both R1 and R2.

Another possibility is to change just R1 or R2 so that the pole stays the same but the gain changes. We increase the gain.

Re-catch stability: no additional component, but simple value adjustments

By changing the value of R_1 or R_2

(i.e. the gain)

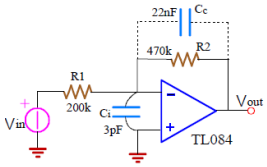


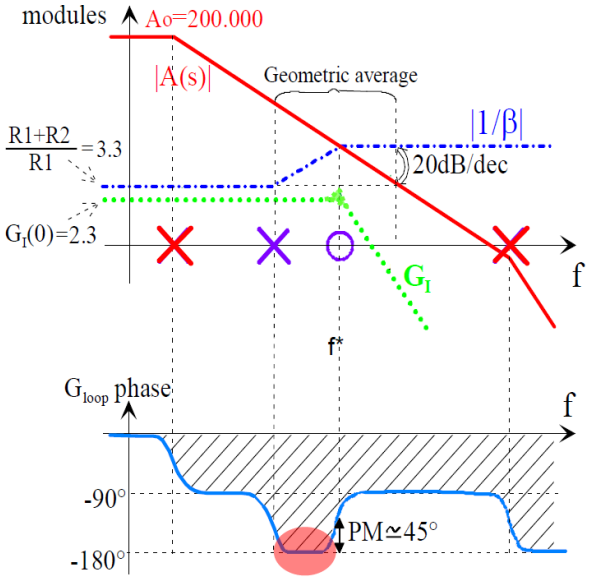
The last possibility is to regain stability by adding components.

Compensation of instability by adding components

POLITECNICO MILANO 1863

Compensation is needed !

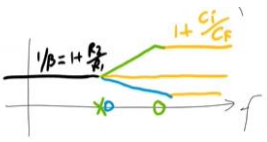




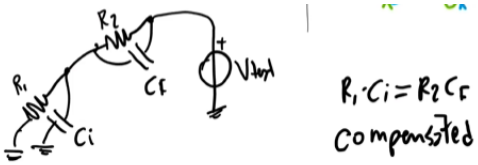
The opamp is ok because it is compensated, it is the beta(s) that is bad.
To guarantee a good crossing, we may want the beta to saturate, so we want a zero in the beta that causes saturation, but it must be positioned correctly to gain stability, so that the crossing with A(s) is 40-20 or 20-20.

A possible way to introduce a zero is by adding a capacitor in parallel to R2 (Cc). In fact, every time we have a C in parallel to a R and the signal has to propagate through there, this portion of the circuit generates a zero with tau = R*C. Instead, we don't have two poles because the two capacitors are in parallel and not independent. Hence we have just one pole with tau = (Ci + Cc)(R1||R2).

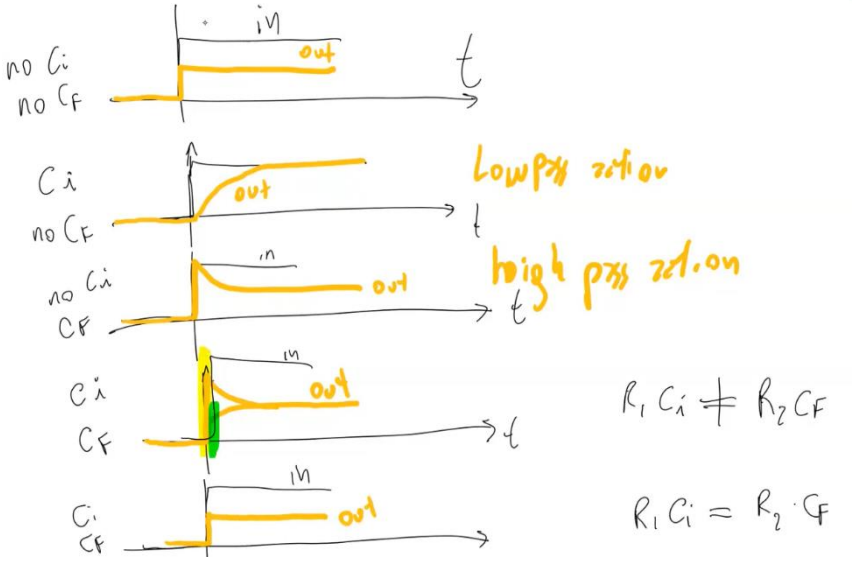
In the beta network, we may have different possibilities. In DC, 1/beta is 1+R2/R1, while at AC impedance depends on the two capacitors, because their impedance is much smaller than the one of the resistors, so it will be 1+Ci/Cc. Depending on the values of Ci and Cc we may have a higher, equal or lower value with respect to 1+R2/R1.



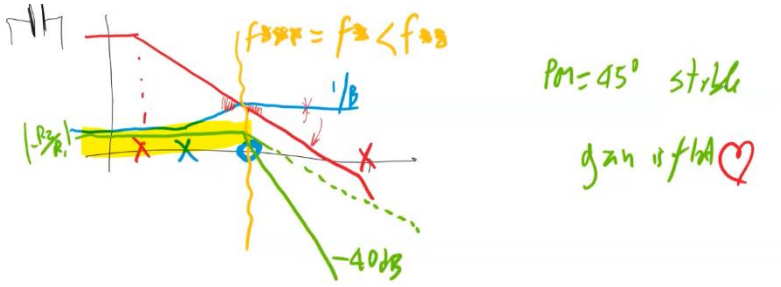
In general, in a voltage partition network, if the product between R1*C1 = R2*C2, the voltage partition is said to be compensated. This means that if the previous equation holds, then the Bode diagram, is flat.



The compensation is important because if for instance we enter in the circuit with a step and we have no Ci or Cc, the output of the circuit will be for sure a step attenuated (in the time domain). If instead we have just Ci, Ci causes LP filtering action. If we have Cc but not Ci we have a HP filtering action. If we have bot Cc and Ci and the network is not compensated, then either the LP will prevail or the HP will prevail, so we will have a trend with a step and then regime or with an exponential increase.
Eventually, if the Ci and Cc are matched, if we enter with a step the output will still be a step because we will have a perfect compensation of the LP transition and of the HP transition.



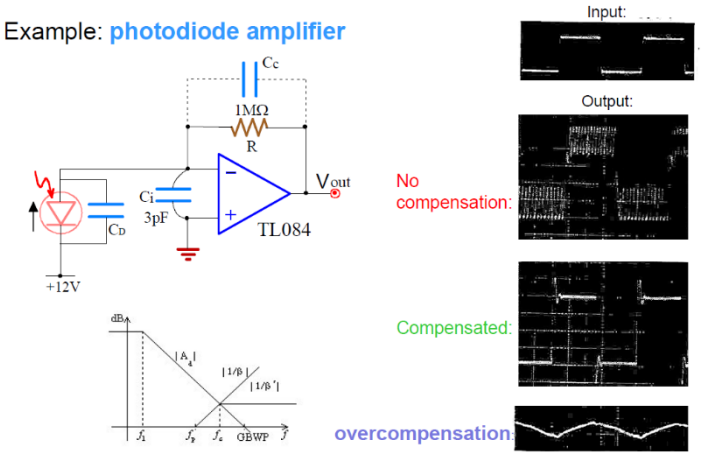
Hence with the Cc capacitor we can regain stability. The zero is not to be placed at too low frequencies, nor at high frequencies, because the best way is to place it where A(s) and 1/beta(s) cross each other.
Thus the real gain is the green one; in fact, the ideal gain is the one of a non-inverting stage in DC, then it collapses with a pole that is the zero of the beta network. If Cc was placed so that the zero is before f*, than in the operation range of my amplify I would introduce a LP filtering action that is not required.



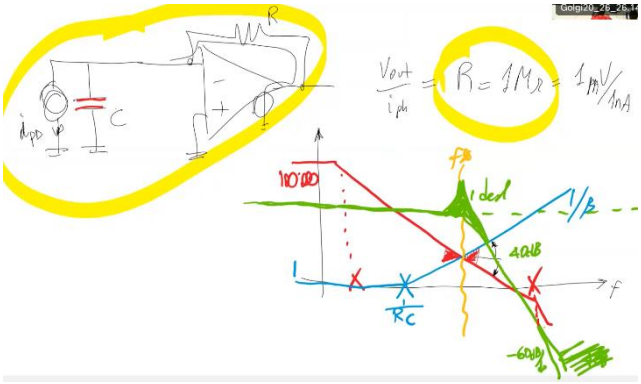
The gain dies with -40 dB/dec with a minor peaking because we are marginally stable, and Cc > Ci.
If we draw the phase plot, we notice that it is the zero that recovers the stability. However, just before the zero we have a portion of the plot where the phase margin is 0, so theoretically we should be unstable in that region where Gloop is higher than 1. However, there is not the problem because it is only where Gloop = 1 that the phase is important, because if we have poor phase margin even before it doesn't matter.

Eventually, if Cc is too big, the zero is at a so low frequency that we see the low pass filtering action, and the output transition to the steady state value becomes so slow that the stage is unusable.

TRANSIMPEDANCE AMPLIFIER

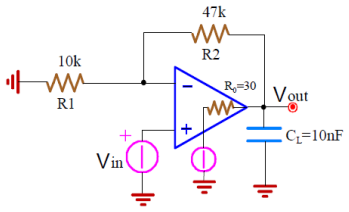


This stage is introduced to demonstrate that the ideal gain is not coincident with 1/beta. The photodiode can be modelled with a current generator in parallel with a parasitic capacitance. The Iph in DC passes through the feedback resistor R and Vout/Iph = R.
If we study the stability, A(s) is compensated, and the beta is 1 at DC, then we have a pole.

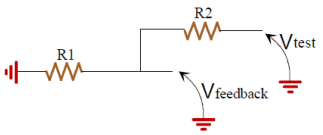


Hence a standard transconductance amplifier is always unstable, because the crossing will be bad, 40-40. Since we have peaking, we have two c.c. poles. Hence for sure we need to introduce a Cc capacitor. If Cc is too small, we don't grant enough compensation. If it is too big, we are overcompensating.

EFFECT OF AN OUTPUT CAPACITOR



It threatens stability because it alters the forward gain A(s)



$$A \cong A(s) \cdot \frac{\frac{1}{sC_L}}{R_0 + \frac{1}{sC_L}} = A(s) \cdot \frac{1}{1 + sC_L \cdot R_0}$$

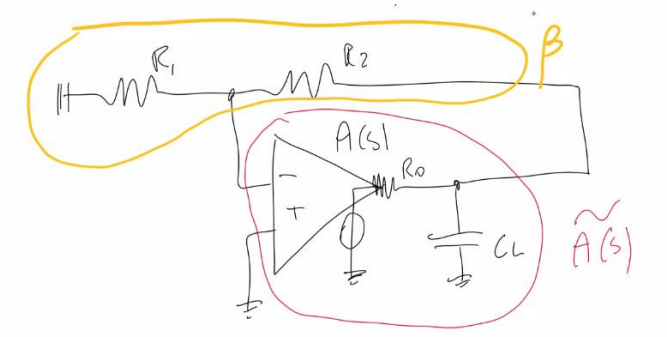
$$\beta = \frac{R_1}{R_1 + R_2}$$

… due to the additional output pole: $f_{out} = -\frac{1}{C_L \cdot R_0} = 530kHz$

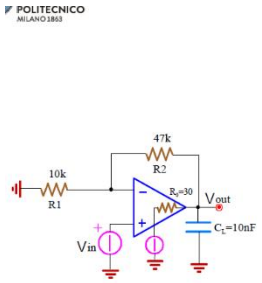
The output capacitance, as well as it was for the input one, could be added on purpose or it can be a stray effect.
The amplifier can be either in a non-inverting or inverting configuration and for instance if we drive the load with a long coaxial cable we introduce a stray capacitance.

Of course, if the opamp was ideal with a zero value of Ron (output resistance), then the Cl won't have any effect; however, if we use an OA with a given value of Ron, it is not negligible.
So we start with a compensated opamp but then, due to Ron and the Cl we add a pole. The new A(s) can be considered not just the A(s) of the opamp but the combination of the opamp and Ro-Cl network.

The idea is to consider the following circuit with an OA with A(s), then we add a pole due to Rout (Ro) and Cl and then we have the feedback network due to R2 and R2. We consider a new A(s) and the remining part is the beta network.

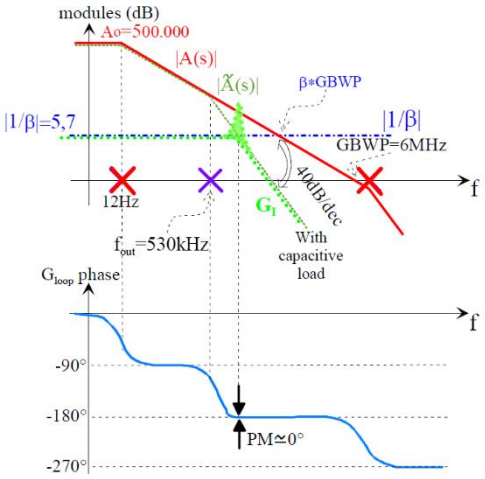


Let's plot A(s)-tilde. It will be due to the typical response of the compensated opamp but then, due to Ro and Cl we have a pole that is at the frequency with tau = Ro*Cl and so we decrease with -40 dB/dec after this pole. Now the A(s)-tilde is similar to the frequency response of an uncompensated opamp with the second pole before the GBWP.



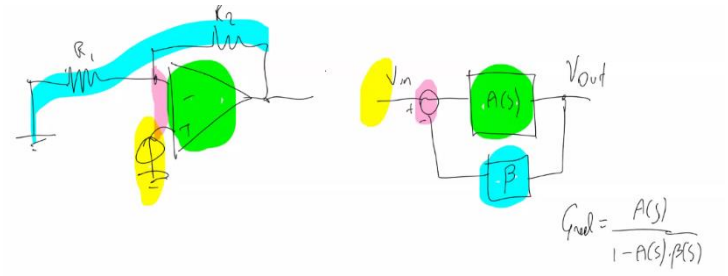
$$f_{out} = -\frac{1}{C_L \cdot R_0} = 530kHz$$

$$f_{polo} = \sqrt{f_{out} \cdot \beta \cdot GBWP}$$

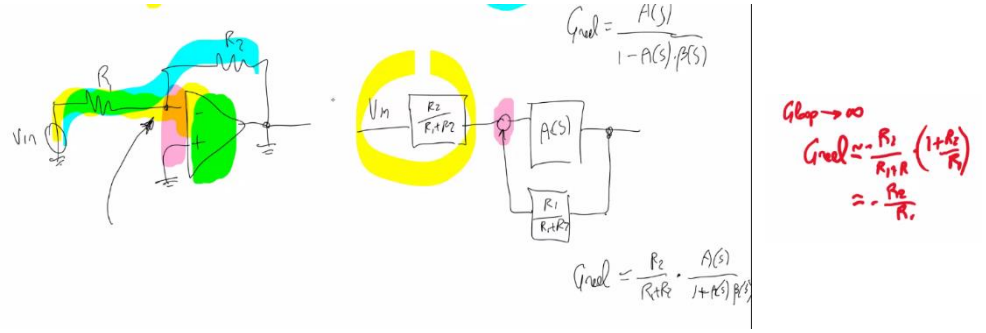


Now we should study beta, but it is simply a resistive network, so it will be constant in frequency. We see that we are unstable because the closure angle is 40-40 → there will be peaking. In this case the ideal gain is equal to 1/beta, just by chance.
As for the real gain, it is equal to the ideal one up to when A(s)-tilde and 1/beta touch, then it dies by -40 dB/dec.

If we study another circuit, the inverting configuration, in this configuration we cannot model the circuit with the typical block diagram of automation and control. This modelling applies only if we study the non-inverting configuration.



If we use the inverting configuration, Vin gets attenuated by R1 and R2, so the signal that reaches the loop is not Vin but its attenuated version. Then we have the subtracting node and then we enter in the opamp with A(s), then we have the attenuating network beta and then the signal gets subtracted from the previous one. Hence Greal is different.

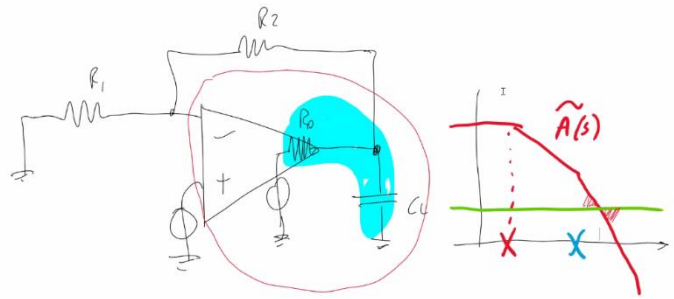


In this case, when Gloop is very very high, in the noninverting configuration, if Gloop → inf, the Greal tends to 1/beta. In the inverting one, it tends to -R2/R1.

So the classical automation control blocks diagram happens only for the non-inverting configuration, in all the other cases and circuit we cannot do this. So we have to compute the ideal gain and then study Gloop, that is instead the classical automation and control diagram with A(s) and 1/beta(s). Ideal gain is 1/beta(s) only for the non-inverting configuration.

Compensation

Let's try to compensate. And easy idea that we could use is (I know that it is unstable because we have a resistance Rout and a capacitor Cl, so the overall A(s)-tilde (the red circle) goes as in the red plot, so since the 1/beta crosses at -40db, the system is unstable). To compensate, we could add a Cc capacitor in parallel to R2. But this Cc causes beta to be equal to 1 at HF, hence Cc is not recovering stability.

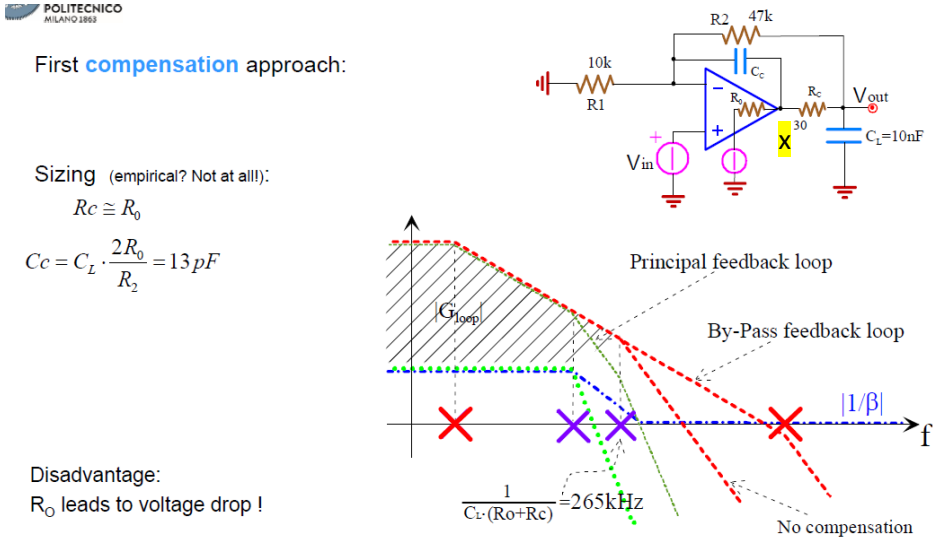


Anyhow, the crossing will still be bad, so Cc cannot guarantee stability. Cc tries to modify the 1/beta(s), but it decreases so much that is the crossing is so bad, because the Ro and Cl create the pole. So to compensate we can do something different.

The best idea is to feed the node between the voltage output generator and Ro and feed it back at HF (e.g. with a capacitor) to the input. But this cannot be done because the node is inside the amplifier, so we cannot have access to it.

Hence what we do is that we put a resistor (Rc) in series with Ro, eventually even worsening the effect of Ro because the pole decreases in frequency, and then I use a capacitor to feed back to the input.

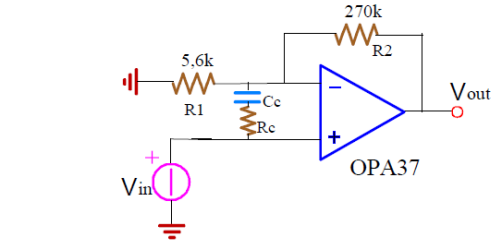
So nothing changes in DC but in HF Cl goes in short, so A(s)-tilde should die, but since we have Rc in series with Ro, we can have access to node x and we can take it and feed it back to the input through Cc to regain stability.

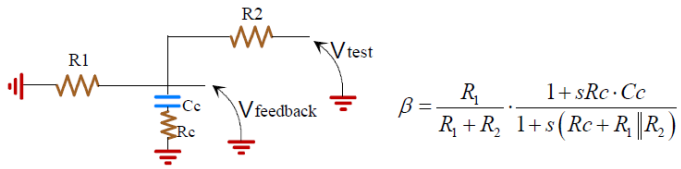


The advantage of this solution is that at AC we will have just A(s), not A(s)-tilde, the 1/beta is 1 at HF and now we grant stability.

COMPENSATION WITH NEGATIVE FEEDBACK

This approach is suitable for the compensation of any pole of A(s)
It improves the Slew-Rate too

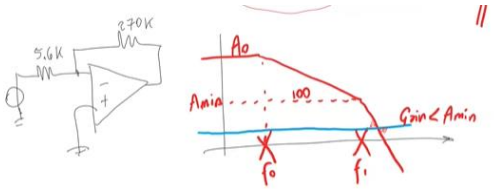




We introduce a further negative feedback. Let's imagine that the stage in the image is unstable. Let's imagine to have an uncompensated opamp (so A(s) of the opamp has two poles). If we introduce the couple Rc-Cc we can regain stability.

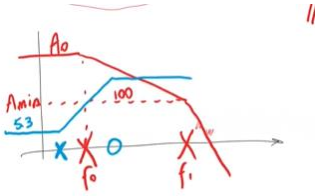
Let's imagine A(s) is uncompensated and let's study beta. At DC the capacitor is open and beta is R1/(R1+R2), the usual one. Then in AC we have the capacitor shorted.

To compensate this stage, and Amin = 100, for example, and I want an amplification of 2.3, the crossing will be bad and the stage unstable because Gain < Amin. To regain stability, we could increase the gain to have a good crossing. But I don't want to change the gain, so maybe I could use a

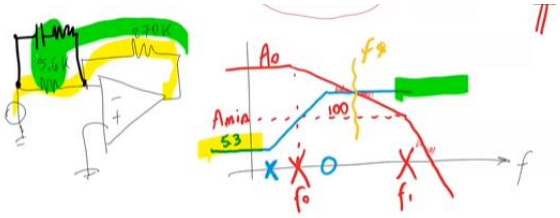


compensated OA. If I don't want to change neither the gain nor the OA, we can change the position where the crossing occurs.

So at DC we use the gain we want, e.g. 5.3, but then we introduce a pole and a zero in the beta network to shift the gain up and compensate everything. To do so, having a beta high at low freq and low at high freq (in the plot I have 1/beta), I cannot use a capacitor in feedback, because it causes 1/beta to be ok at LF, but then it goes to 1 at HF, so it's even worse. So Cc is not the solution because it cause the stability to be even worse.



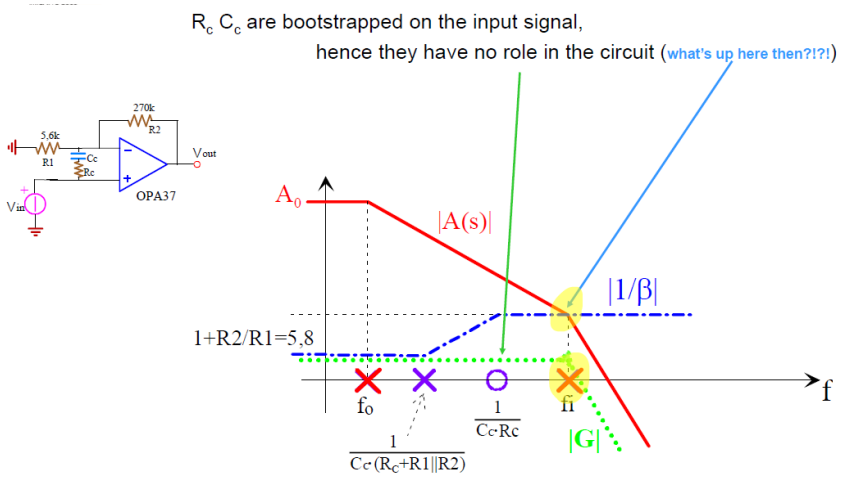
If we want beta to decrease, we could place something in the input branch, because we want beta to decrease at HF is we want a high 1/beta. Hence we place a R-C series.



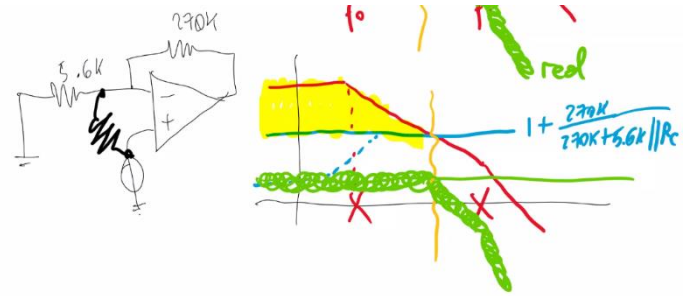
This solution could work, but the problem is that we are changing the gain of the circuit. The new ideal gain is no more flat, but it has a pole and a zero, and we want it to be constant.

The new ideal gain is not flat, but it is flat, rises and goes flat again, and this is not good for the customer, because the customer what a constant gain. So if I want the ideal gain to be constant, I can include the compensating network connecting it between the terminals of the OA. If so, thanks to VG the two don't play any role in the ideal gain, and it remains constant in frequency.
If epsilon is 0, they don't play any role in the ideal gain. As for the Gloop, I have the OA, it's A(s) and to study beta I shortcircuit Vin so the two branches in input are in parallel. Thanks to this smart connection I've compensated the stage.

So ideal gain is something, the 1/beta(s) is something different and also the actual Gloop is different. So ideal gain is never equal to 1/beta except for the ideal non-inverting configuration.

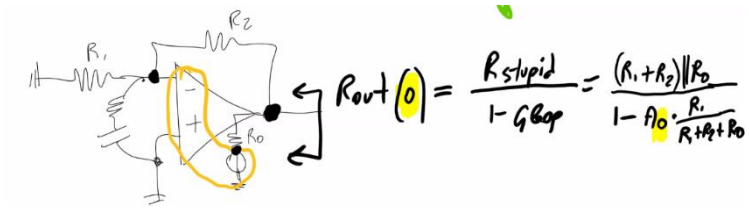


If we remove Cc, thanks to the feedback, ideally in Rc no current flows if Gloop is strong enough, so Rc doesn't impact in the computation of the ideal gain. However, if the ideal gain is not changed, the beta is changed. A(s) is still the same, but now beta is always high and constant → we are loosing the low 1/beta at low frequency, so now Gloop is lower than before.



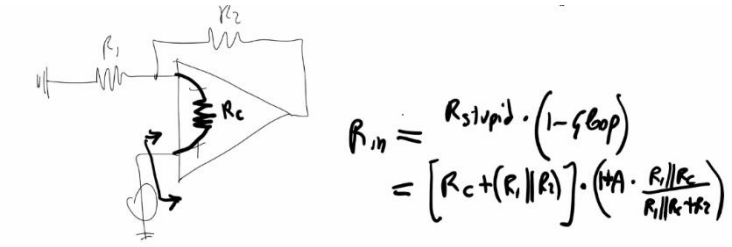
However the circuit is still stable and the ideal gain (green) is the same as before, and the real gain as before.

Having a smaller Gloop means that the feedback is less strong. Having a high feedback is a good thing because if we have for instance a real OA with its Ro and we want to compute the Rout, the Rout(0) is equal to the Rstupid/(1- Gloop), where the Rstupid is the impedance I see from the output (A0 is A(s) at 0 freq from the input to the upper node of the output generator, so Ro must be considered).



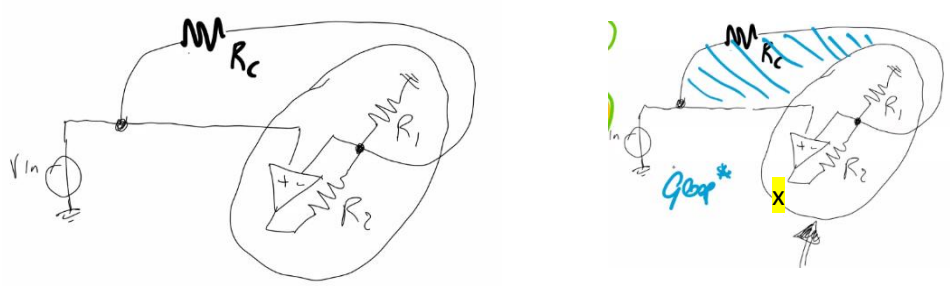
If the Gloop is sufficiently high, the Rstupid is reduced a lot. If we remove Cc permanently, now also at DC the equation changes, because we won't have R1*A0 but (R1||Rc)*A0, so Gloop is lower and Rout is higher than before, because the action of Gloop is not that much strong.

What about the input impedance if we place Cc and Rc (or just Rc)? Is it no more infinite?
No, it is still infinite because epsilon across the Rc Cc series is zero, since we have Vin both at the + and – terminals. The impedance we see is the stupid one but magnified by 1 – Gloop.



So if Rc is infinite, the input impedance is infinite, but if it is not infinite, thanks to the feedback it gets really big.

Let's redraw the schematic with Rc outside the OA and modify the circuit a little bit.



I have a circuit, a loop, a node and so the input impedance is the stupid one divided by 1 – Gloop. But if I divide by 1 – Gloop, it means that the result is less than before, because Rin was Rstupid was magnified by Gloop. However, the two equations are correct anyway! The two equations in fact refer to two different Gloop.

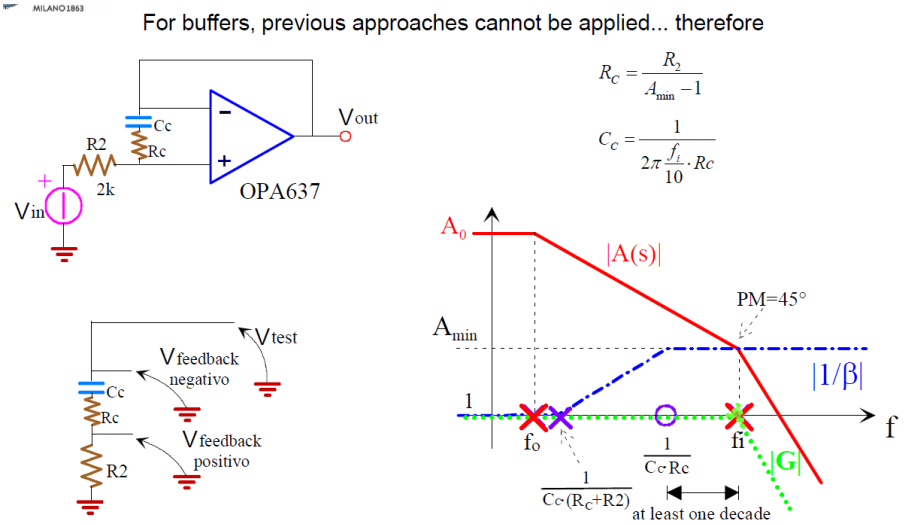
The two equations are correct but the Gloop is not the same for the two, they differ from each other. In the previous equation when the impedance Rc gets magnified by Gloop, Gloop is given by the feedback of the OA (negative feedback).
In this other case the loop is given by another net, the new Gloop = Gloop* (blue). Let's compute Gloop*
The gain of the part inside the black circle is not 1 + R2/R1, because our 'output' is not at x, but it is at the – terminal of the opamp. Ideally, epsilon is 0, so the gain is +1.

So we can compute Rin.

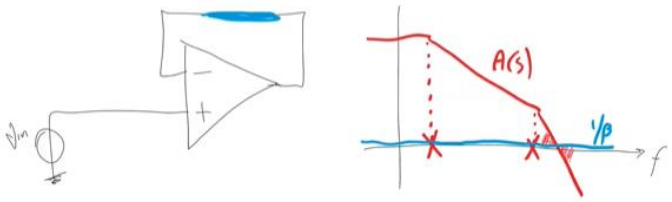
$$R_{in} = \frac{R_{stupid}}{1 - G_{loop}^*} = \frac{R_c + (R_1 \| R_2)}{1 - A^* \cdot 1}$$

Hence we have a division by 0, and Rin becomes infinite. And this is the exact same result than before, when the resistance is amplified by Gloop and so still becomes infinite.
In reality, the gain A* is not exactly 1, but it is 0.99, because the loop is not ideal, so we are dividing by 0.01, so we get a very high value but not infinite.

COMPENSATION WITH POSITIVE FEEDBACK

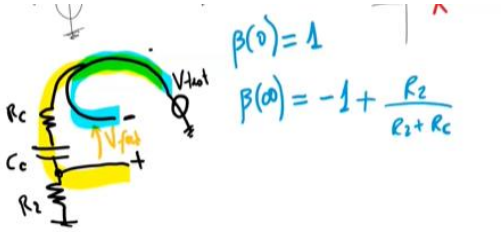


We have an opamp and let's suppose we want to use it as a buffer. The problem is that the opamp I bought is uncompensated, se we are out of stability.

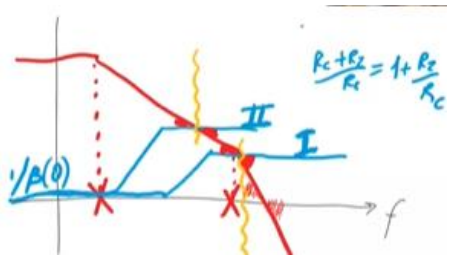


Hence we may want to add a compensation network. If I add a compensation capacitance in feedback it is not good, I'm changing the circuit. So can I eventually introduce the Rc-Cc couple? No, because at 0Hz beta is 1 and at HF it is still one, so not compensated. Hence we need to introduce also an additional resistance. Now it works.

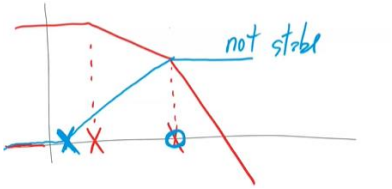
Now in DC beta is 1 (capacitor open), instead, at HF, the capacitor is short circuit and the beta network is the one two images ago. The v_feedback I'm interested in is the one between the input terminals of the opamp. Beta(inf) is the beta at the minus node (-1) plus the beta at the plus node (we can also compute simply the voltage divider and we achieve the same result).



In this case the beta is the sum of the contributions because the + terminal of the opamp is not connected to ground like in all the other previous cases.
Since now beta is < 1 at HF, 1/beta will be high, so eventually we can cross the uncompensated opamp in the 20-20 or 20-40 crossing. We prefer the first case because f* is higher, and the higher f* is, the better is the bandwidth.

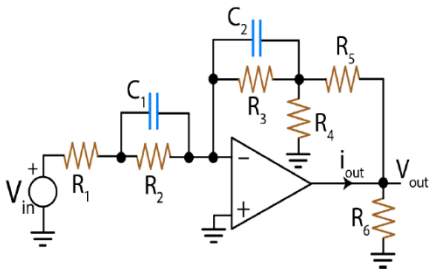


If we have a situation like below, we are still out of stability because we are still maintaining a 40-40 closure angle.



The circuit is stable if the zero of the beta network is at least 1 decade before the second pole of the uncompensated opamp, so that the phase margin is 45°.
Cc can be removed, but it is better to keep it to have a very large Gloop in DC, because if we remove it 1/beta will be constant and equal to the HF value in the case we have the Cc.

Example 1



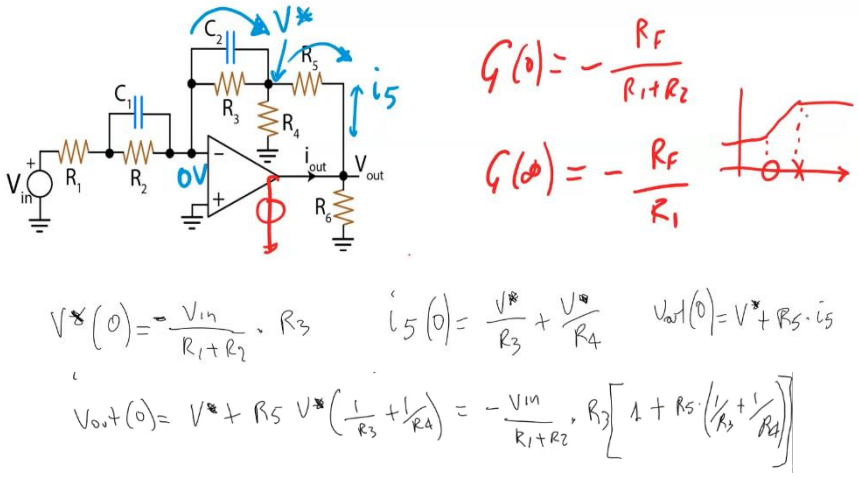
R_1=10kΩ, R_2=100kΩ, R_3=220kΩ, R_4=1kΩ, R_5=10kΩ, R_6=47kΩ, C_1=1μF, C_2=100pF.

a) Plot the ideal $|v_{out}(f)/v_{in}(f)|$ gain.

b) Compute i_{out} when V_{in}=-100mV.

Let's check if we have a negative or positive feedback. We enter at the minus node, the signal propagates at the output and then it returns back to the input and it is negative, so the minus terminal will be VG.
We can compute the input current. If in feedback we would have had just a resistor Rf, than the DC gain would have been -Rf/(R1+R2) and at AC, the gain would be -Rf/R1, so it would have been a high pass circuit.

We have now a resistor R4 connected to ground, that is used to increase drastically the gain. The input current i_in(f) = Vin/(R1+Z1). At DC, i_in = Vin/(R1+R2), at HF it is Vin/R1. Than this current flows through the feedback. Voltage at node V* above R4 can be computed in DC and AC.
Once we have it, we can compute the current that flows in R5, i5; consequently we can also compute Vout.



The current that the opamp should provide is not just the input current, but also the one that flows through R4. We can finally write the gain in DC.

$$G(0) = \frac{V_{out}}{V_{in}}(0) = -\frac{R_3}{R_1+R_2}\cdot\left[1+\frac{R_5}{R_3 \| R_4}\right] = -\frac{220k}{110K}\left[1+\frac{10K}{1K}\right] = -2\cdot 11 = 22$$

The gain doesn't seem so high, but it is already increased by R4.
As for the rest, we have two capacitors, so we are prone to think that we will have two poles. For sure R1 and C1 introduce a pole because they are in the path of the signal towards ground (actually, towards VG), and the same for C2 and R3.
When computing the pole values, the concept of virtual ground applies, so C1 sees R1||R2. Since we have a parallel configuration of a C1 and R2 we will have also a zero with a tau C1*R2.

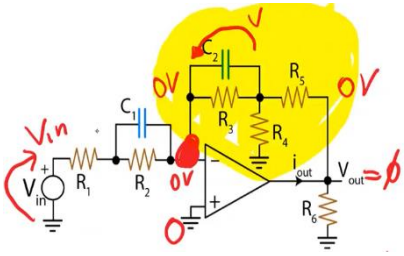
$$pole_1 = \frac{1}{2\pi C_1 \cdot (R_2 \| R_1)} = \frac{1}{2\pi 1\mu \cdot (100K \| 100K)} = 18\,Hz$$

$$zero_1 = \frac{1}{2\pi C_1 \cdot R_2} = 1.6\,Hz$$

$$pole_2 = \frac{1}{2\pi C_1 \cdot R_3} = 7.2K\,Hz$$

Now let's compute the pole and eventually the zero introduced by the other capacitor. No current will flow in the input branch because Vin is shorted, so no current flows in the feedback.
Again, now we should consider if C2 introduces also a zero. It will do for sure because at DC the gain is different from AC when C2 is in short circuit. At AC, V* = VG, so there is no current in R4 and so the gain will be provided only by the current in the input branch and through R5. Since there is a current in R5, the gain won't be 0. Hence C2 should introduce also a zero.

The procedure to understand if the capacitor introduces a zero is to suppose a signal on the capacitor, a signal in input and check whether the output is 0. Is this possible? The circuit of course is operating, so the VG is provided. Let's study the yellow circuit.



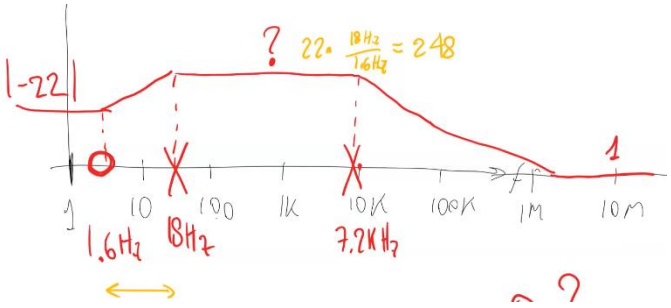
The current that flows in the capacitor is V/(1/sC2). This current should be such that the output is 0. So the resistances R3, R4 and R5 are seen as in parallel.

$$\frac{V}{1/sC_2} = -\frac{V}{R_3} - \frac{V}{R_4} - \frac{V}{R_5} = -\frac{V \cdot (R_4R_5 + R_3R_5 + R_3R_4)}{R_3 \cdot R_4 \cdot R_5}$$

The zero is at the frequency where this equation holds.

$$s = zero_2 = \frac{\frac{1}{R_3} + \frac{1}{R_4} + \frac{1}{R_5}}{C_2} = \frac{1}{2\pi C_2 \cdot (R_3 \| R_4 \| R_5)} = 1.8MHz$$

Let's now plot the Bode diagram. Of course, the zero could also have been computed with the graphical representation of the Bode plot.



The gain at infinite frequency is:

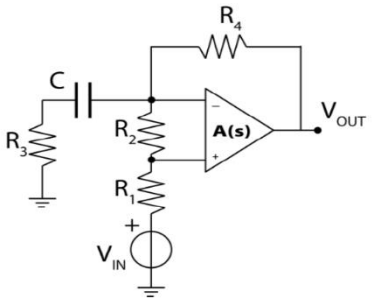
$$G(\infty) = \frac{V_{out}}{V_{in}}(\infty) = -\frac{R_5}{R_1} = -1$$

NB: R6 has no effect because the opamp is producing its Vout that is the one we need to find, and then it's the opamp that is providing the additional current on R5, but it is on the top of the feedback network, it is additional. It is wrong to say that the current that flows from the feedback goes in R6 and then in the opamp, because it is the opamp that provides the current to the feedback and to the load.
Vout then depends only on the feedback and input network and not on the amplifier and load resistor.

As for request b), Vin = 100mV is a DC value.

Vout = 22*Vin = -2,2V. To provide this voltage, the opamp should be able to provide the current -2.2/R6 and the other current that flows in the feedback, that is the current that we have in input and the extra current that should flow in R4. So we should recompute the value of V* to get the current. We already have the formula for it from the beginning, and we get V* = +0.2V.
Hence we can compute the current in R5. So we finally get to i_out = i5 + i6 = 0.5mA + 47uA = 247uA.

Example 2



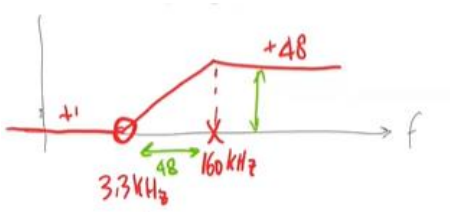
A_0=100dB, GBWP=100MHz.
R_1=22kΩ, R_2=1kΩ, R_3=1kΩ, R_4=47kΩ, C=1nF.

a) Plot the Bode diagram of the **ideal** and **real** $v_{OUT}(f)/v_{IN}(f)$.
b) Compute the range of GBWP values that guarantees stability with a P.M. better than 90°.

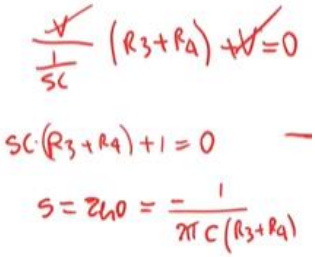
The opamp is compensated because the GBWP is specified.

a) Again, we have a negative feedback. So epsilon is 0V, so no current through R2, and so no current in R1 and hence I will have Vin both on the – and + terminal. Since in DC the capacitor is open, there is no current in feedback and G(0) = 1.
G(inf) = 1 + R4/R3.

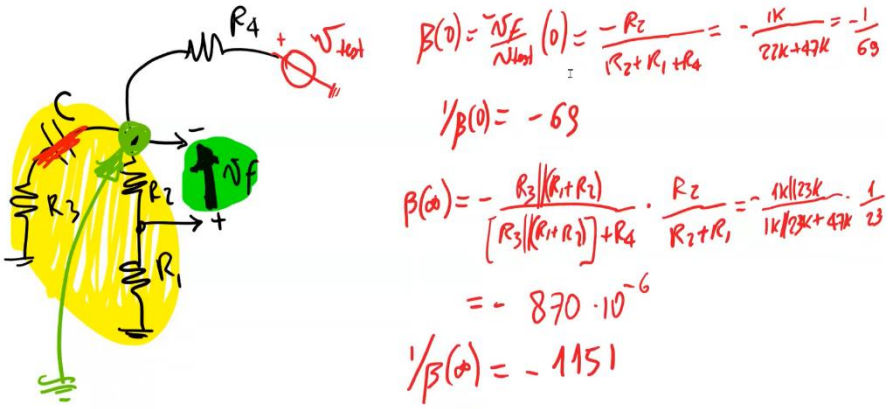
Hence the circuit will have a pole and a zero. To compute the pole, Vin is off, so since R2 is between 0V and 0V, C sees only R3. f_pole = 160 kHz. The zero can be computed dividing 160kHz by 48, so at 3.3kHz.



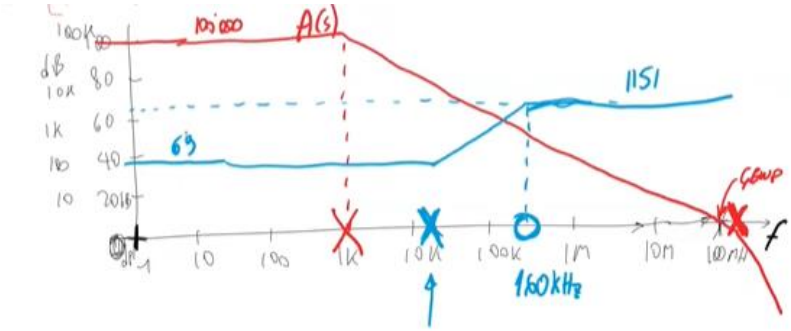
To compute it analytically, I put a voltage on the capacitor and the output to 0. The current in the capacitor will flow in R3 but won't go in R2 because epsilon is 0, so it will flow in R4. – terminal will be at Vin (not VG).



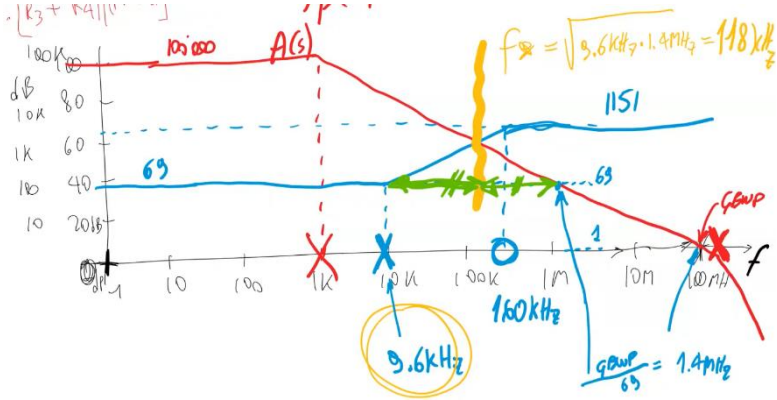
To compute the real gain we need now to study Gloop. We have the following beta network.



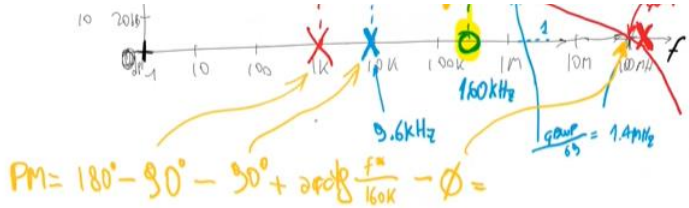
As for the poles and zeros of the beta, the pole has a Req = R3 + (R2+R1)||R4. As for the zero, we have a node and an R-C hanging from that node (C and R3). So it is easier to compute the zer0, because its tau is C*R3.
Now we can plot the Bode diagram to retrieve the real gain and to compute stability.



To compute f*, we prolong the DC value of 1/beta so that we get a triangle (green one) that has the base divided in half by f*. So we compute f* with the geometric average.

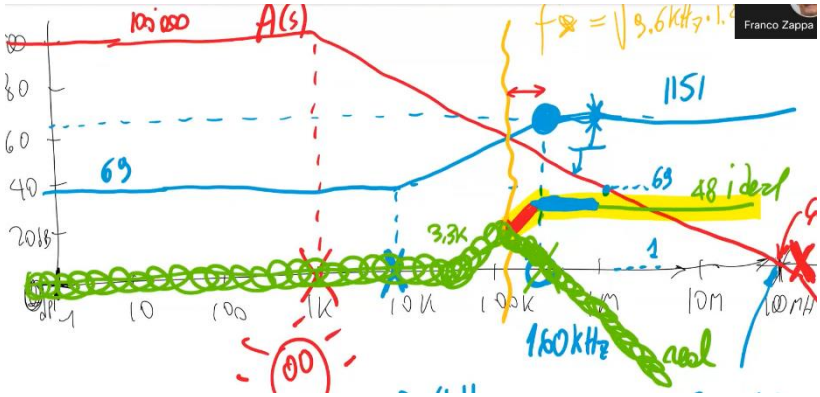


We see that the closure angle is not good, 40-40, even if there is a zero of beta that tries to help. We can compute the phase margin.

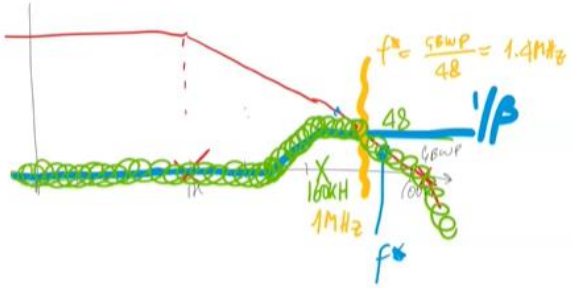


We are left with a phase margin of 36°, so the system is unstable. Let's now plot the real gain. Firstly, we need to plot the ideal gain.

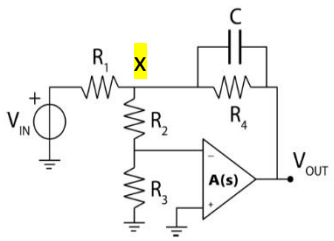
When we reach f* we drop by -20 dB/dec and then still -20 dB/dec after the zero of the ideal gain.



To solve the problem we need to let the circuit reach the HF gain. A solution could be to change to opamp to have the crossing between 1/beta and A(s) with 20-20. But it is better not to change the opamp, so we can reshape the beta so that the crossing with A(s) happens at the frequency I want. To decrease 1/beta we have to increase beta at HF, and to do so we should not kill beta at HF. At HF, beta dies because there is the capacitor C. Maybe, in this circuit it was an incorrect thing to add that capacitor.
To regain stability we can change R2, which is not impacting on the ideal gain, but it can change beta. Reducing 1/beta, f* increases, and to increase beta we have to increase R2. If we completely remove R2, the real gain will be the same, but beta will be 1 at LF and at HF it will be just the R3 and R4 partition, 48. So it is overlapping with the ideal gain.
Maybe now the situation is ok.



Example 3



Compensated OpAmp: A_0=120dB, GBWP=10MHz, I_B=10nA, V_{OS}=5mV. R_1=47kΩ, R_2=33kΩ, R_3=22kΩ, R_4=680kΩ, C=330pF.

a) Plot the **real** $v_{out}(f)/v_{in}(f)$ gain and comment stability.
b) Compute the output static errors due to the OpAmp.
c) Let the OpAmp be a Norton Amplifier instead, with A_i=5, compute the $v_{out}(f)/v_{in}(f)$ gain.

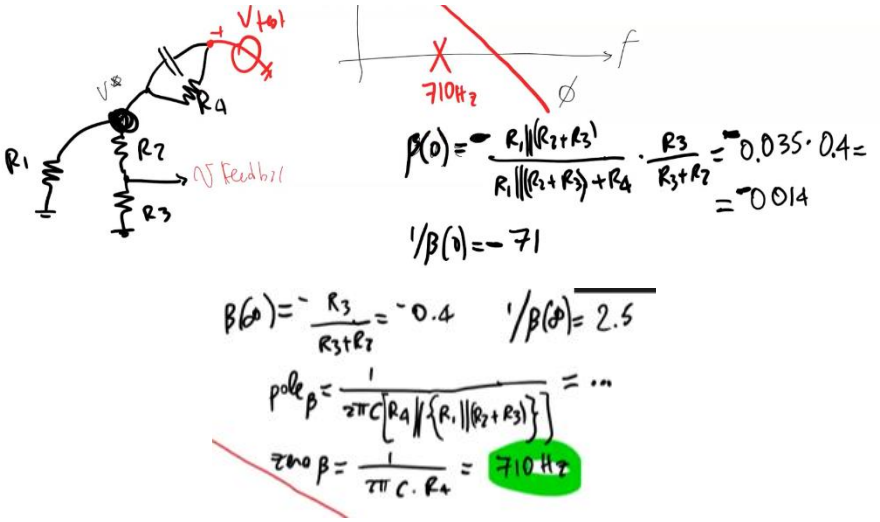
Opamp is still compensated.

a) Again we have a negative feedback. Because of this, I can say that I have virtual ground at – terminal. Hence no current flows in R3 and R2, so also node x is at 0V.

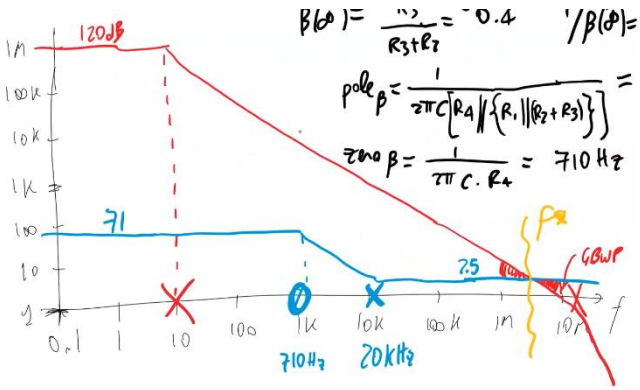
G(0) = -R4/R1 = -14.5
G(inf) = 0

The Bode diagram of this circuit is the one of a LP filter, so the circuit has no zero and one pole. The Req of the pole is C*R4. f_pole = 710 Hz.

As for the real gain, we need to compute the 1/beta(s). Of course everything is negative because we return to the – terminal of the opamp.

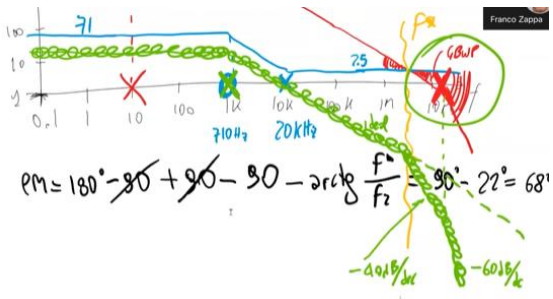


We can now plot the real gain. f* is very close to the second pole of the opamp, so we might be unstable. So let's compute the phase margin.



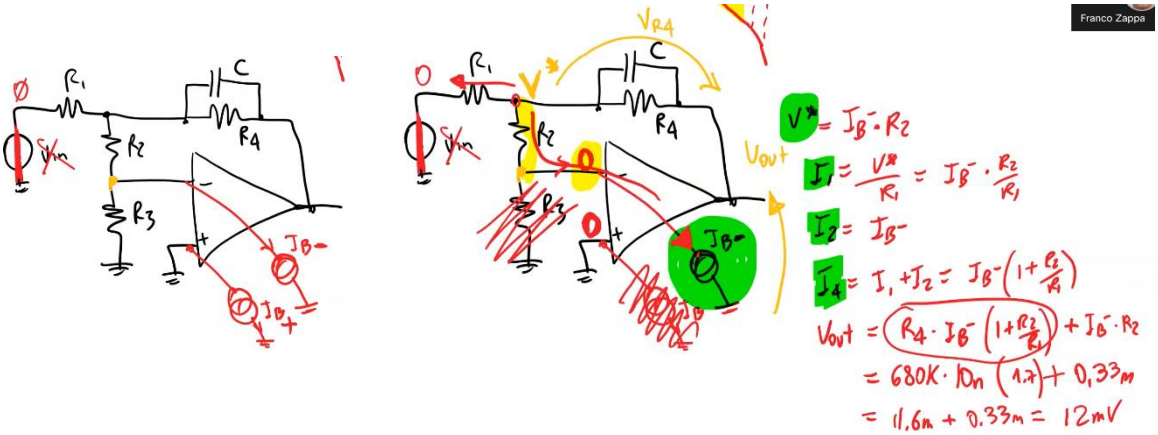
$PM = 180° - 90 + 90 - 90 - \arctg \frac{f^*}{f_2} = 90° - 22° = 68°$

The phase margin is ok, very good.



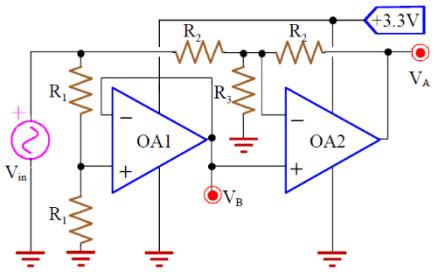
b) Output static errors of the opamp. Since the opamp has components inside, the impedance in input is not infinite, but there is some. Moreover, even if the input was infinite, maybe at DC we still need an input current for the opamp to operate, e.g. in the case of BJTs. The two bias current causes a current intake but the opamp, but still with infinite impedance.

When we consider the effect of the bias, we switch off the input and we consider just the bias effects, with the superposition of effects, hence considering one bias current at a time.



If I consider Ib+, the other one will be open circuit. Ib+ is pumped into a ground, so it doesn't give rise to any signal, so no error at the output. As for Ib-, due to virtual ground we have no current in R3. If so all Ib- flows in R2; if the opamp is sinking current, V* is positive. So we will have a current I1 in R1. Since the bias current is a DC contribution, the capacitor will be open. To kill this contribution, we should add a resistor, properly sized, between PS and V*.

Example 4

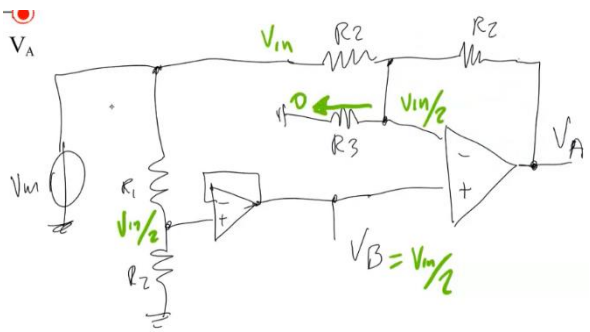


OpAmp with A_0=100dB and GBWP=100MHz. R_1=47kΩ, R_2=220kΩ, R_3=110kΩ.

a) Compute the **real** $v_A(f)/v_{in}(f)$ and $v_B(f)/v_{in}(f)$ gains and the input impedance.

b) Compute the output static error on V_A, due to I_B=10nA and V_{OS}=5mV of both OpAmps.

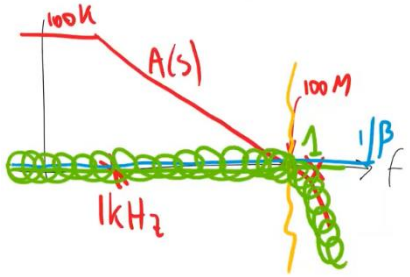
a) OA1 is a buffer that goes to the + terminal of OA2. So we have two local negative feedbacks, but is there a global one? No, because no one can counteract on the node where the voltage source Vin is placed. Let's redraw the circuit. The second stage has VG.



$$V_A = \left(\frac{V_{in}}{R_3} + \frac{V_{in}/2 - V_{in}}{R_2}\right)\cdot R_2 + \frac{V_{in}}{2} = \frac{V_{in}}{2}\left[\left(\frac{R_2}{R_3} - \frac{R_2}{R_2}\right) + 1\right]$$

$$G_{an} = \frac{V_A}{V_{in}} = \frac{1}{2}\left(1 + \frac{R_2}{R_3} - 1\right) = \frac{1}{2}\cdot\frac{R_2}{R_3} = +1$$

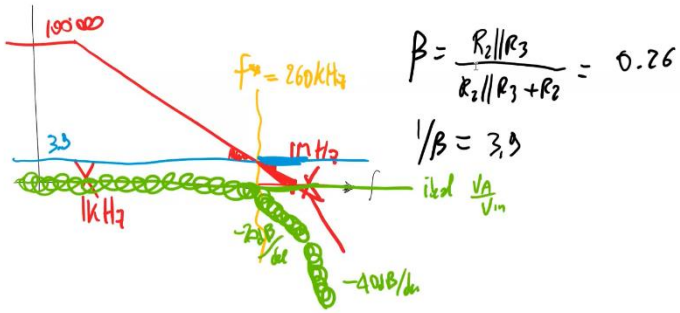
To compute the real gain of OA1 we know that it is a buffer, so I expect it to have a gain of 1, but it is not, because we have a A(s) and a beta(s). The beta for OA1 is 1, just the feedback network. The real gain is as below.



If the second pole of the amplifier is too close to GBWP I run the risk of having a phase margin poor, of 45°. For this reason we have a minor peaking.

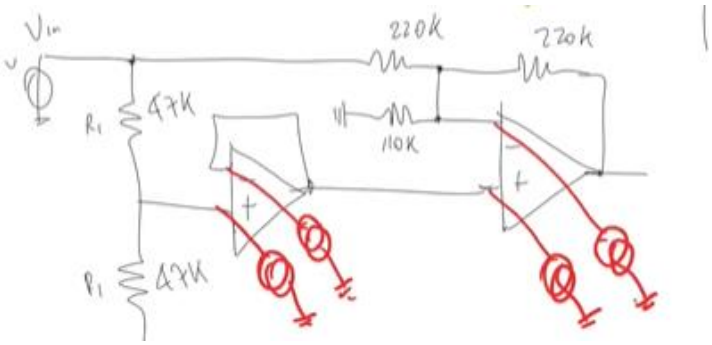
As for the real gain of Vb/Vin, it won't be 0.5, but at GBWP = 100 MHz it will experience two c.c. poles, so a minor peaking and a -40 dB/dec decrease. The two poles are due to a crossing 20-40 in OA1.

Let's now study the second opamp with its beta. f* will be at a frequency 3.9 times before GBWP.



Hence the real gain won't be flat, equal to 1 for almost all the frequencies of interest, but then there will be a first pole at f* and a second pole where the second pole of the opamp is.

b) Static error due to Ibias and Vos.

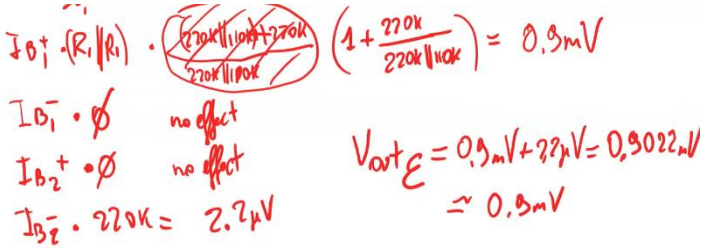


Let's compute the Vout due to the bias generator. Ib+ of OA1 have a current that flows in a parallel of 47k||47k (Vin is off). Once we have V+ of OA1, that will be also at the + terminal of OA2. That value get's amplified to the output with a gain of the non-inverting configuration. Eventually, the contribution due to Ib+ gives 0.9mV.

As for Ib- of OA1, the current comes from the opamp itself, and since the opamp is providing 0V at the output (Vin is 0), so that voltage doesn't move and because of this, the contribution of Ib- on the output is nihil.

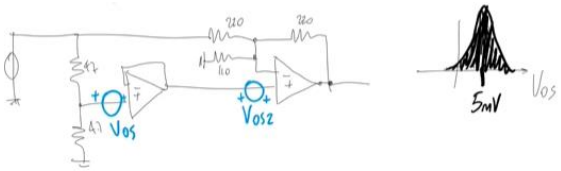
Let's now consider Ib+ of the second stage. Same reasoning as before because it pumps again in a voltage generator of OA1 output. Also for this current the contribution is 0.

Finally, for Ib- of the OA2, we have V- of OA2 at zero because of the propagation of the Vin = 0. If so, there is no current in the 220k and 110k resistors. Hence the Ib- is provided by the voltage source of the opamp and causes a voltage drop across the feedback resistor.

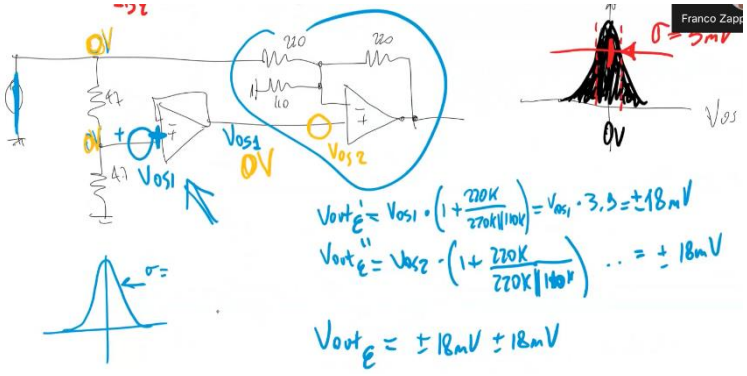


We can see that the opamp that mainly affects the offset in the case of the bias current is OA1.

As for the offset, it is due to the mismatch between components. Offset voltage is unpredictable, it can be positive in one or the other side, differently from the bias current that can be positive inward going or outward going. The Vos is a statistical error, with a statistical distribution (e.g. Gaussian) around a given value.

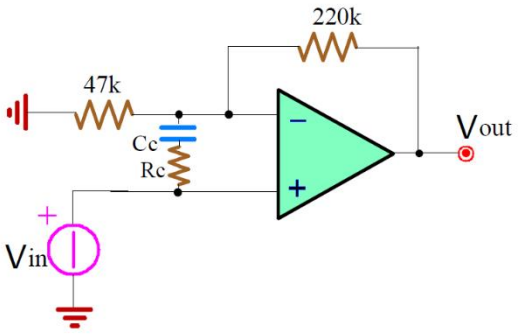


If we consider the offset of OA1 (input still off), we have an amplification due to the second stage, so the error is Vos1 * G2.
As for Vos2 it is exactly the same of Vos1, because in the second opamp enters just Vos2.
Total error is +-36mV.



We can notice that Ibias contribution is negligible with respect to Vos contribution.

Example 5



Uncompensated OpAmps with A_0=120dB, f_0=5kHz and f_1=50MHz, I_B=1nA and V_{OS}=3mV.

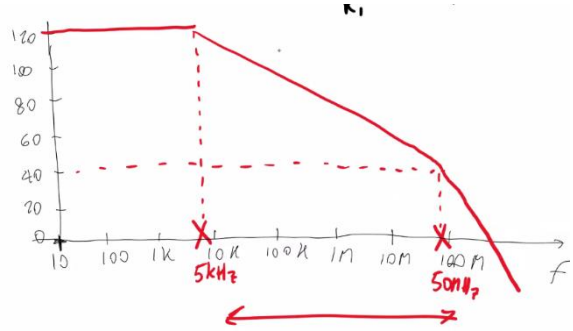
a) **Without** C_C and R_C, compute stability and PM.

b) Properly size C_C and R_C to attain PM=90°.

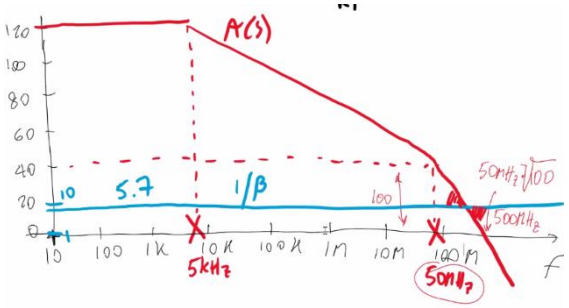
We start from the input signal, and then it touches node +, it gets amplified and the output is sent to the input → we have a feedback. Is it positive or negative? If + increases, the output increases, so due to the feedback network also – increases, so the error signal decreases and we have feedback → we can apply the concept of virtual ground, so the voltage on the + and – terminals are the same, so no drop on Cc and Rc.

Gideal = 1 + R2/R1 = 1 + 110/47 = 5.7

a) Let's compute the phase margin. The first pole is at 5kHz and the second at 50MHz, the opamp is uncompensated. Let's draw the Bode plot. The attenuation between first and second pole is 1000, that is 4 decades, so I will decrease by 4 decades in amplitude of the gain.



The crossing frequency is 50MHz*sqrt(100) = 500MHz. This in red is A(s). Now we have to study beta. Since Rc and Cc has to be neglected it is 5.7 and constant.



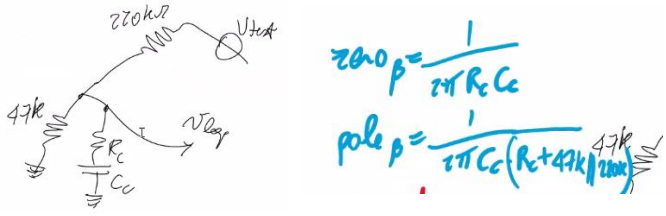
We can see that the crossing with A(s) is at -40dB/dec, so we can compute the phase margin. To compute f*, the crossing frequency, the ratio in gain is 100/5.7 but we have -40dB/dec, so f* = 50MHz*sqrt(100/5.7) = 209 MHz.

PM = 90 – 77 = 13° → not stable.

$$PM = 180 - 90° - \arctg\frac{f^*}{50MHz} = 90° - \arctg\left(\frac{20MHz}{50MHz}\right)$$

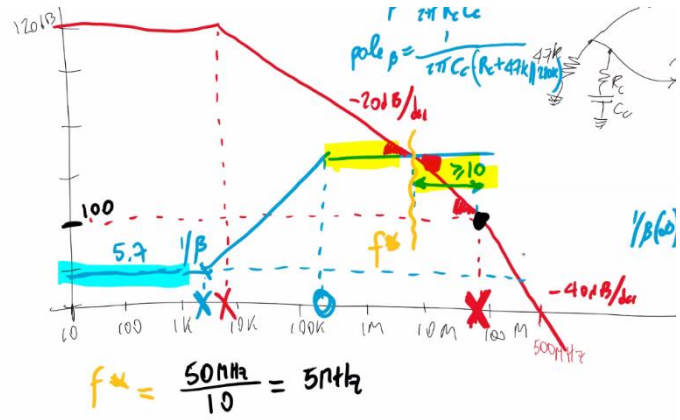
The ideal gain is 5.7, that just by chance overlaps with 1/beta, and let's plot the real gain, that is equal to the ideal up to f*, where it then follows A(s) because 1/beta is constant and A(s) is -40. If 1/beta was -20 after f*, the real would go with -20 after f* (REG).

b) Let's see the role of Rc and Cc. We are adding something that is not changing the ideal gain, because the voltage across them is ideally 0, but it impacts on the beta. In DC 1/beta is the same because the capacitor is open, but then we have a pole and a zero. Every time we enter in a circuit and we have a RC to ground, that RC introduces a zero = 1/(2*pi*Cc*Rc). To compute the pole, we switch everything off.



We have to size Cc and Rc to cross A(s) in the -20 region but not in a decade within the next pole, otherwise we are still experiencing the effect of the following pole. We have to stay one decade before the second pole of A(s).

The important thing is that the crossing is with a 0 slope, because if it is with a +20 slope for 1/beta we are not compensating anything.



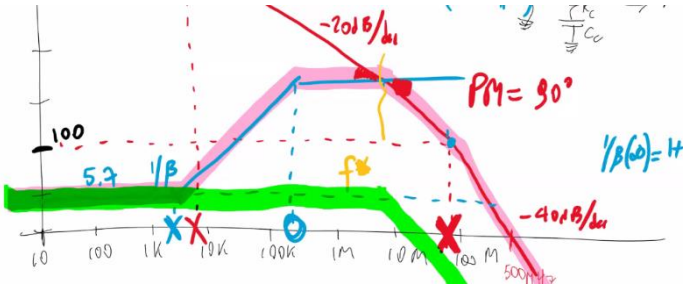
Moreover, as to the 1/beta at infinite frequency, should be 1k that is 10*Amin, because we know that the gain difference between the value at f* and at f of the second pole is 10.

$$1/\beta(\infty) = 10 \cdot A_{min} = 1K \simeq 1 + \frac{220K}{R_c}$$
$$R_c = \frac{220K}{999} = 220\Omega$$

Then since I want the zero one decade before to be sure that we don't suffer of the presence of the second pole, we can compute the value of Cc.

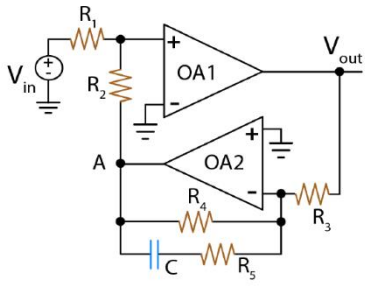
$$\frac{1}{2\pi C_c R_c} = zero\,p = \frac{f^*}{10} = 500kHz \qquad R_c$$
$$C_c = \frac{1}{2\pi \cdot 500kHz \cdot 220} = 1.4nF$$

In this situation, we are sure that the PM is more or less 90° (we are adding the two poles and then a zero of 1/beta). Then of course we have another pole, but we are interested in the phase shift only at the frequency where Gloop = 1.



The pole of this circuit is when the real gain decreases, that is at f*, and this is the pole of the closed loop configuration.

Example 6

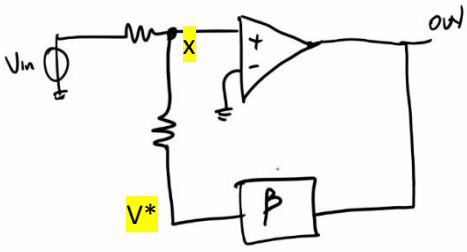


OpAmp with A_0=120dB and GBWP=20MHz, with I_B=10nA and V_{OS}=5mV.
R_1=1kΩ, R_2=50kΩ, R_3=2kΩ, R_4=50kΩ, R_5=1kΩ, C=10nF.

a) Plot the Bode diagram of the $v_{out}(f)/v_{in}(f)$ **real** gain, when OA2 is still **ideal**.
b) Discuss circuit stability when also OA2 is **real**.
c) Compute the output error due to bias currents and offset voltages of both OpAmps.

If we apply something positive, the output of OA1 goes positive, and then we have a standard inverting amplifier in the feedback, so the output of OA2 (stage with G = -25 in DC and -0.48 in AC) will go negative (node A) and hence we have a negative feedback because the negative on node A opposes on the positive increase in the + terminal of OA1.

OA1 is an amplifier with a very large gain because it is open loop, so the gain is A(s), while OA2 has a feedback that causes a gain that depends on the frequency. We have the following configuration.



The beta network is not just resistors in this case but it is an amplification, and since it is a negative amplification, the output can return to the positive terminal of OA1.

a) So OA1 has infinite gain (or at least A(s)), and if so, between its two input the epsilon is 0, so there there is the concept of virtual ground → + terminal of OA1 will be at 0V. If + terminal is at 0V, this voltage is set by the feedback, whatever we apply at Vin.
Thanks to feedback, we apply a signal Vin that tries to move node x but then Shrek moves the output so much that the beta stage will respond causing a negative signal to appear so that there is no movement at node x, remaining 0 (virtual ground).
A current will flow in R1 that is Vin/R1. Voltage V* = Vin/1k * 50k.

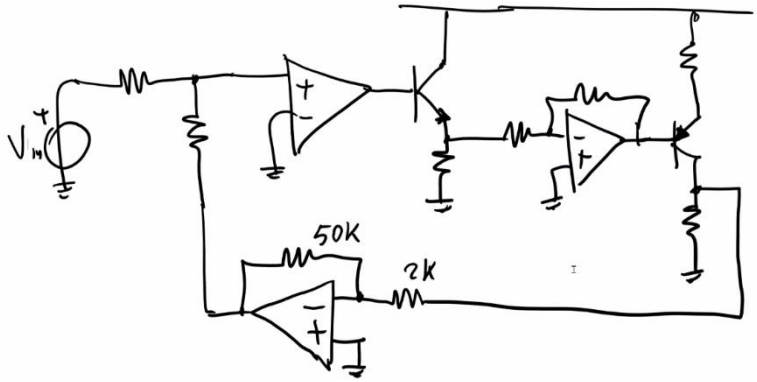
We are interested in Vout that is, at DC: V* = -25*Vout.
At AC: Vout = V*/-0.48.
As for the final transfer function:

$$\frac{V_{out}}{V_{in}}(0) = \frac{-50}{-25} = +2$$

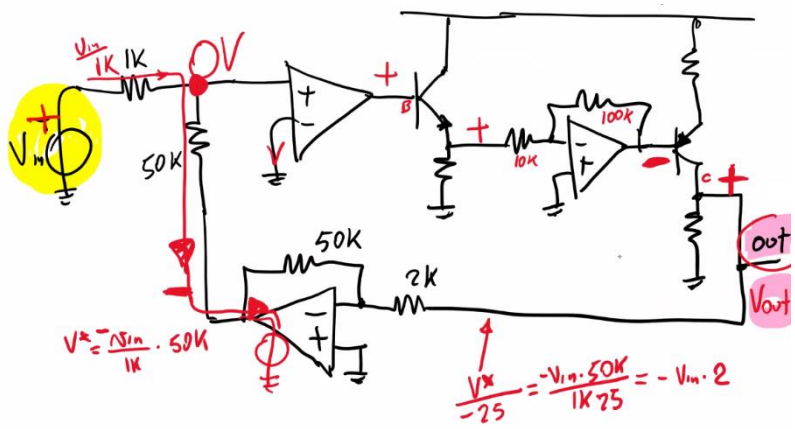
$$\frac{V_{out}}{V_{in}}(\infty) = \frac{-50}{-0.48} \simeq +100$$

Let's plot the ideal gain, it will have a zero and a pole. To locate them, we cannot exactly say where they are, since C causes something but not in a clear way. If there were a capacitor in series to R1, I would exactly know the position of the pole and zero.

To understand where pole and zero are, we must study stability. First of all, let's complicate the circuit.



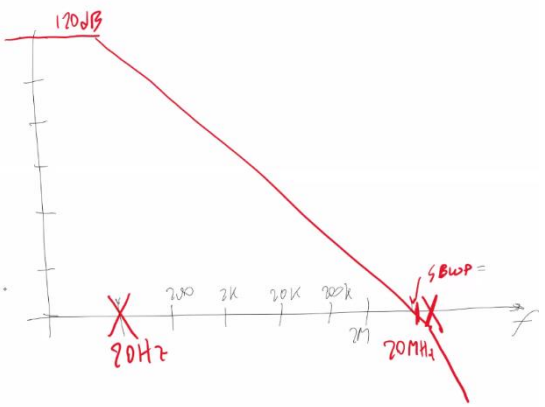
We start from Vin and then I move. I cannot go in OA2 output because it is a voltage source the output of an opamp, so I go in OA1. Given the +, the output increases, and since we are increasing the base of a transistor, the emitter is increasing as well (common collector configurator, emitter follower). But then we have an opamp with an inverting configuration. Then we have a BJT where we are entering in the base and exiting at the collector we are in an inverting configuration. The infinite gain is still related to OA1, because the gains of OA2, OA3 and the two BJTs are limited. So on OA1 applies the concept of virtual ground. Let's compute the gain. The current in input flows in the OA2, so V* = -Vin/1k * 50k.



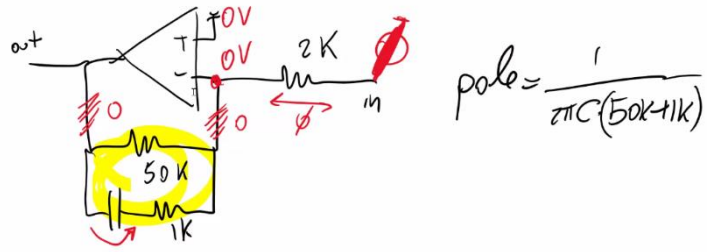
So we don't need to pass through the upper network to get the output. The signal propagates firstly in the upper branch and then in the feedback, but we can analyze it in the way we want.
So I don't care if the opamp is ideal or if there are non linearities, I don't care about what A(s) is (indeed, A(s) is not necessarily only the one of the single opamp, but it can be the whole forward branch), because the important thing is the beta network, that is the one that sets the real gain

(from the output back to the input). We can choose whatever beta and A(s) network we want, the important thing is that the f* is always the same, as well as for the close angle in the crossing between 1/beta and A(s). Now we return to the original circuit.

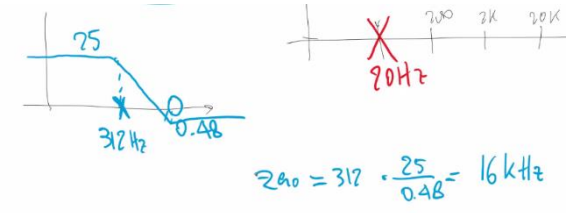
OA1 is the open loop opamp and let's call it A(s). It is uncompensated.



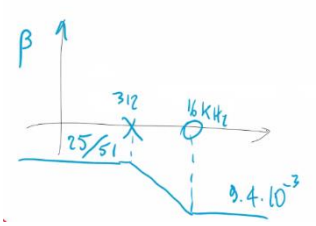
Now let's plot beta. We cut at the output of OA1 and place Vtest. Then we have the gain of the classical inverting configuration. We have also a pole and a zero because we have the C capacitor. The pole is not the one of the overall circuit, nor of the beta, but of the inverting configuration in which OA2 is. In the CL config of the inverting, we recall that the pole is given by the two resistances in series (we have to switch of all the input signal when computing the pole).



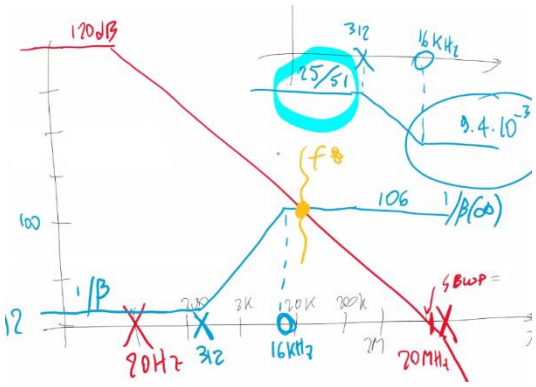
fp = 312 Hz.
As for the frequency of the zero, it is a factor 25/0.48 higher than the pole.



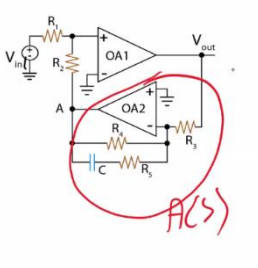
Apart from the gain of the inverting stage, the beta network is also composed by the resistive divider that we have in output to the inverting stage. So the gain of OA2 must be multiplied by an attenuation factor R1/(R1+R2) = 1/51.



Overall, the beta network has a gain of 25 * (1/51) in DC and 0.48 * (1/51) in AC.

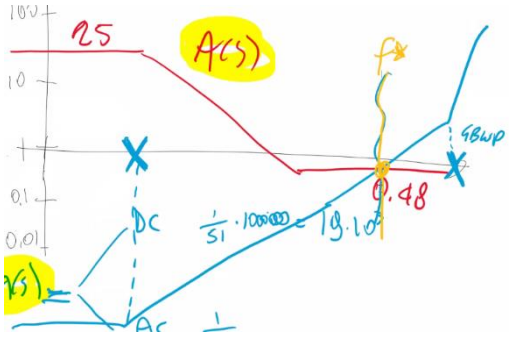


But I don't want beta, but 1/beta, so I have to revert all the gain values. Now we have to compute the f* = GBWP/106 = 189 kHz. The zero is at lower frequency, so the plot is correct. If f* was 189kHz and zero was at a higher frequency, the plot was wrong.

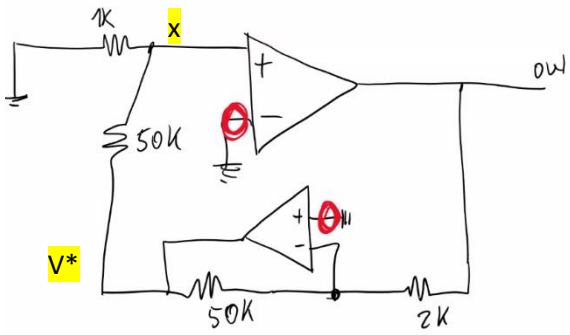


We could also study the circuit in another way. Now I change what is A(s) and what is beta. In fact, if we switch off the source, output disappears, so the loop is a loop, so we can define whatever we want.

In this case, A(s) has a gain of 25 at LF and 0.48 at HF. Now we have to plot beta, that is firstly an attenuation due to the resistive divider, and then the t.f. of the opamp. In DC, beta = 1/51 * 1'000'000 = 19'000, while in AC it dies as the t.f. of the opamp. We want the inverse of beta, of course. DC: 51*10^-6. Even with this method, I find f* that is exactly the same of the previous case.



b) (It is request c).
Bias currents are a DC concept, so the capacitors are open. Offset can be put where I wish, it is easier if placed as in the image.

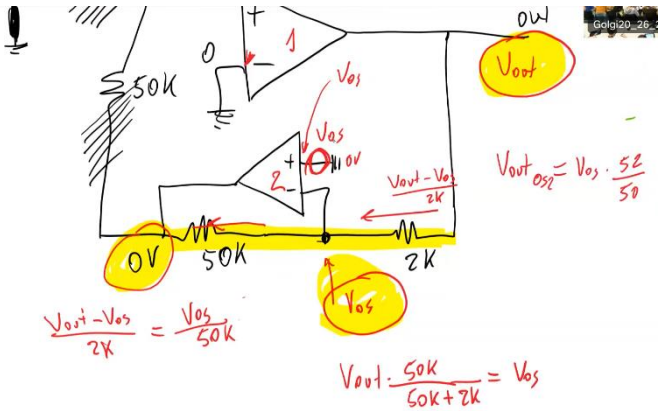


Let's consider Vos,OA1. Vin is off, so node x I have Vos. So a current flows in R1 and through also R2. Then it goes to the source that is the OA2, whose output generator is providing the current. V* = Vos/1k * (50 k + 1 k).

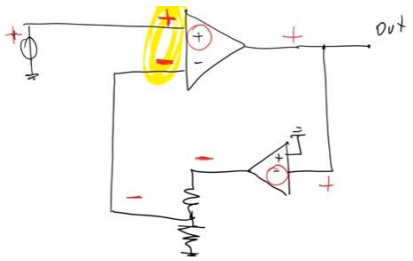
Hence the output error due to Vos1 is Vos1*2 more or less.

(1:26:10)

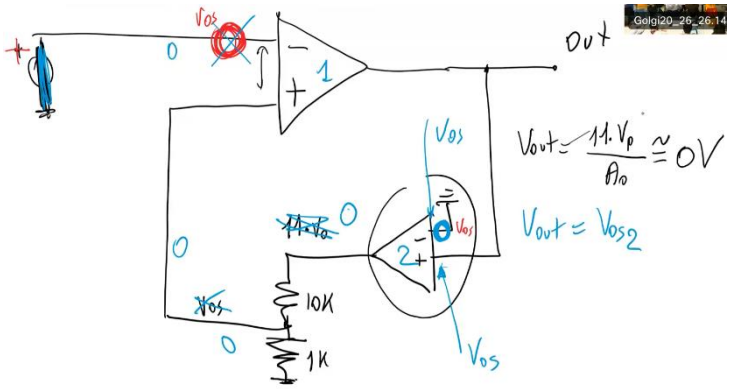
Let's consider now the second offset related to OA2. The typical error is that we suppose the Vos propagates clockwise in the network to the output with all the different values of gain and everything, but to investigate the circuit I have to move the other way round.
So the input is 0, so I have no current in the 1k resistor and no current in the 50k resistor. Hence output of the feedback opamp is 0. We will have a current in the 50k resistor due to Vos. And this current must be equal to the one in the 2k resistor.



Let's consider the following circuit.



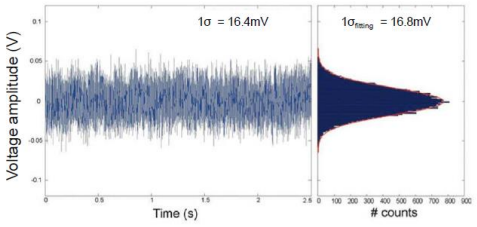
There is not a negative feedback, the feedback is positive. To have a positive one it's sufficient to switch the input terminals. The contribution of the offsets of the two opamps are the following.



NOISE

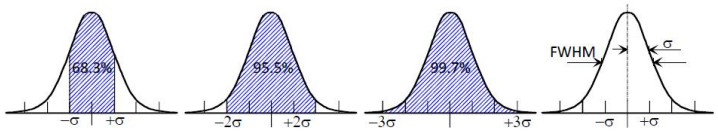
Noise is a disturbance that can be considered white or not. It is a problem both in analog signals and digital signals. In fact in digital signals it can add time jitters.
Noise is random in its nature, it is not deterministic; most of the noise we are dealing with have an average value equal to 0, so it is no sense to consider the specific value of noise at a specific time instant, but better to speak about noise distribution.

Noise often follows a Gaussian distribution. The full area of the Gaussian is equal to 1.



Distribution: $p(x) = \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(x)^2}{2\sigma_x^2}}$

Note that: $\int_{-\infty}^{+\infty} p(x)dx = 1$



Peak-to-peak value: $x_{99.9\%} = \pm 3 \cdot \sigma$

Full-Width at Half Maximum: $FWHM = 2.35 \cdot \sigma$

Hence the random event has a probability to have a given number and the sum of all the probabilities is equal to 1. However, often the noise is concentrated within a specific range. This is according to the bottom image. In a gaussian it is no sense to speak about a peak-to-peak value, but whenever we have a random noise that can be described as a gaussian, we can consider the FWHM. Almost all the noise will be within +- 3*sigma. The peak-to-peak value is in within this range. So given the sigma, we can define the FWHM of a gaussian is 2.35*sigma.

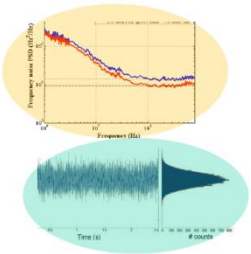
Noise can be described in the time domain or in the frequency domain.

Instantaneous value is nonsense, better to measure the power: $Power = \frac{1}{T} \cdot \int_0^T |x(t)|^2 dt$

"Ergodic" process (time average = samples average), Gaussian, with nil mean value

Variance, "power", mean squared value: $< x^2(t) >= \sigma^2$

Parceval's theorem: $\sigma^2 = < x^2(t) > = \int_0^\infty S(f)df$



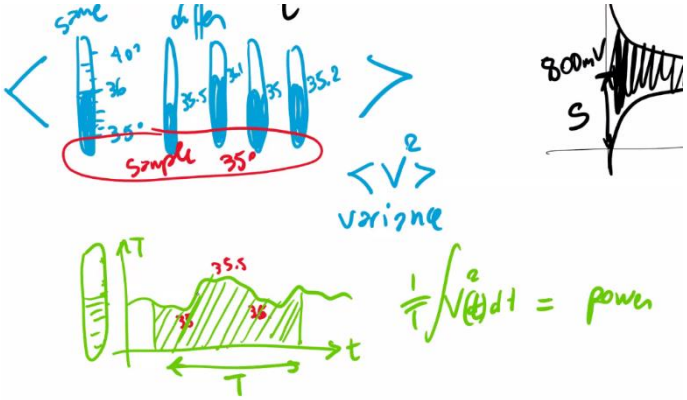
Root mean square, *rms*: $x_{rms} = \sqrt{\sigma^2}$

The power of a signal can be computed in two ways. We can integrate the value of the voltage measured between 0 and time T in dt and divide by T to get the voltage average value. For a noisy signal, the average value is 0. If we have a gaussian centered around a value, e.g. 800mV, the average value is 800mV.

If we compute the integral of the square of the voltage between 0 and T and then we divided by T, we don't get the average voltage but the average in the time domain of our signal. This is also called variance of the signal, average power of the signal.

$$\langle V^2 \rangle = \frac{1}{T}\int_0^T |v^2(t)|\,dt$$

Variance is a value related to a set of values. Imagine we have a thermometer with a possible measured value between, 35° and 40°, and we take the measurement with different (or the same several time) thermometers of the same quantity, and then we average them, we get the variance of the experiment.



Conversely, we can take one thermometer and see how the temperature varies over time. If we compute the area and we divide by T, we get the power of that signal.

If we consider ergodic processes, the time average is equal to the samples average, so the two previous quantities in the examples are the same. Moreover, the Parceval's theorem says that if we have a signal, if we have the time dependent waveform of the noise and the spectrum of the noise, if we compute the power in the time domain, that is the variance, is equal to the integral of the spectral density of the system.

If we compute the square root of the power we are computing the root mean square. Sigma is the quantity that if we take 3 sigmas on the right and on the left of the gaussian curve and we compute the area, we get the 99.7% of finding the signal in that range. In fact, the peak to peak amplitude of the signal is more or less 6 times signa.

Two noise sources: $v_t(t) = v_1(t) + v_2(t)$ x

Mean total value: $< v_t(t) >= 0$

Total variance: $< v_t^2(t) >=< [v_1(t) + v_2(t)]^2 >=< v_1^2(t) > + < v_2^2(t) > +2 < v_1(t)v_2(t) >$

… in case of NO correlation : $< v_t^2(t) >=< v_1^2(t) > + < v_2^2(t) >$ Such as the effect overposition principle

… in case of TOTAL correlation $v_1(t) = v_2(t)$: $< v_t^2(t) >= 4 < v_1^2(t) >$

100% ERROR ! but … NEGLIGIBLE ! ("only" 41% on rms value)

Therefore… let's consider all noise sources as uncorrelated

CORRELATION AMONG NOISE SOURCES

Let's suppose we have a resistor. The resistor is noisy, with a spectral density that is a number, that in the case of 1 kOhm resistor at RT is 4nV/sqrt(Hz).
If then we take another resistor that is noisy and we put it I series, the equivalent resistor has a value of 2k, but the noise is not 8 nV/sqrt(Hz), we don't get the sum, because it is the power and variances that sum up.

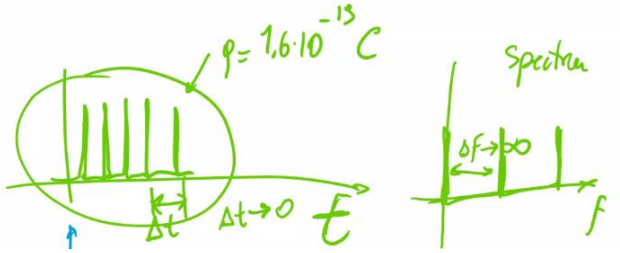
$$\sigma_{Tot} = \sqrt{\sigma^2 + \sigma^2} = \sqrt{(4n)^2 + (4n)^2} = 5.6nV$$

We might have another issue. If we study a circuit, in the circuit we may have different components, resistors, opamps, capacitors and so on. Each component (except for capacitors) will have its own noise, and we should consider the sum of all the noise contribution. But how to do so? Should we sum the amplitude or the power?

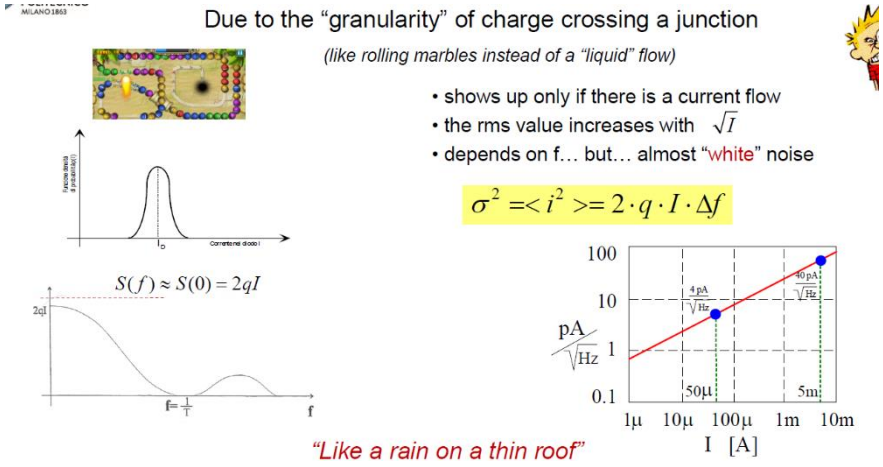
If we have a signal that is the sum of two signals, like in x. Mean total value is 0, because both the integrals are zero, since the two processes are random noises. But then when we compute the variance, we have to compute the power 2, that is not just the sum of the power of the two signals, because there is also the third contribution. **In case of no correlation**, the third contribution is negligible, so **we can apply the principle of the superposition of effect in the powers, not on the individual noise**.
If signals are correlated, the sigma of the two signals is the same, and the overall sigma_tot is two times the sigma of the signal. Even if we had, by chance, considered the two signals uncorrelated, the error committed in doing this is not so big, so in principle we can consider all the noise sources uncorrelated.

SHOT NOISE

It is due to the granular nature of charge crossing the junction. The charges passing across the channel varies depending on the time instant at which we are looking the situation. We would see several deltas related to the passage of the electrons, and in the frequency domain this still remains with a comb shape.

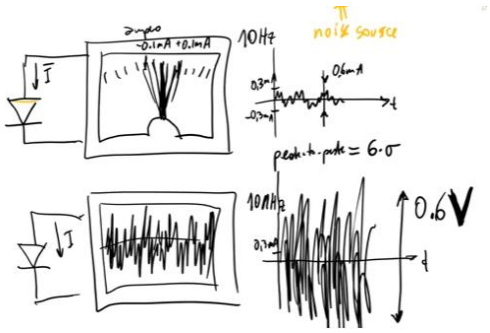


Electrons are random in movement due to Brownian motion of charge, so the real spectrum in the frequency domain will be similar to a cardinal sin of events, which has a zero crossing in the thousands and thousands of GHz → where we study the noise we can consider the shot noise to be flat.



This noise increases with the amount of current. If the current is low, the noise is low. If the current increases, then even the noise increases. The proportionality factor of increase is 2*q.
The shot noise 2qI is not depending on frequency, nor on a resistor value, it depends just on current. Of course, it is proportional to the bandwidth, we need to multiply by delta_f of the instrument that is measuring noise to get the sigma square. But the noise contribution is just 2qI.

So e.g. we have a diode through which an average current is flowing, but we measure the current with two instruments; the first one could be an amperemeter, and the second one could be an oscilloscope. The oscilloscope input has a much higher bandwidth with respect to the amperemeter (e.g. 10MHz vs 10Hz of the latter). So if we look at the same noise on the oscilloscope we would see a much broader nose, a peak to peak equal to 0.6V.

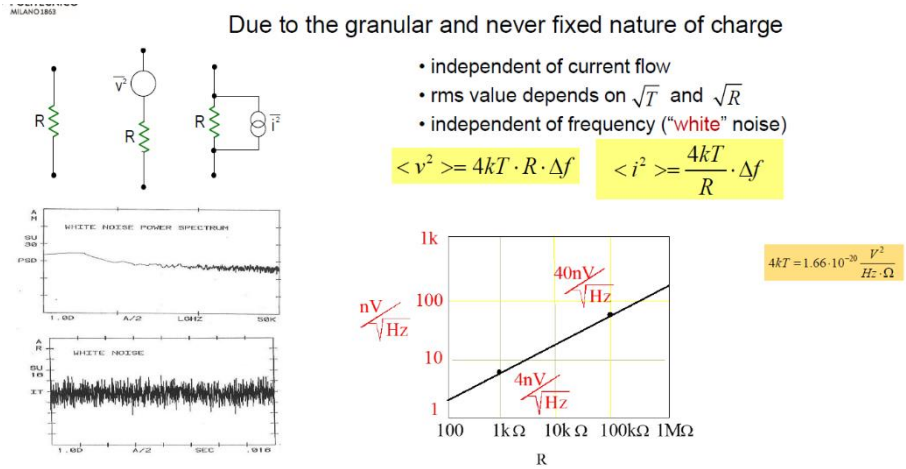


Hence depending on the bandwidth of my instrument we will have the total power of our noise. Hence if the average value of the current increases, also the noise will increase, but if I consider signal over noise (SNR), this increases with sqrt(I).

NB: let's remember that for a 50uA current the corresponding shot noise is 4 pA/sqrt(Hz).

THERMAL (JOHNSON) NOISE

Let's consider a resistor and no current in it. I should see 0V across the resistor, but instead I see some voltage fluctuations across it due to the Brownian motion of electrons. This noise increases with temperature and value of resistance.

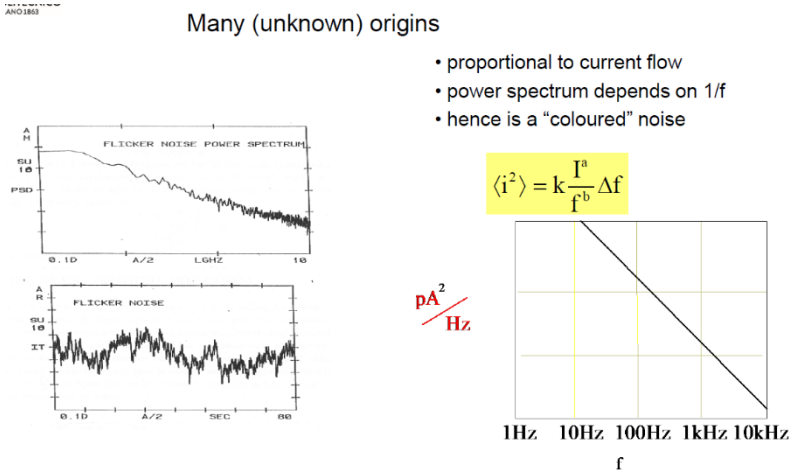


4kTR is the intrinsic noise of the resistor. 4kT = 1.66+10^-20 (at RT), and it is similar to q, that is the electron charge. Then we have also to consider the bandwidth.

Again, if we increase R the power increases linearly but the rms value increases by the sqrt(R) (this is in terms of voltage noise).
But if we look at the noise in terms of current, it seems to decrease with an increase of R. But since the power (v^2) increases by R, it makes sense that the current decreases with R.

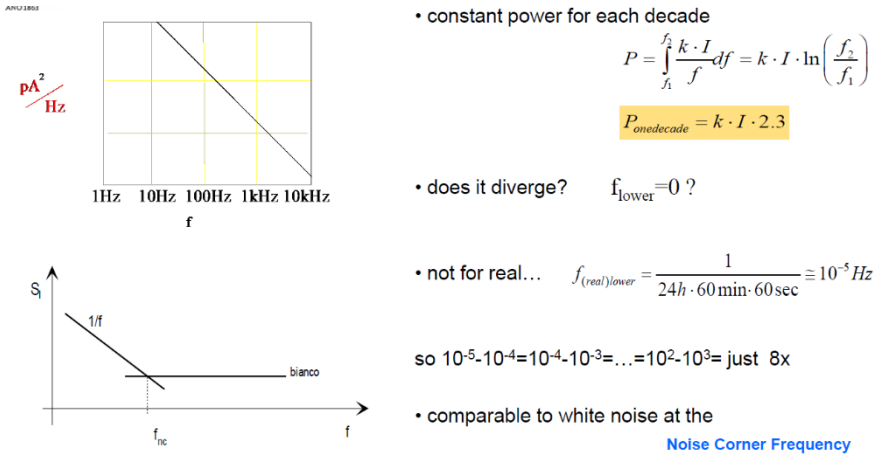
NB: for a 1k resistor the noise is 4 nV/sqrt(Hz).

FLICKER NOISE (1/f)



It has many unknown origin and it is proportional to the current flow. Its power spectrum is not constant, it is not white because it decreases with frequency. Usually coefficients a and b are 1, but they can change a bit depending on the real physical mechanism.

The important thing to compute is the **power within one decade**. If we compute the power within one decade of that spectral density, that is kI/f, k and I are constant so what remains is the integral of 1/f that is the logarithm. The flicker noise is not constant, we have not to multiply it for delta_f as in thermal or shot noise.
Once we found k and the current in the device, then the power within a decade is P = k*I*2.3.

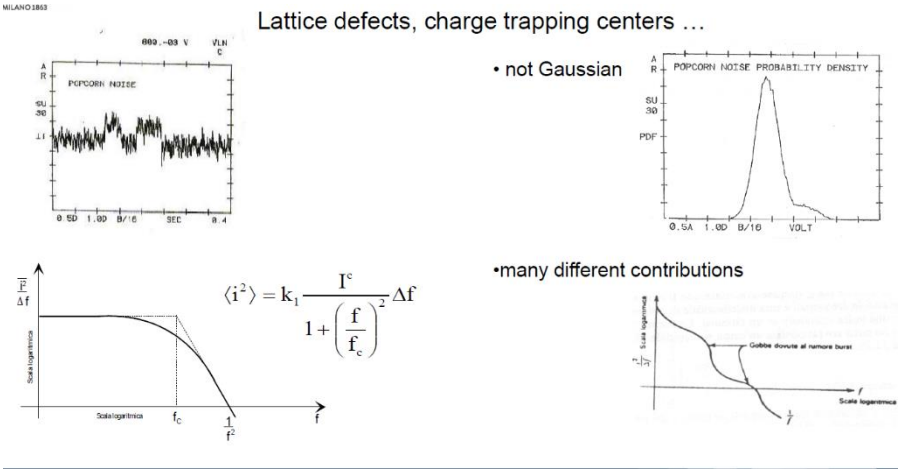


Does this noise diverge?
In theory yes, because if the lower frequency is 0, the logarithm will diverge, even though practically it doesn't. In fact it doesn't make any sense to set f_low to be 0, we will have never 0. Then once we have also set the upper frequency, we count the number of decades that we have in between the two frequencies and then we are done, we multiply the P value for that number and we have the flicker noise.

So we have a contribution that varies with frequency due to flicker noise and another that is constant, the thermal and shot noise (white noises). Hence to compare which noise is dominating, instead of

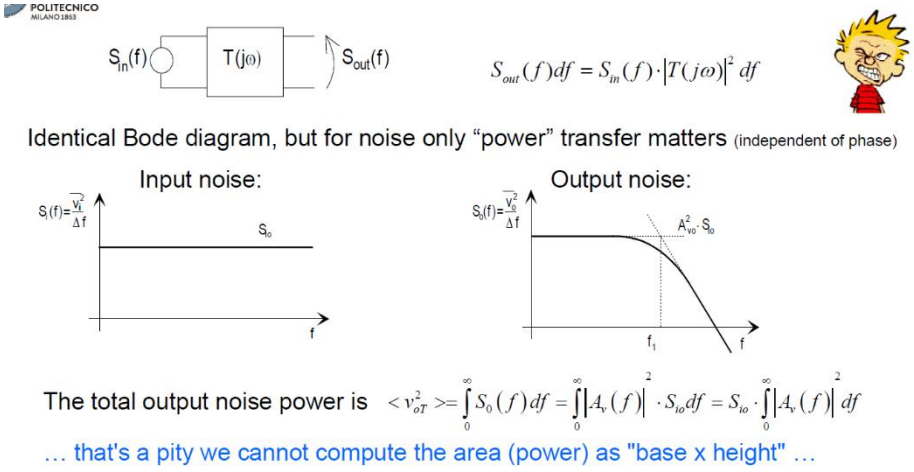
computing the coefficients of the 1/f noise and the level of the white noise, we can quote the corner frequency at which the two noise spectral densities are equal.

BURST NOISE



It is a noise that keeps moving up and down, usually between two levels or more. If we plot the probabilistic density of this noise it's not one gaussian so there are also a secondary gaussian and so on. Hence it is a more complex noise that can be modelled as a constant contribution for low frequencies and then at HF the noise drops as (1/f)^2. It is not like flicker noise because it doesn't go down as 1/f and, compared to flicker noise, it is constant at LF.

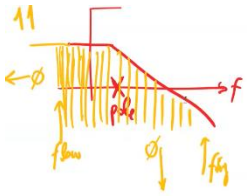
NOISE I/O TRANSFER FUNCTION



We have a noisy source with its noise spectral density and an amplifier with its amplification gain. If the input is a voltage source, the out is the input multiplied by the gain. But since we are considering noise, we have to multiply by the square modulus of the gain (power gain), because we are considering power.

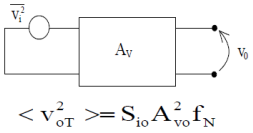
If the input noise is white, the output noise is white if the amplifier has no poles. But if the amplifier has a pole, the output spectral density will show the typical drop due to the pole of the amplifier. The pole causes the voltage gain to drop by -20dB, but since we are in power gain we have -40dB.

Eventually, if the input sigma is constant, the output sigma may not be constant, because the typical Bode diagram of the amplifier will have a pole. Since soon or later we will need to compute sigma_out, it means that we need to compute the square root of the output power, which depends on the input power that is frequency dependent and that should be multiplied by the transfer function of the amplifier to the power of 2, everything then under the integral.

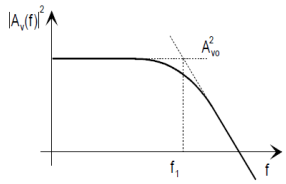


This is equivalent to computing the area of the right image. In computing the area, if the amplifier has 1 pole and we integrate from 0 to inf, we end up with 1,57*f_pole. This means that if we have an amplifier whose gain is 1 and has a pole, we can compute the integral that is equal to the height of the gain multiplied by a base that is at a frequency given by f_n = 1,57*f_pole. This is the noise equivalent bandwidth for a single pole amplifier.

NOISE EQUIVALENT BANDWIDTH



$$< v_{oT}^2 >= S_{io} A_{vo}^2 f_N$$



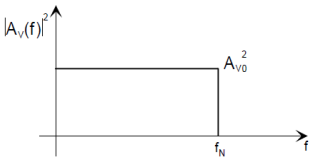
The pole is NOT the noise bandwidth f_N=Δf to consider !

Noise Equivalent Bandwidth: $f_N = \frac{1}{A_{vo}^2}\int_0^\infty |A_v(f)|^2 \, df$

Very simple cases:

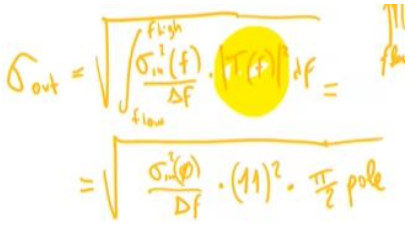
1 pole: $f_N = \int_0^\infty \frac{df}{1+\left(\frac{f}{f_1}\right)^2} = \pi \cdot \frac{f_1}{2} = 1,57 \cdot f_1$

2 poles: $f_N = 1,22 \cdot f_1$



OK then, the rms value is SQRT("base x height") of such an equivalent rectangle

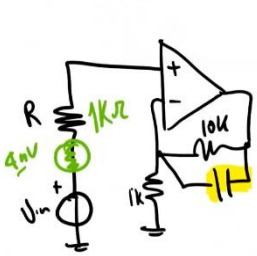
Hence to compute sigma_out we can consider the sigma_in at 0Hz, the gain at 0Hz and then multiplied by the noise equivalent bandwidth.

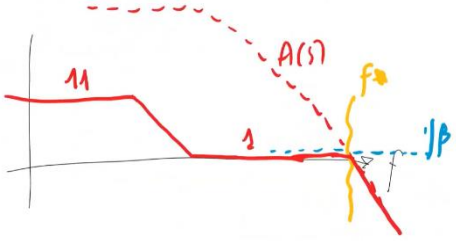


So we just need to consider the input noise, the power gain at DC and the noise equivalent bandwidth.

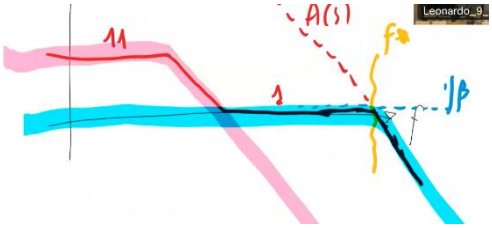
If we consider infinite poles happening at f_pole, than f_n = f_pole, because the decay slope is -inf dB/dec, so we have a rectangle.
If instead the poles of the amplifier are at different frequencies, we should overestimate the noise, considering the bandwidth limited just by one pole, that is the dominant pole at the lower frequency.

In some circuits we may have an amplifier with a capacitor, like below.

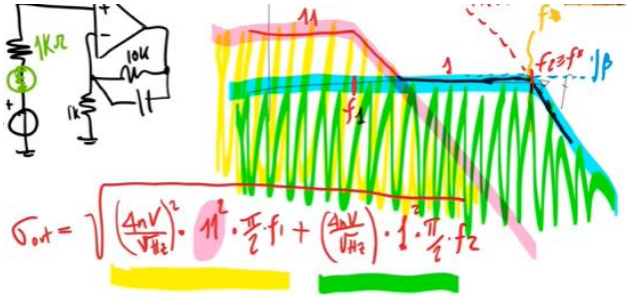




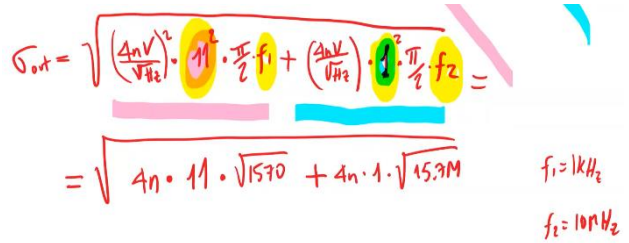
In this case I also have a zero. In this case we consider the Bode diagram as composed by two different Bode diagrams.
One that is the pink one and one that is the blue one.



Hence we will have like two poles and again sigma_out will depend on the two contributions.



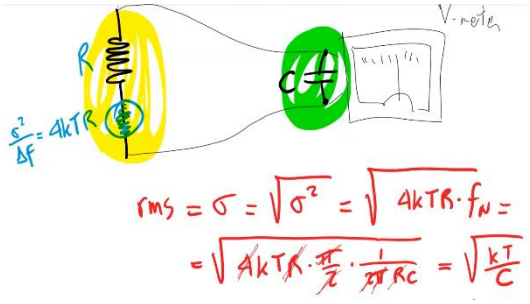
So the first contribution has a large gain but low noise equivalent bandwidth, vice versa for the second contribution. Eventually, instead of computing the power two of the members, we can take the, as power 1 and then the square root of the bandwidth.



One of the two contributions may prevail and the other one could be negligible.

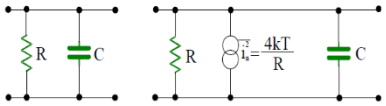
Example

Let's apply the noise equivalent BW concept to a very simple amplifier. Let's consider a noisy resistor and we want to measure it with an instrument, e.g. a voltmeter. The voltmeter will have its BW due to the parasitic capacitance C. Let's compute the output rms noise, but to do so we need to compute the noise equivalent BW of the circuit. To do so, it will be pi/2 divided by the pole, which is 1/(2*pi*R*C). The result is the one below.



Hence the noise in output of the instrument will be proportional to C and no longer proportional to R. This doesn't mean that the C is the noisy component, it is the fact that the rms noise depends on the BW that introduces the C in the noise computation, because the noise increases with R and the BW reduces with R, so the product is independent on R.

Simple RC network:



Full computation:

$$T(j\omega)=\frac{R\cdot\frac{1}{sC}}{R+\frac{1}{sC}}=\frac{R}{1+sCR}$$

$$S_0(f)=\frac{4kT}{R}\cdot|T(j\omega)|^2=\frac{4kT}{R}\cdot\frac{R^2}{1+\omega^2C^2R^2}$$

$$<v_{0T}^2>=\int_0^\infty S_0(f)df=\frac{4kT}{R}\cdot\int_0^\infty\frac{R^2}{1+\omega^2C^2R^2}df=$$

$$=4kTR\cdot\left[\frac{arctg(\omega CR)}{2\pi RC}\right]_0^\infty=\frac{4kTR}{2\pi RC}\cdot\left(\frac{\pi}{2}\right)=\frac{kT}{C}$$

"Equivalent" computation:

$$<v_{0T}^2>=4kTR\Delta f=4kTR\cdot\frac{1}{4RC}=\frac{kT}{C}$$

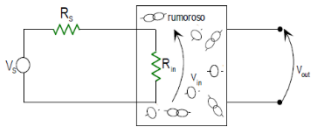
for C=1pF we get 65μV_{rms}

So what? Does rms noise depends only on C? Is C noisy instead of R?

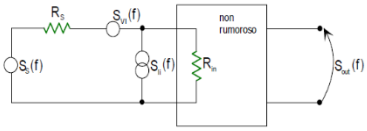
Hence the larger the C the lower the noise, because if C increases the BW decreases.

NOISE EQUIVALENT GENERATORS

Real circuit:



Equivalent circuit:



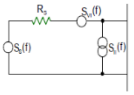
Total input-referred noise (no R_{in}):

$$S_{in}(f)=4kTR_S+e_V^2+i_I^2\cdot R_S^2$$

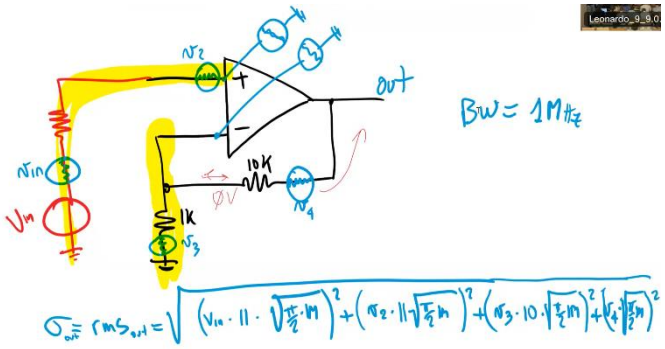
Total output noise:

$$S_{out}(f)=|T(j\omega)|^2\cdot\left\{[S_v(f)]\cdot\frac{R_{in}^2}{(R_S+R_{in})^2}+S_I(f)\cdot\frac{R_{in}^2R_S^2}{(R_S+R_{in})^2}\right\}$$

No need to compute the output noise… let's stop at the input!

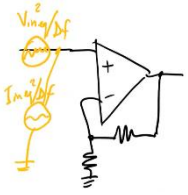


Instead of using the individual noisy sources in a schematic, it is a great advantage to consider the noise equivalent generators. For instance, we may have an inverting stage with all the noises related to the resistances, and current noises in the amplifier. The output total noise is due to the contribution of all of them. We can compute the contribution of all the noise sources to the output.

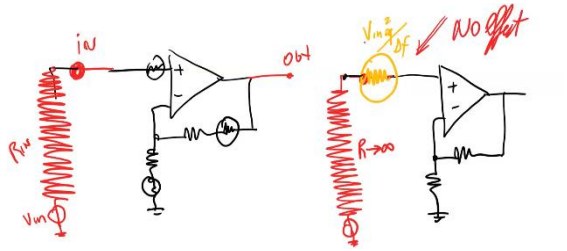


If we change the value of resistances, but keeping the same gain, what will change is the contribution of V3 and V4. If we change Rin, then the input contribution will change.
So if we leave the noise sources spread along the circuit is very difficult to decide how to optimize the signal and the circuit.

The best solution is to consider the stage to be ideal, noiseless, and we consider the noise sources to be applied at the input. So we have a voltage input equivalent noise source and a current equivalent noise source. Then we connect our source.



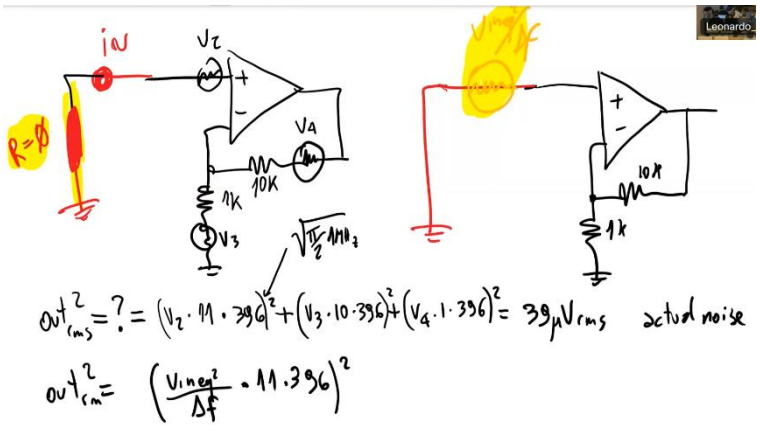
Modelling all the noise sources with just a voltage or current equivalent noise generator is not enough, because if for instance at the input I place a very large input resistance, if R increases up to infinity, the noise generator would have no effect because at the output I'm not finding any value because the node to which the voltage noise generator is attached is floating, so at the output I find no noise. Of course this is not true because the circuit has noise.



So the other possibility is to use just a current generator for the equivalent noise. But this is still a problem if we have a source with a very low R, because noise current recirculates in the shortcircuit and no noise is presented at the output. So we need to use both the generators.

Voltage equivalent generator

To compute it, we choose a R of the source equal to 0. If so, the current noise generator dies thanks to the short and it has no effect on the circuit. So we need to shortcircuit the input to compute the voltage equivalent generator. Then we compute the total output rms noise in the real circuit; then we shift to the model circuit and the output noise for the model circuit is the second line.



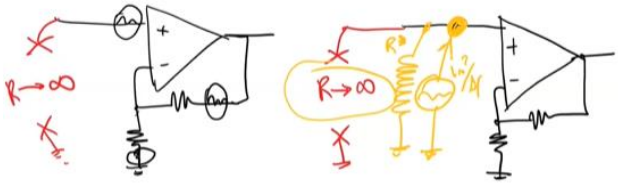
Now, if we want the model circuit to be equal to the real circuit, we have to equate the two expressions.

$$\frac{V_{in\,eq}^2}{\Delta f} = \frac{(39\mu V_{rms})^2}{(11 \cdot 396)^2}$$

The unit of measure is (nV/sqrt(Hz))^2.

Current equivalent generator

In this case, the input is floating, so that the voltage noise equivalent generator has no effect. Then the computations are the same than in the case of the voltage generator.
However, there may be situation where the circuit is "bad", like in the image below. In fact, Iin pumps into a high impedance node, infinite impedance, so that node moves to infinite voltages, so the contribution is really dramatic. So we introduce a high value resistance, we don't consider it infinite.



Of course, only in this case the equation could diverge and we need to introduce the resistance R*. Then we have to compute the limit.

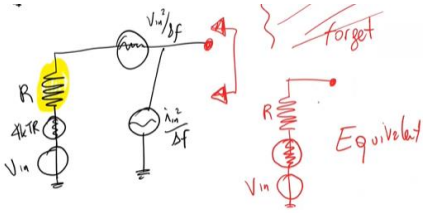
$$\text{output}_{rms} = \left(V_2 \cdot 11 \cdot 396\right)^2 + \left(V_3 \cdot 10 \cdot 396\right)^2 + \left(V_4 \cdot 1 \cdot 396\right)^2 = 50\mu V$$

$$\text{output}_{rms} = \frac{i_{in}^2}{\Delta f} \cdot R^2 \cdot 11^2 \cdot 396^2$$

$$\lim_{R^* \to \infty} \quad i_n^2/\Delta f = \left(\frac{50\mu V}{R^* \cdot 11 \cdot 396}\right)^2 = 0$$

We see that in this circuit, just by chance, the current equivalent noise generator has no effect. Hence this specific circuit can be modelled with just one noise equivalent generator, but this is just by chance. In fact, if we had considered the noise bias current of the opamp, the limit would be different from 0.

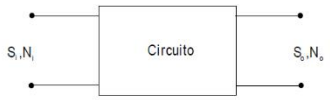
Once we have the two noise current generators we can compare them to our source. The source will have its resistance R with its noise and the input generator; I can forget about the rest of the circuit and consider just the following.
We want to compute the Thevenin equivalent of this circuit; in this way we achieve the equivalent input source, the one in red, where the total noise is not just the noise of the source itself, but we have also the contribution due to the current and voltage noise equivalent generators.



The new equivalent noise generator will be the following.

$$\frac{V_{in\,tot}^2}{\Delta f} = 4kTR + \frac{V_n^2}{\Delta f} + \frac{i_n^2}{\Delta f} \cdot R$$

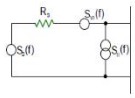
SIGNAL-TO-NOISE RATIO AND NOISE FACTOR



Signal-to-Noise Ratio: $\frac{S}{N} = SNR = 10\log_{10}\left(\frac{V_{in}^2}{v_{n,i}^2}\right) = 10\log_{10}\left(\frac{V_{in}^2}{\left(4kTR_S + e_V^2 + i_I^2 R_S^2\right)\cdot \Delta f}\right)$

Noise Factor: $F = \frac{SNR_{in}}{SNR_{out}} = \frac{S_i}{N_i}\cdot\frac{N_0}{S_0} = \frac{N_0}{G\cdot N_i} = \frac{\text{Total output noise}}{\text{Total output noise just due to } R_S} =$

$$= \frac{\text{Total input noise}}{4kTR_S}$$

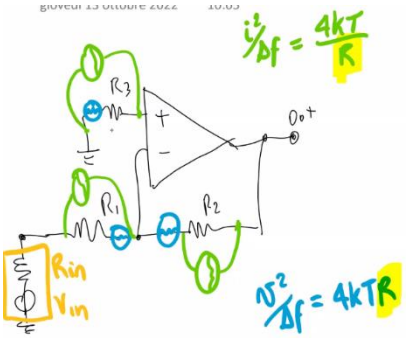


Noise Equivalent Temperature: $T_{eq} = (F-1)\cdot T_0$ just to tell how "hot" the noise is

The noise factor is given by the ratio between the total noise of the real circuit divided by the noise of just the source. Hence the noise figure is a mathematical number that compares the total noise of the circuit with the noise of the source. NF tells how bad the amplifier was in increasing the noise of the source. The extra noise added by the amplifier is a NF quantity higher than the source noise.

Recap

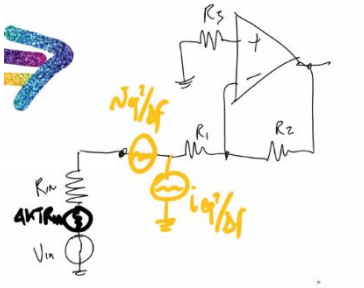
Given a real circuit, e.g. the one in the image, all the components are noise and they can be described with their noise generator (either voltage or current noise). The PSD i^2/delta(f) = 4kT/R for the current generator.



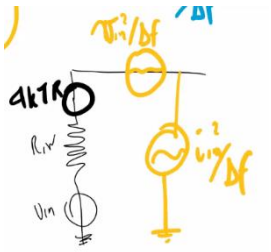
We can move from this circuit to another one, not considering all the components. As soon as I discover the output to be noisy, I want to change something to reduce the extra noise. But if I study all the individual components it is different to understand which components gives the noise → let's move to the noise equivalent.

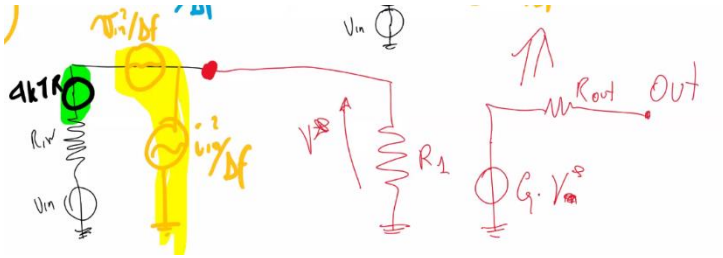
I want to consider the circuit to be noiseless because I computed the equivalent noise for all the noise source, and I get a voltage and current noise equivalent spectral densities. Then I attach the input (e.g. microphone) and I have to consider also the intrinsic noise of the source.

This second circuit is identical to the previous one.

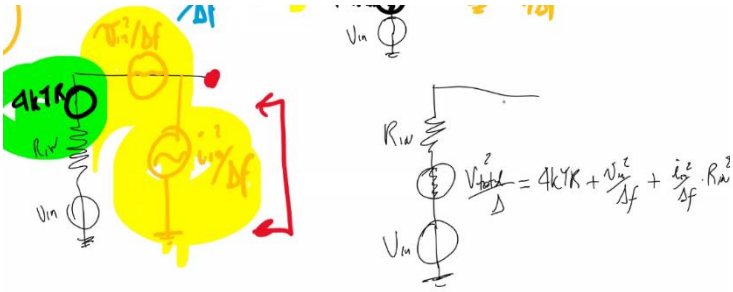


If they are equivalent, we can stop at the input stage and avoid the opamp, considering only the source and the input stage. In studying this circuit I'm not committing any error, because the output noise is not needed to be computed, because if I want to compute the importance of the intrinsic noise of the source compared to the noise added by the amplifier it is enough to stay at the input stage → we can forget what happens in the rest of the network (red part, R1 is the input impedance).





If we stop at the input we can compute the total input noise, considering the Thevenin equivalent of the circuit (not considering the rest of the network).



The one on the left is the total noise at the input, and it is mathematical, we cannot measure it, it is the one used to have the equivalent circuit to behave like the real one with all the noisy components inside. In fact, the equivalent input noise generators are a mathematical modelling, we have indeed the same output noise, but the noise I find at the input of the model circuit is higher than the one in the real circuit. So the noise we compute thanks to the equivalent is not real, it is just a modelling, the upper result is a model, we will never measure it in input of the stage.

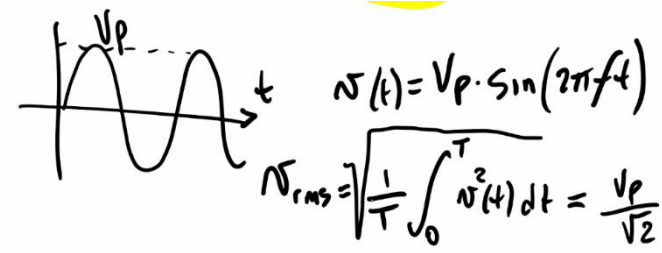
This modelling is useful because it let us define some figures of merits, to understand if the amplifier is good for the application I want. I can define for instance the S/N.

SIGNAL TO NOISE RATIO

We cannot compute directly the ratio between the signal in V and the noise that is a spectral density, so we have to consider power both at the numerator and denominator.

What is the power of a sinusoidal signal whose peak is Vp?

If we compute the integral between 0 and T of the square, we get the rms value of the signal and to get the PSD we need to take the square root. So the power of the sinusoidal is equal to the peak value divided by sqrt of 2. It is an equivalent DC voltage with this value, it dissipates the same power of the sinusoidal with a Vp value.



S/N = (Vp/sqrt(2))^2 / [(Vtotal^2)/delta(f)] * delta(f)

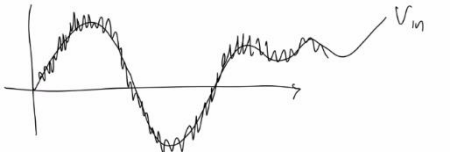
Once we find the input sigma^2 we can compute the ratio.
So every time we have to compute the SNR we have to compute the 10*log of the ratio between powers (to express it in dB). At the denominator I have the sigma rms of the noise to the power of 2.

If instead of considering powers as above, we prefer to consider peaks or rms values of the input, we can do it, we can compare the input rms value with the noise rms value (green), that is sigma. In this case we have to consider 20*log, and I will get the same number.

27dB, for instance, means that at the input the signal is the black one and the noise is 3 mV rms, a factor 27dB lower than the signal.

$$SNR = 10 \cdot \log_{10} \frac{\left(\frac{V_{in\,peak}}{\sqrt{2}}\right)^2}{\sigma_{rms}^2} = 10 \cdot \log \frac{\left(100mV/\sqrt{2}\right)^2}{(3mV_{rms})^2} = 27dB$$

$$= 20 \log_{10} \frac{V_{in\,peak}/\sqrt{2}}{\sigma_{rms}} = 27\,dB$$



However, the SNR considers the full noise, so it doesn't tell me how much the signal is bad, since it considers the source noise and the amplifier together. So it is better to define the noise factor, which compares the intrinsic noise of the source with the one of the amplifier.

NOISE FACTOR

It can be computed at the output or at the input, it is the same. It is the total noise at the numerator divided by just the noise of Rs.

$$NF = \frac{\text{total noise}}{\text{noise of just } R_S} = \frac{4kTR_S + \frac{V_{in}^2}{\Delta f} + \frac{i_n^2}{\Delta f} \cdot R_S^2}{4kTR_S}$$

$$NF = 10 \log_{10} \left(1 + \frac{V_n^2/\Delta f}{4kTR_S} + \frac{i_n^2/\Delta f}{\frac{4kT}{R_S}}\right)$$

Then we have the Noise Figure, that is the dB version of the noise factor.
If we buy an amplifier with no noise, NFideal = 1 = 0 dB. It means that the amplifier is not introducing any noise.

Instead, NF = 4 = 6 dB means that 10^(6 db/10) = 10^0.6 = 4 it means that the number in the parenthesis is 4, so the amplifier introduces a noise equal to a factor 3x compared to the noise of the source (microphone).
So if e.g. Rin = 1k, the intrinsic noise is 4kTR = (4 nV/sqrt(Hz))^2, but the total noise will be the intrinsic noise multiplied by 3 due to the amplifier plus the intrinsic noise of the resistor, that is in the end the intrinsic noise multiplied by 4.
So the noise figure defines how good is the input with respect to the amplifier.

The noise figure relies on Rs, that is a real number, T that is the temperature and on noise spectral density that are mathematical values obtained studying the ideal circuit. So the NF can never be measured, it can

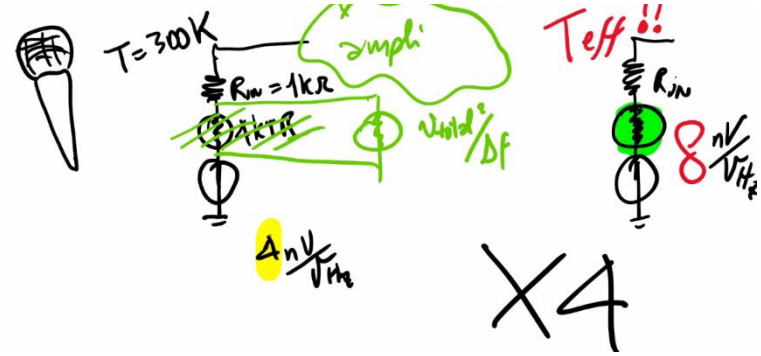
just be calculated. In reality, despite being the passage just mathematical, the noise figure can be measured.

We have a real circuit, and all the resistances are noise. Let's measure the SNRin, that is equal to the signal divided by the noise, and the SNRout.
If everything were ideal, the output noise should be equal to the input noise, but it is higher because the amplifier adds a contribution.
We can compute SNRin/SNRout = Sin*Nout/Nin*Sout = Nout/(Nin*G), where N is the noise.
Nout is the total output noise, and Nin*G is the output noise just due to Rin, so we are getting NF.

So NF has a mathematical equation but we can also measure it measuring SNRin at the input, SNRout at the output and then performing the ratio.

NOISE EQUIVALENT TEMPERATURE

Moreover, the input microphone (source) has its Rin and related noise, but instead of the noise 4kTR related to Rin, let's consider the total noise adding also the amplifier contribution. Now the resistor is noisier wrt the simple resistor.
This means that my resistor has a higher noise, so the resistor is not like at T = 300 K as it was alone, but it is at a higher Teffective (Teff). If previously it was at 4 nV/sqrt(Hz), since we are a factor 4 (in terms of power) of improvement and since the rms is the sqrt of power, we get 8.

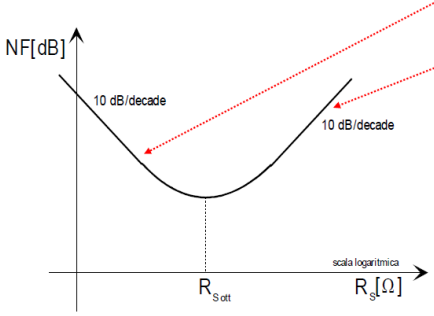


So the amplifier at 300K is so noise when connected to the amplifier as if it was alone at 900K.

NOISE FIGURE - Computations

Noise Figure:

$$NF = 10\log_{10}\left(\frac{4kTR_S\Delta f + <v_i^2> + <i_i^2>R_S^2}{4kTR_S\Delta f}\right) = 10\log_{10}\left(1 + \frac{<v_i^2>}{4kT\Delta f}\cdot\frac{1}{R_S} + \frac{<i_i^2>}{4kT\Delta f}\cdot R_S\right)$$



Minimum at:

$$R_{S,opt} = \sqrt{\frac{<v_i^2>}{<i_i^2>}}$$

$$NF_{opt} = 10\log_{10}\left(1 + \frac{<v_i^2>}{4kTR_{s,opt}\Delta f} + \frac{<i_i^2>R_{s,opt}}{4kT\Delta f}\right) = 10\log_{10}\left(1 + 2\cdot\frac{\sqrt{<v_i^2>\cdot<i_i^2>}}{4kT\Delta f}\right)$$

x x

Let's draw the noise figure, that is the one enlightened in yellow. The first term decreases with Rs, while the second one increases with Rs. When the first term prevails and it's much higher than the other, 1 is negligible and also the last one, and this happens for low value of Rs.

For low values of Rs, we have a high number in the parenthesis related to, so the equivalent voltage noise of the amplifier gives the major contribution, whereas for high values of Rs the other terms becomes dominant. At low Rs, NF decreases as 45°, at high Rs it increases with 45° (factor 10 dB).

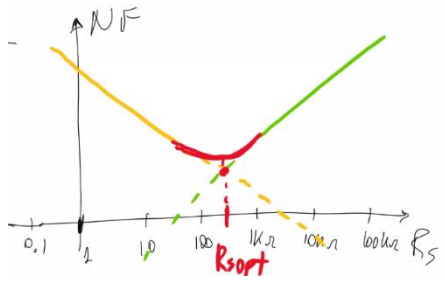
The optimal Rs can be found by computing the derivative. If we do the ratio of the equivalent voltage and noise generator and we take the square root we get the optimal Rs. Then we can also find NF optimal, in which contribution x and x will be the same.

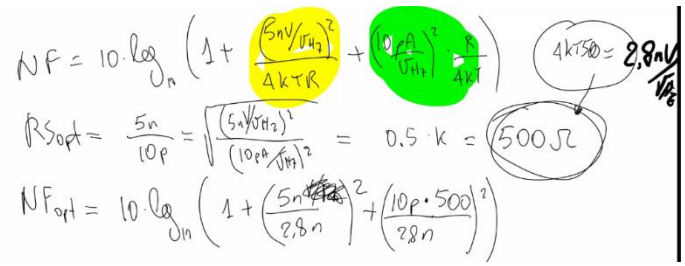
In the end, the amplifier itself with all its noise is equivalent with a voltage equivalent and a current equivalent and I attach them to the input where I attach also the microphone (source). Let's suppose current equivalent is 10pA/sqrt(Hz) and voltage is 5 nV/sqrt(Hz), and if Rs = 1k, noise of it is 4nV/sqrt(Hz).

Let's plot the noise figure.

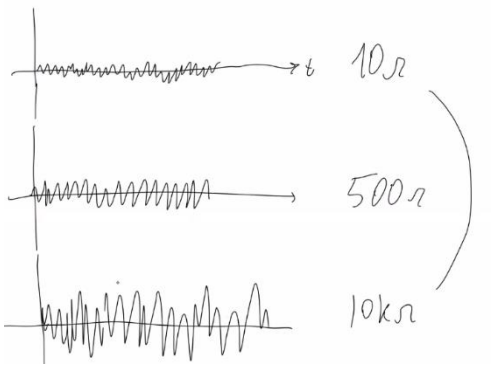
$$NF = 10 \cdot \log_n \left(1 + \frac{(5nV/\sqrt{Hz})^2}{4kTR} + \left(\frac{10pA}{\sqrt{Hz}}\right)^2 \cdot \frac{R}{4kT}\right)$$

At low values of Rs the second contribution is the most important one, whereas for high values it is the last one.

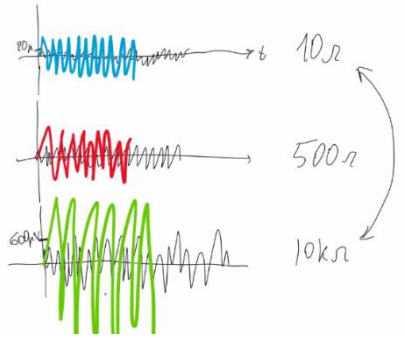




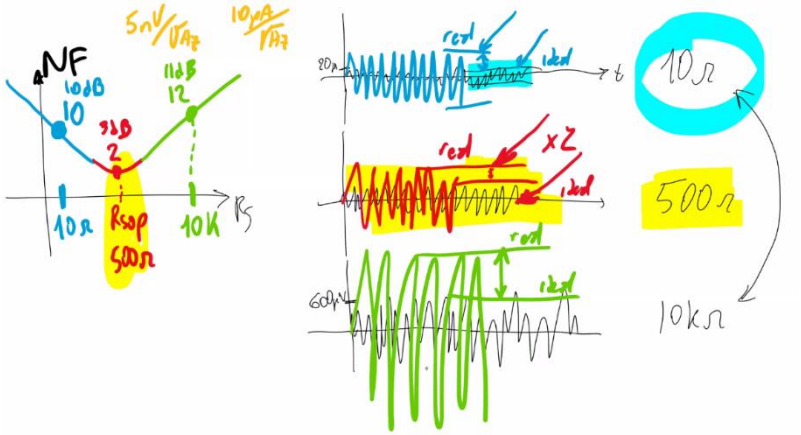
We discovered that the NF has the trend in the left image.
If for instance we use a resistor smaller or greater than the Rs = 500 Ohm optimal. Nevertheless, the amplifier gives 5nV/sqrt(Hz) and 10pA/sqrt(Hz). Without the amplifier



If then we connect the amplifier, the real noises are:



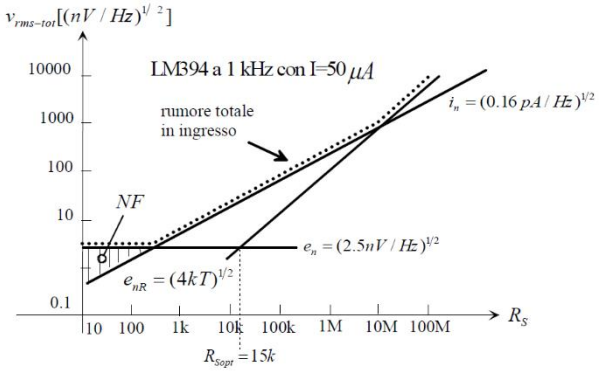
The ratio between the real value and the ideal value is in the Ropt case. If we choose a resistor too low, the problem is that we have a high noise, we are not decreasing it. The same if the resistor is too high. At Rs,opt we can pretend the noise to increase by a factor 2, whereas in the case of too low R, the ratio between ideal and real is x10, and x12 with a too high resistor.



TOTAL INPUT NOISE

Total input noise (no R_{in}):

X $S_{in}(f) = 4kTR_S + e_V^2 + i_I^2 \cdot R_S^2$



NF shows how electronics worsens the intrinsic noise of the source
NF is NOT a signature of the total noise !

The minimization of the noise factor does correspond to the minimization of the noise? No, because the total input noise is equation x. If we plot it, the total input noise is the noise of the resistor plus the noise of the amplifier plus the current equivalent noise of the amplifier multiplied by Rs^2. e_v is not depending on r, so it is constant, then the resistor contribution that increases with Rs and the current contribution that increases as Rs^2.

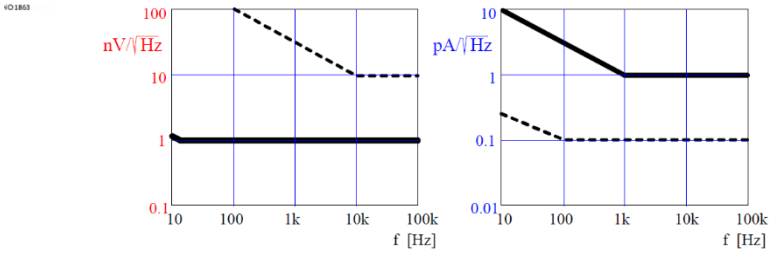
When the two contributions touch, we get the Rs,optimal, and where the total noise gets as close as possible to the ideal noise. The minimum noise is towards Rs = 0 because we have only the equivalent

voltage generator, and the current generator gets killed by the shortcircuit, and it remains only the voltage noise generator. If instead we move toward very high values of Rs, then the contribution that dominates is the current noise generator, the other two are neglibigle.

The minimization of the noise figure is to choose the source R as similar as possible to Rs,optimal if we have to choose the amplifier. Conversely, if the input is already defined, choose the amplifier so that its noises (current and voltage) in ratio are in the order of Rs of the source. If so, the total output noise from the system has an extra contribution form the amplifier is minimum with respect to the one from the source.

Instead, to reduce the noise in output form the amplifier, choose and Rs that is very very low. The lower Rs gets, the lower the noise becomes, saturating to the equivalent input voltage noise of the amplifier.

Example 1

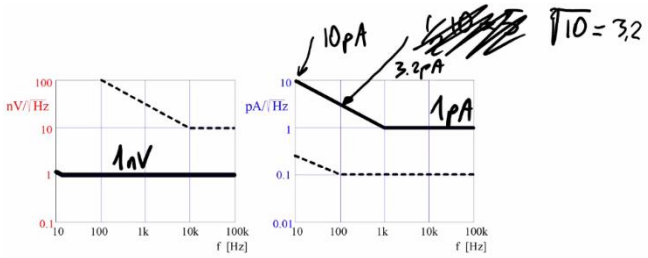


MOS (dashed lines) and BJT (solid lines) transistors at $I_C=I_D=1mA$

a) Draw the NF plots vs. $R_S=10\Omega \div 10M\Omega$, at 10, 100, 1k, 10k and 100kHz

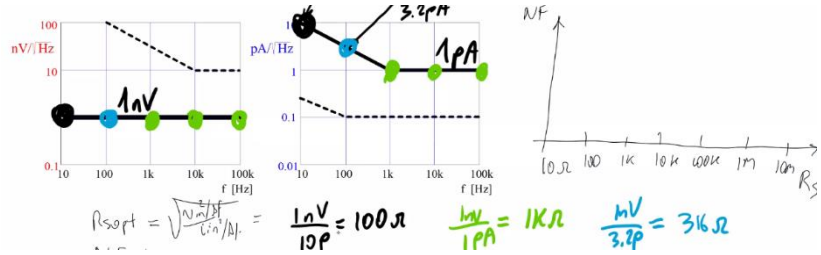
b) Select the lowest noise transistor for $R_S=10k\Omega$

We have two transistors; solid lines are voltage generator and current generator for BJT, dashed for MOSFET. The voltage noise generator for the BJT is lower than the one for the MSOFET and white. Instead, for a MOSFET transistor, since the gate current is low, the current noise generator is lower.
The current generator for the BJT is higher, I can see some flicker noise. As for the MOS, since the gate current is ideally 0, the shot noise should be lower but as for the voltage noise at the input it is higher. Moreover, MOS transistors have a higher Flicker noise in terms of voltage.



a) If we consider the BJT transistor, I can see that we have 1nV and 1pA, with 10pA of flicker noise in current. The value in the middle part is half of a decade, since we are in a log-log plot, so it is sqrt(10). Now let's draw the noise figure; we have to compute Rs,opt and NFopt. Rs,opt depends on Vin and i_in, as seen before. Let's consider the point at 10 Hz; it Is 1nV/10p, so 100Ohm.

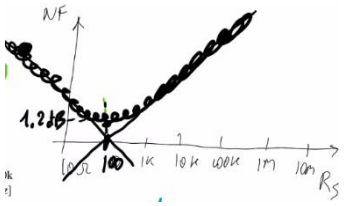
Then the noise figure in the green points for the current noise is always the same, and there Rs,opt = 1kOhm.



Then we can compute the NFopt. $NF_{opt} = 10\log\left(1+\frac{2}{\frac{a_{in}\cdot i_{in}}{4kT}}\right)$

NFopt = 1.2dB, and Rs,opt = 100Ohm.

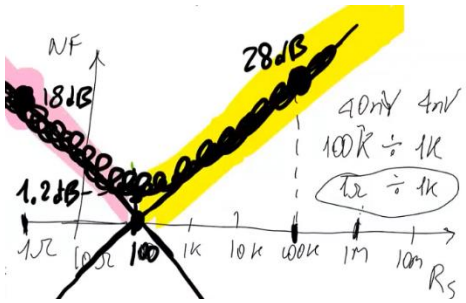
The last ratio in the parenthesis has a different value depending on the chosen point. For high values of Rs, we can find a point, the same for low value of Rs and then knowing the trend of the curves we can plot the value of NF.



So we compute Rs,opt and NFopt and then we choose values of Rs smaller and bigger than Rs,opt, we compute the NF value. E.g., in the case of Rs = 100kOhm it is the following.

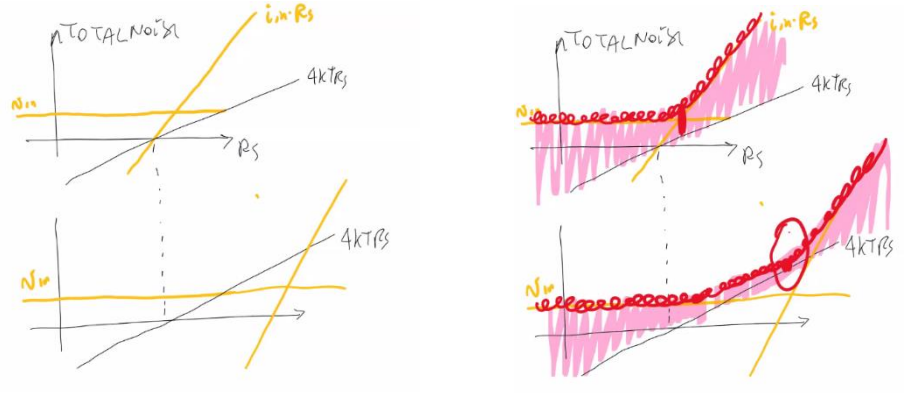
$$NF = 10\log\left(1+\left(\frac{1n}{40n}\right)^2+\left(\frac{10p\cdot 100k}{40n}\right)^2\right)\simeq 28dB$$

Then we repeat the same procedure but with a lower value of Rs, e.g. 1Ohm, and we put another point in the NF vs Rs plot. Then we go down by 45° and the two trends will cross in correspondence of Rs,opt.
The yellow portion depends on the current, while the pink portion depends on the voltage.

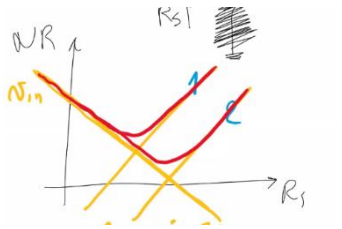


When we move alongside a plot of the noise current where it is the current that changes, the right shoulder of the NF plot decreases and it is shifted downward if the noise current is decreasing (e.g. in the case of the BJT from the black to the green points). Hence if we increase i_in noise, the Rs,opt decreases.

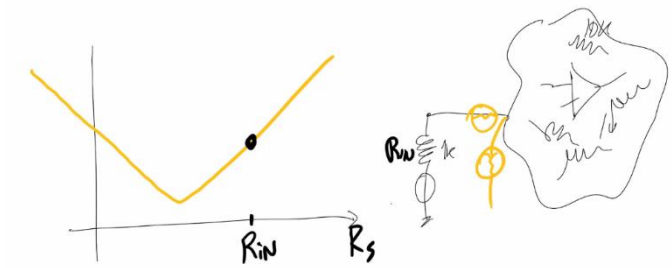
If we want to compute the total noise, we may be in the case where vin is constant, then we have the effect of i_in*Rs and then 4kTRs (upper image). Otherwise, we can have also the case in the bottom image, is i_in is very low. In the first case the total noise is almost given by vin and i_in, in the bottom case 4kTRs appears in a certain region, so the position where we have the minimum distance between noise and Rs is different. So the noise figure (pink) is different, and the minimum noise figure is lower.



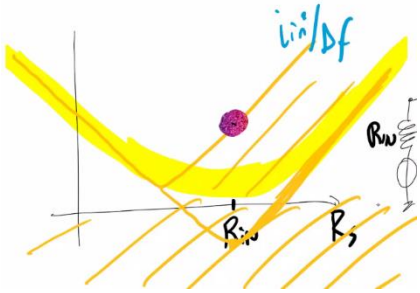
If we plot the NF in the two cases, Vin is the same, so the left shoulder is the same. What changes between the two is the right shoulder.
I can confirm that the NF in case 2 is lower than NF optimum in case 1, simply because i_in decreased.



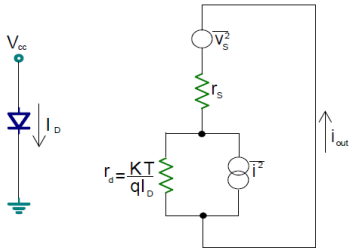
Better to change the amplifier or the source?
Now, I bought an amplifier, the microphone with 1k resistance (not the ideal 500 Ohm). Due to the Vin and Iin of the amplifier I discover that the curve is the yellow one, but Rin is not the optimal one. We can hence change the amplifier so that we center Vin and Iin to have Ropt in correspondence of Rin or we can change Rin.



Since we are on the right part of the plot, is Iin to be reduced so that Rin eventually is in correspondence to Rf. However, we will never be able to go to negative values of NF, the limit is 0 dB (we have 1 + … in the parenthesis).



NOISE IN DIODES



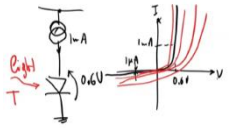
$$< v_S^2 >= 4kTr_S\Delta f \qquad < i^2 >= 2qI_D\Delta f + K\cdot\frac{I_D^a}{f}\Delta f$$

$$S_{out} = 2qI\cdot\frac{(1/g_m)^2}{(1/g_m+R)^2} + \frac{4kT}{R}\cdot\frac{R^2}{(1/g_m+R)^2}$$

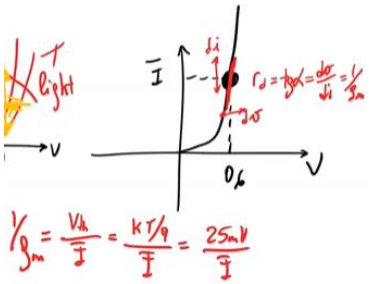
$r_D=1/g_m$ is NOT noisy… but it seems so: $< i_S^2 >= 2qI_D\Delta f = 2q\frac{kT}{kT}I_D\Delta f = 4kT\left(\frac{1/2}{g_m}\right)\Delta f$

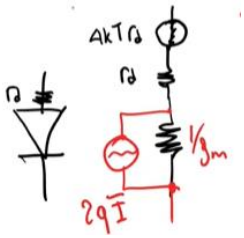
Equivalent Noise Resistors: $R_{eq.I} = 2\cdot\frac{1}{g_m}$ $\qquad R_{eq.V} = \frac{1}{2}\cdot\frac{1}{g_m}$

I may have a situation where I pump a current in a diode. The diode has a drop of 0.6V across it and I know the typical V-I curve of a diode. But if the temperature changes of light, the current voltage characteristic may drift.



When we operate a diode, we operate in a specific working point, and then we perform the small signal analysis around that point. And the resistance of the diode is the tangent of the curve in the w.p..

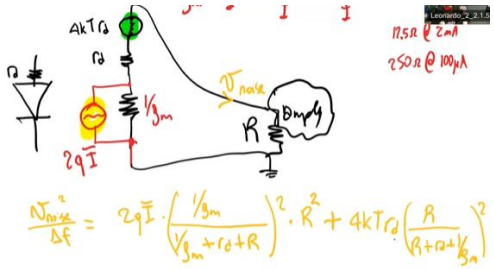




Hence a diode can be model with its 1/gm but also a parasitic real resistance r_d. r_d is a real resistor, while 1/gm is a small signal parameter.

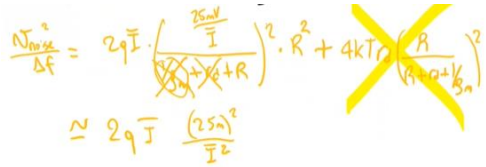
There is a current generator in the diode model due to the current that flows in the junction and it is a shot noise 2qI, where I is the polarization current. Then there is also the thermal current noise due to the r_d component (or we can also consider its voltage noise). More precisely, the I of the shot noise considers also the Flicker noise, but we can neglect it by now and consider only the two white noises.

Let's compute the noise of the diode that enters in the amplifier. The amplifier has its internal R and I want to know the current that enters in the amplifier, to then retrieve the noise. We apply the principle of superposition of effects.



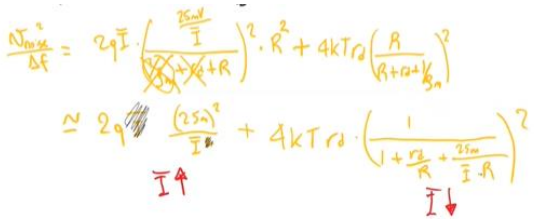
The total noise that enters in the amplifier is given by the yellow formula. The I and r_d to be used in the diode as to be chosen carefully. It is not true that if I and r_d both increase the noise increases. In fact,

the first contribution displays a 1/gm dependency over I, because 1/gm = Vth/I (Vth = thermal voltage). Of course it is also at the denominator in the parenthesis, but it is negligible with respect to r_d and R. So the first contribution decreases if the current increases.



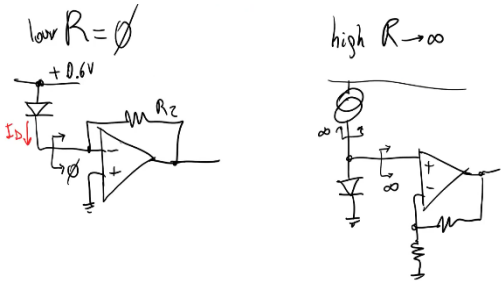
So if I increase the current I, the shot noise increases but the 1/gm drastically decreases. So the shot noise that reaches the amplifier decreases if we increase the current.

As for the other contribution, should operate with high or low I for this one? To decrease the denominator the current I should decrease.

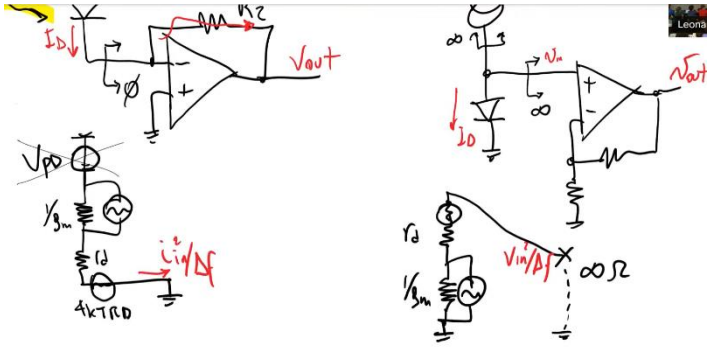


So the correct value of I to be chosen depends on r_d, R and see which contribution between the two prevails. Moreover, the first contribution is independent on R, while in the second contribution, if R increases it is good (R is the internal resistance of the amplifier).

Let's see two possible condition of R and the Id to be used in the diode.

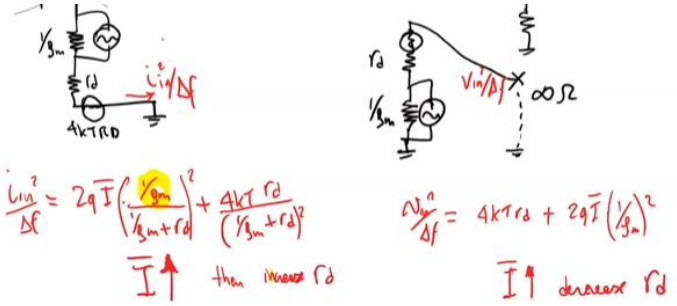


In the low R configuration, we have the diode, its 1/gm, the shot noise, the r_d and then virtual ground (we are not considering signals, just noises). Instead, in the other circuit we have the circuit on the right.



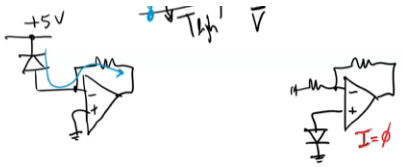
So in the first case it's better to compute current noise equivalent generator, in the second case the voltage one. If we calculate the current noise equivalent generator, we see that we should increase or decrease the 1/gm depending on how much r_d is big compared to 1/gm. In this left case it is better to increase the current I because 1/gm drastically decreases and shot noise is killed. Then we increase r_d so that the second contribution decreases.

These are the golden rules if we are measuring the diode with a transimpedance amplifier (left case). Instead, in the right case it is important to increase I and decrease r_d.



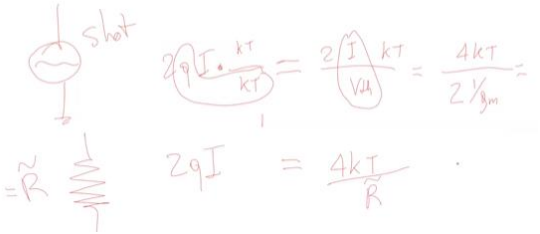
The left case is called photocurrent mode (we set the voltage to be constant), while the one on the right is called photovoltaic mode (we set the current to be constant).

We can also use another approach, operating the diode in the reverse bias regime (left) or in a circuit where the current I is equal to 0 (forward bias and no current at all).



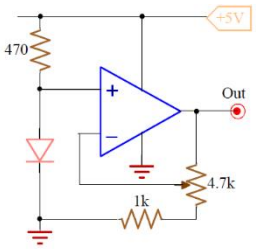
Shot noise remark

Shot noise is 2qI, but if I multiply and divide by kT, kT/q = Vth and Vth/I = 1/gm.



So it seems that the shot noise is equivalent to the noise of a real actual resistor whose R should be equal to 2*(1/gm). This is, however, from a mathematical standpoint → shot noise can be modelled as a thermal noise.

Example 2



OpAmp: A_0=100dB, GBWP=10MHz, ~~4nV/√Hz and 5pA/√Hz noise~~

a) Compute output rms noise, with trimmer's cursor at the two ends

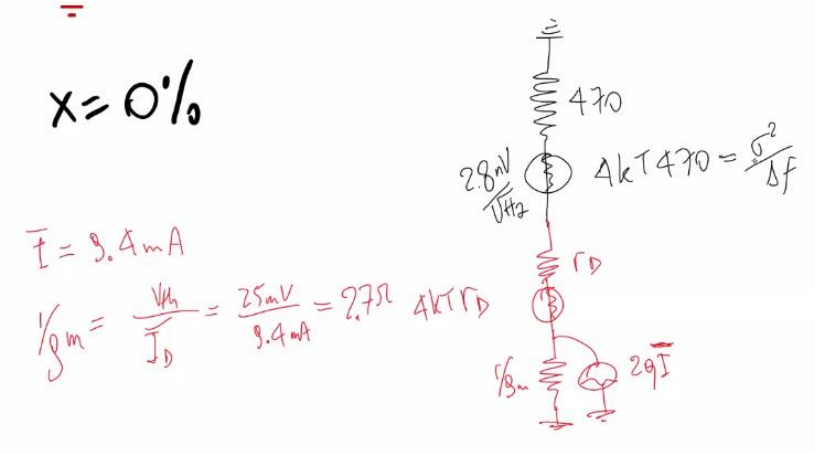
b) Discuss the role of the 1kΩ resistor

Resolution

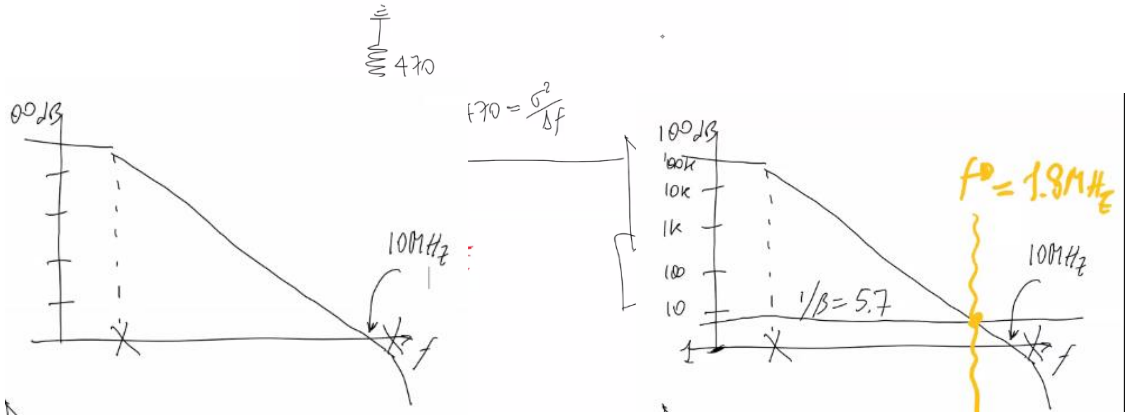
The potentiometer can be either to 0% or 100%.

0%
In this case we have that with 0.6V across the diode we will have 4.4V across the 470Ohm resistance and so a current of 9.4 mA in the diode.

Now let's introduce the diode, that can be modelled with its rd in series with its noise generator, then 1/gm and the corresponding shot noise in parallel to it. 1/gm = Vth/Id = 25mV/9.4mA = 2.7 Ohm



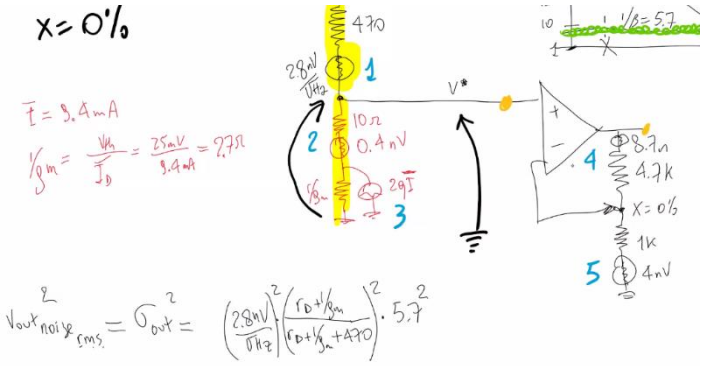
If rd = 10 Ohms, 4kTrd = 0.4 nV/sqrt(Hz).
Then we go straight into the opamp and then we have the cursor, so let's consider the case where it is at x = 0. We can write the noise for these resistors.



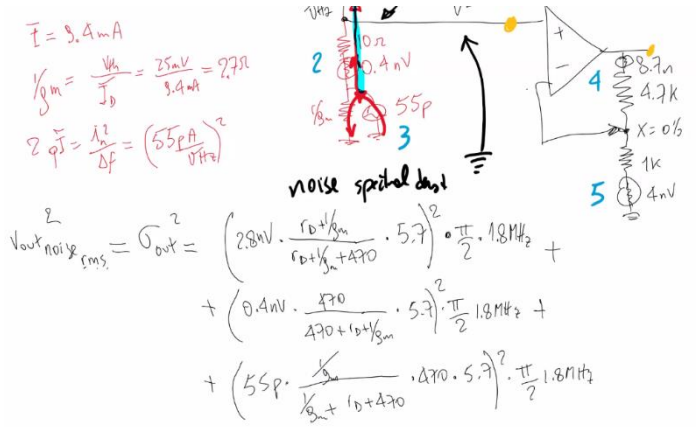
We know the opamp will be compensated, so the second pole will be after the GBWP. The beta of this configuration is 1/4.7, so the 1/beta is 4.7.

So the circuit has a bandwidth set by f*. So if in this configuration (non inverting), the gain will be 5.7 and then at f* I die as Gloop dies. From user standpoint the circuit has a pole at f* and one where the second pole of the opamp is.

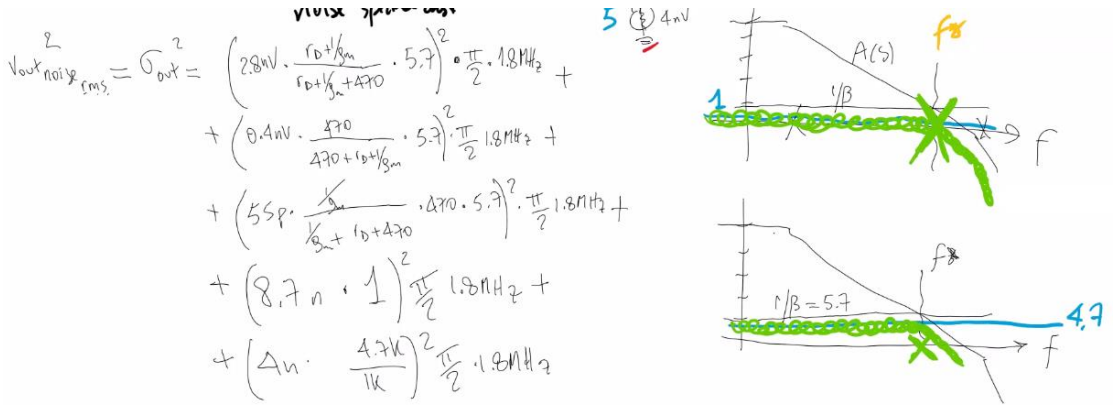
a) We have 5 contributions to the noise, let's consider them all. Let's consider the power two of the output rms, that is the sigma squared out. When we study 2.8nV contribution noise, it won't go directly to V+, because it is in series with the resistors (not other generators, they are not active), so we need a voltage partition. Then I multiply by 5.7^2 that is the gain of the stage.



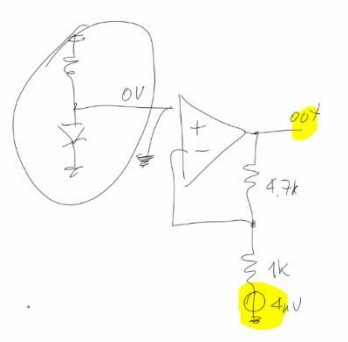
To be coherent, I know that we have all squared, and we have a noise spectral density, so we should multiply for the noise equivalent bandwidth. Which has to be computed. In the case of one single pole, the noise equivalent BW is the pole of the real circuit multiplied by pi/2. This is the first contribution of the first noise source. Then I have to consider the 0.4nV generator, then I multiply by the gain and the equivalent bandwidth.
Then we have the current partition for the last generator, 2qI.



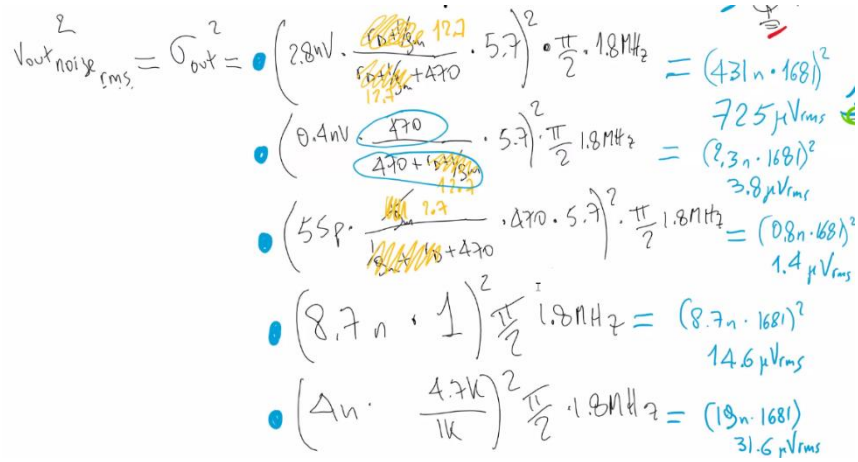
For the 4th contribution, the other noises are off, so V* (V+) is 0, so no current flows in the branch where we have 1k, so the 8.7nV appear straight at the output of my stage. The I have to multiply by the bandwidth. For this, we have to consider the behaviour of the circuit when the input is not the classical input, but the input is the 8.7nV generator. F* is the same as before, and for generator 4 the real gain is 1 up to f*, firstly with 20db. So for the user the pole is at f* again (1 is the ideal gain of this generator to the output.



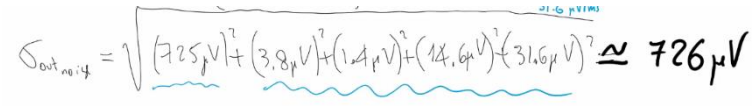
As for the 5th generator, it is an inverting gain stage. Then the equivalent bandwidth, even for this generator is at f* (ideal gain is 4.7, then we plot A(s) and 1/beta that is 5.7. Again, the pole is still f*).



To complete the calculations, let's put numbers paying attention to what is negligible with respect to something other.

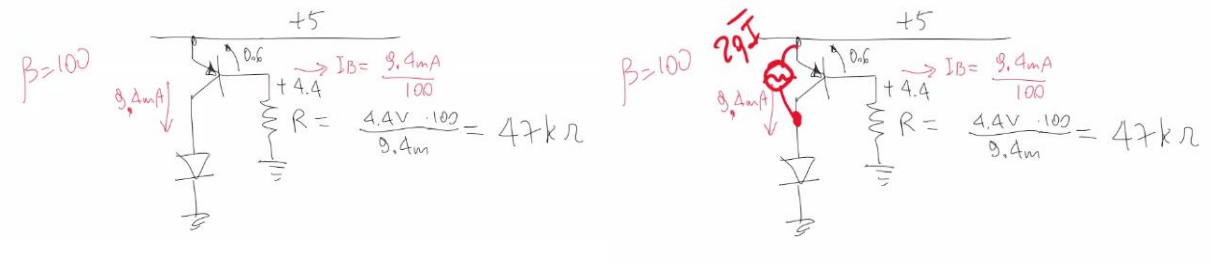


So the final remaining contribution is slightly higher than the dominant one (725).

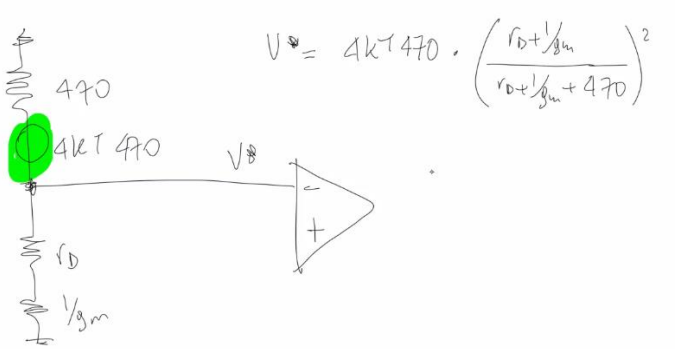


So when x = 0%, Gain = 5.7 and sigma_rms = 0.7.
So to reduce the output noise we don't change resistor, noise or current, but we have to change the 470 ohm resistance. If we change it, the current changes. To reduce the contribution of it, I have to reduce the value properly. A possible solution to reduce the value but having the same current is also to reduce the power supply.
Even better, we can change the circuit. Instead of using a resistor to bias the opamp, which has a high thermal noise, we could use a pnp transistor as a current source. If we have a transistor with a beta of 100, it means that Ib = 9.4m/100, so we need a base current of 4.4/100 * 9.4m.

However, also this current generator has to be modelled with its shot noise of 2qI. And also the resistor on the base has a noise to be considered.



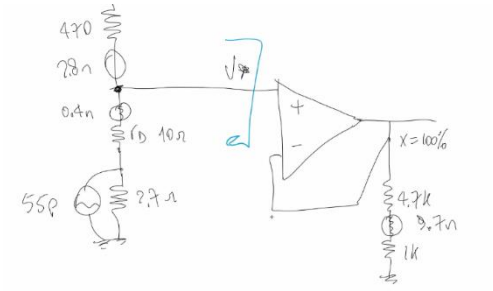
However, 470 ohm was also at the denominator of the voltage partition related to the noisiest generator. This means that probably the reduction of the resistance is not enough. Maybe it is better to increase the 470 resistance instead of decreasing it. Let's see what's better.



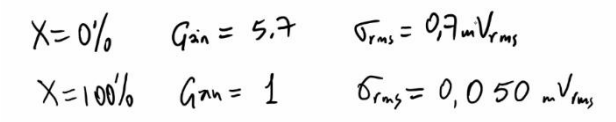
If we reduce too much 470, the noise generator reduces, but the other contributions take into account. If we instead increase 470, we can have it still dominating, but at the same time, thanks to the power 2, the increase in 470 reduces the noise.

100%

We have to consider a brand-new circuit. The two resistors at the output are in series so they act as a 4.7+1 = 5.7k resistor.

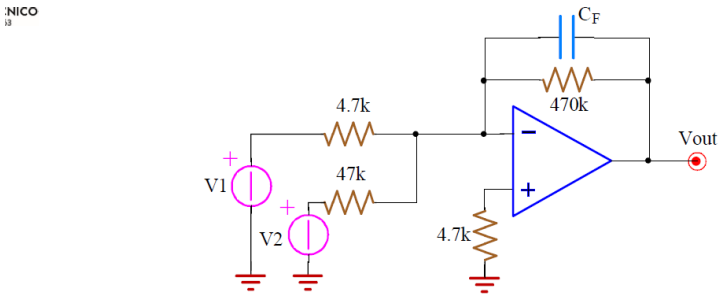


a) All the contributions to V* are the same as before. We have just to consider the contribution of the 5.7k resistance. It is zero because if we turn off all the left part, V* = 0 and Vout = 0, so the resistances are between grounds and their contribution is nihil. The current noise flows inside the opamp and it is the opamp that swallows it. Hence noise is much smaller than before.



b) The role of the 1k resistor is the following. Let's consider a configuration as the following. We run the risk that when the potentiometer is at x = 0%, we are at ground, so the gain is infinite, so we add a resistor at the bottom of the trimmer so that the gain remains finite (REG, change image).

Example 4



OpAmp with $v^2_{in}/\Delta f=(8nV/\sqrt{Hz})^2$ and $i^2_{in}/\Delta f=(0.5pA/\sqrt{Hz})^2$

a) Compute the noise equivalents for input V1

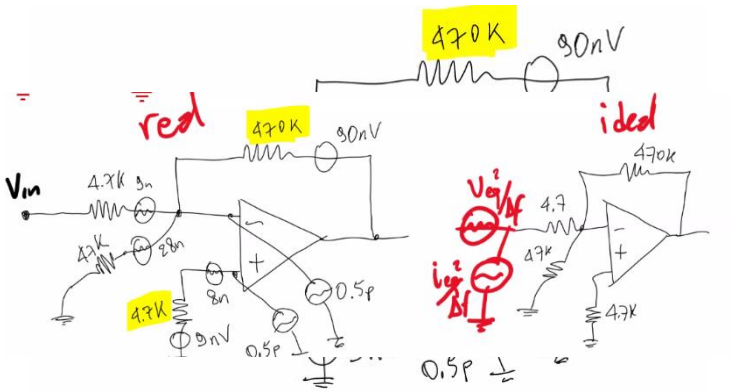
b) Compute the output rms noise

Now also the opamp is noisy.

Resolution

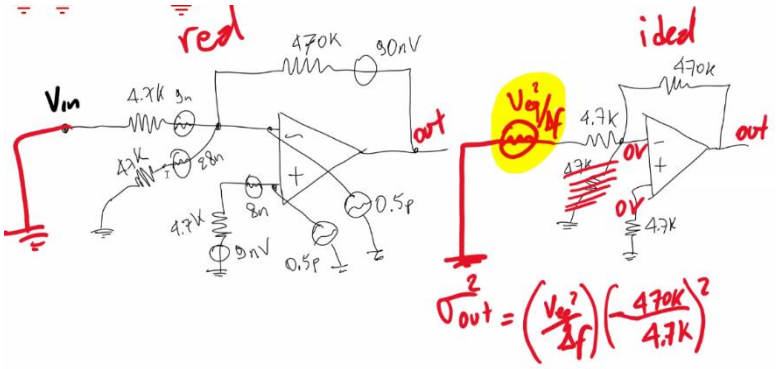
Let's redraw the circuit adding the noise sources.

a) The opamp has current sources.

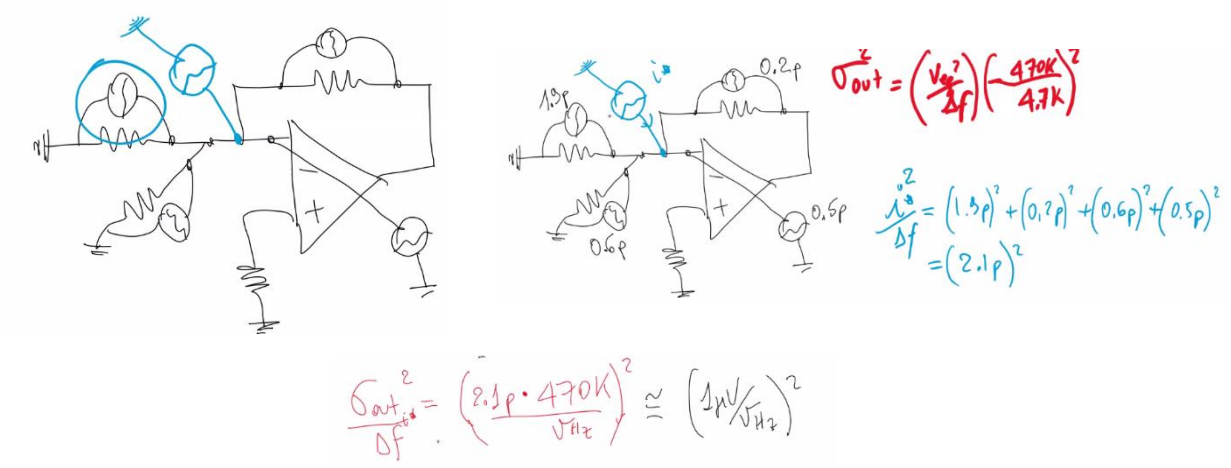


I want to study the real circuit to compute the voltage equivalent and current equivalent generators to be placed at the input and get an ideal net.

To compute Veq, let's connect the input of the two networks (ideal and real) to ground. Thus I can forget the current equivalent generator in the ideal circuit. Then in both circuits I compute the total noise at the output and then I compare the two values.

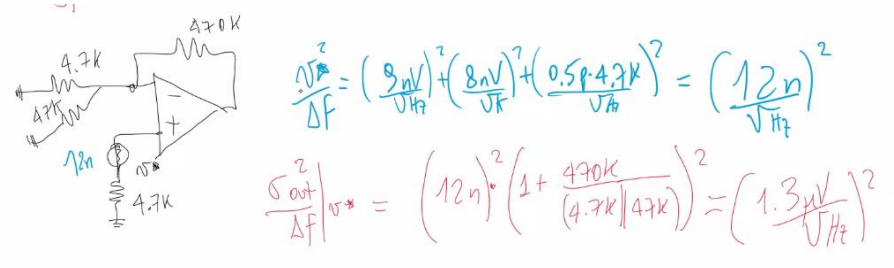


Now let's study all the noise sources for the real circuit. Sometimes, it is better to consider the current noise generator instead of the voltage noise generator. Because of how the network is build, I can consider a unique current generator given by the sum of all the power current generators.



Then I compare this last result with the previous result from the ideal circuit, once also the other input is computed.

Now let's consider the noise sources related to the input +, because so far I've analyzed only the – terminal input. Then I compute the output noise due to V*.

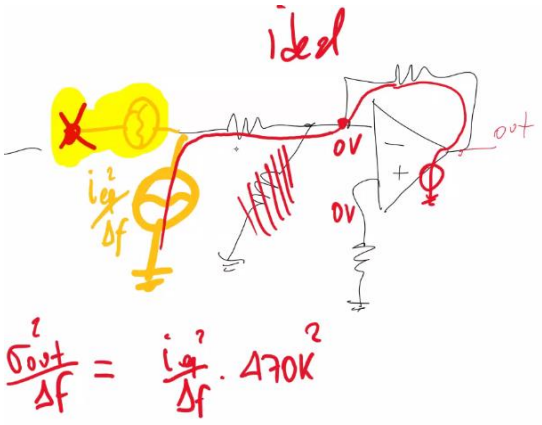


In total, in the real circuit: $\frac{\sigma^2_{out}}{\Delta f} = \left(1\mu V/\sqrt{Hz}\right)^2 + \left(1.3\mu V/\sqrt{Hz}\right)^2 = \left(1.6\mu V/\sqrt{Hz}\right)^2$

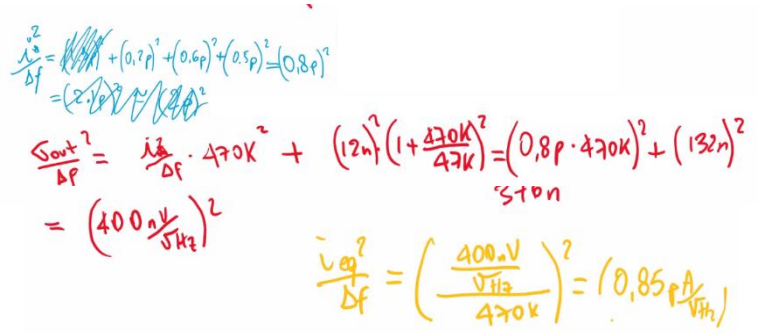
This is the total output spectral density due to the real circuit, and we also have the one in the ideal noiseless model and now we need to equate them.

We have now to compute the current equivalent generator. In the real circuit the reasoning is the same, I put all the noise generators as current and then all in parallel. But in this case the input is floating, so the 4.7k resistor is not playing a role because floating.

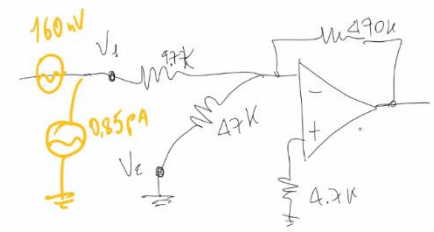
Firstly, let's reason on the ideal network.



Now let's compute the real noise. The part related to the + terminal remains the same in terms of noise voltages. Instead, the macro noise current generator cannot be user, because the input is floating and so the related current noise generator must not be considered. Also the gain of the stage changes.

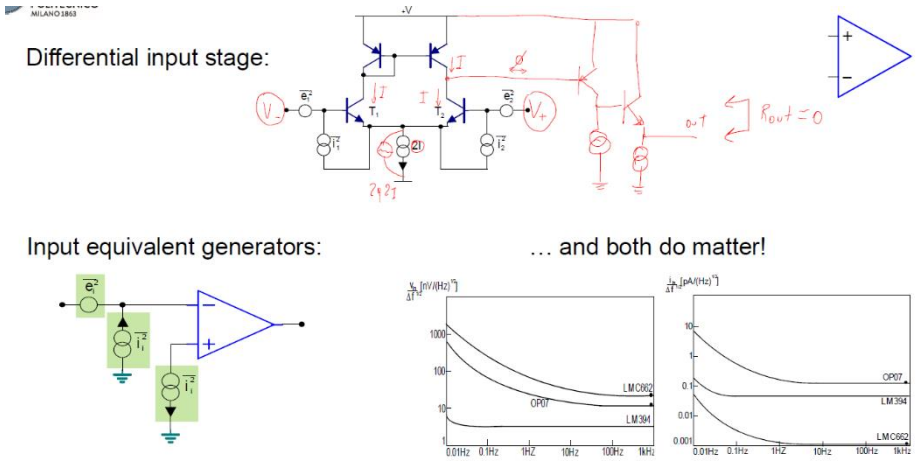


Now we can compare the noise of what we connect (the source) with the input equivalent noise generators and then we can compute the NF.



NB: if one of the two inputs, either V1 or V2 is detached, the corresponding resistance will be floating, and the overall noise will be lower.

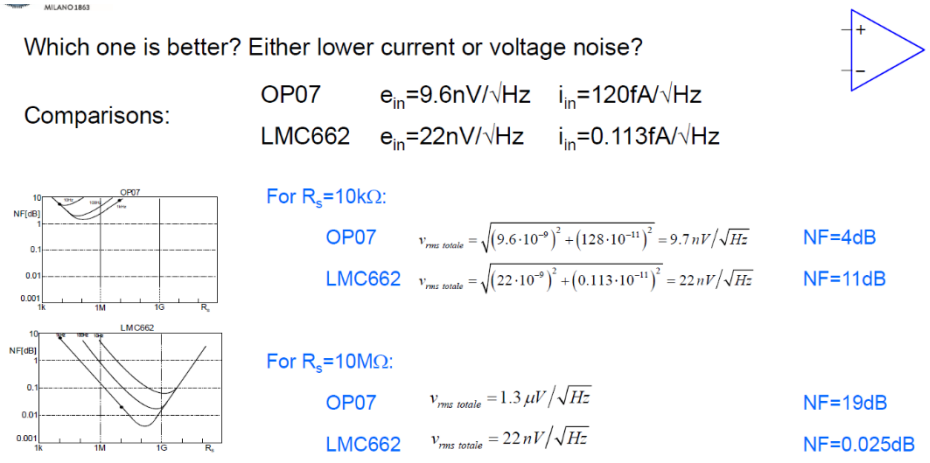
NOISE IN OPAMPS



In an opamp we have BJT or MOS transistors, and each component has its noise. This is the reason why we compute the current and noise voltage generators for the opamp. Since we have two inputs, we have to voltage generators and two current generators. Than the two voltage generators are in series, so it reduces to one in the final representation.

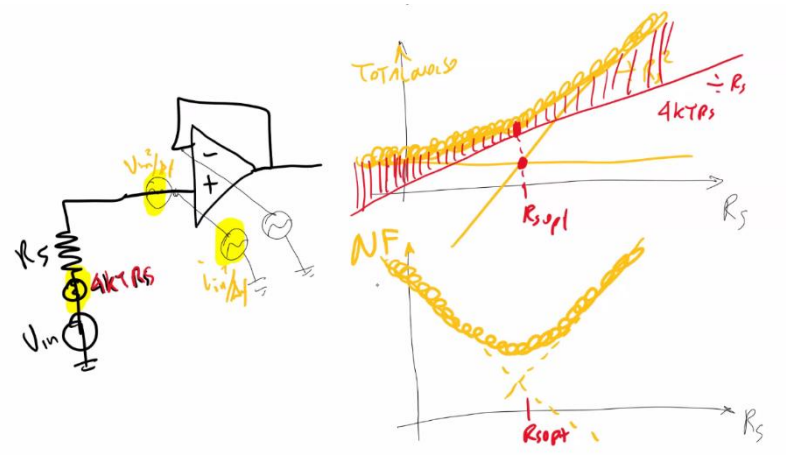
The noise of each passive component attached to the opamp is then added to the equivalent noise current or voltage generator at the input of the network. However, in doing so, we are theoretically over-estimating the noise because we consider the same passive components twice for each generator (e.g. for the feedback resistor).

On the market, we find different opamps, quoted in terms of equivalent voltage or current noise generators. In general, BJT opamps have a higher current noise with respect to MOS ones, but then the trend is opposite for the voltage noise.



To choose the better, we should place the opamp in the real circuit and then compute all the noise figures. So for low value of Rs the first one is better, whereas for high values the other one is better. So depending on the circuit, the voltage or current noise of the opamp matters more or not.

Let's consider a simple buffer to which we connect our microphone with Vin and Rs. If we study the total noise of the circuit, we realize which Rs is better, and it depends on the noise figure (the red distance). Where we have the minimum distance between the total noise and the Rs noise gives the Rs_optimal.

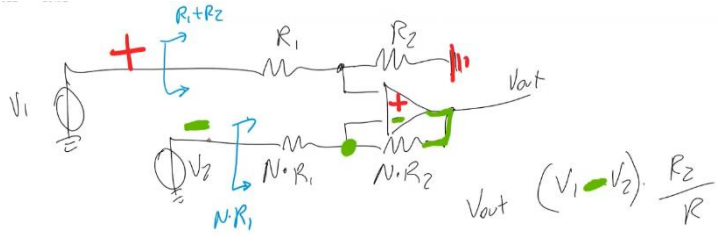


If i_in increases, its shoulder (right one) will shift upward. The same applies for the left shoulder of NF if Vin increases.

INSTRUMENTATION AMPLIFIER – INA

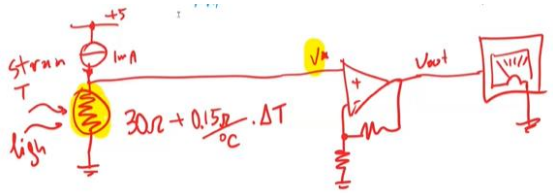
DIFFERENTIAL AMPLIFIER

We have two sources and we want to compute the difference.
The signal that gets positive amplification is that with R2 to ground and in the + terminal.

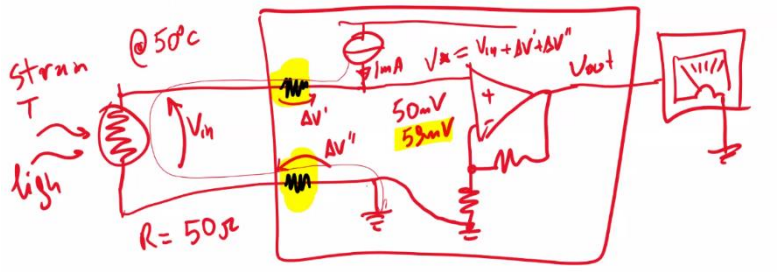


The problem of that in this circuit the input impedance is not infinite.

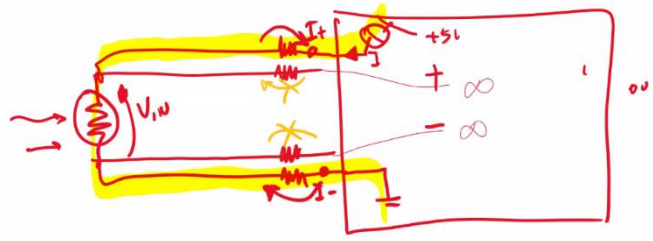
Let's consider a LDR. A possible way to amplify a voltage signal is to use an amplifier.



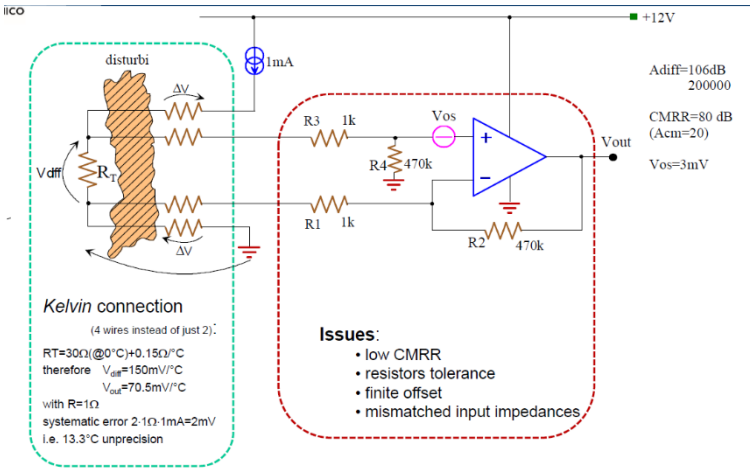
But now the issue is different because in the real circuit the current generator is not where we have the sensor, and the same for the sensor. But in this case there may be a series resistor in series with a wire and so the current generator is giving current not just to the sensor, but also there is a drop on the cable resistances. So instead of seeing 50mV I see 59mV, but this gives an error in reading the temperature.



So this two wire configuration is not good. Better to use the 4 wires configuration. One wire is used to pump current in, one is used as ground (I sink current out from this wire) and the other two are used to measure voltage difference. Now we have a voltage drop across the two power wires, but no voltage drop on the wires used to read the voltage of the sensor. This is true if the voltage difference amplifier, with infinite input impedance.



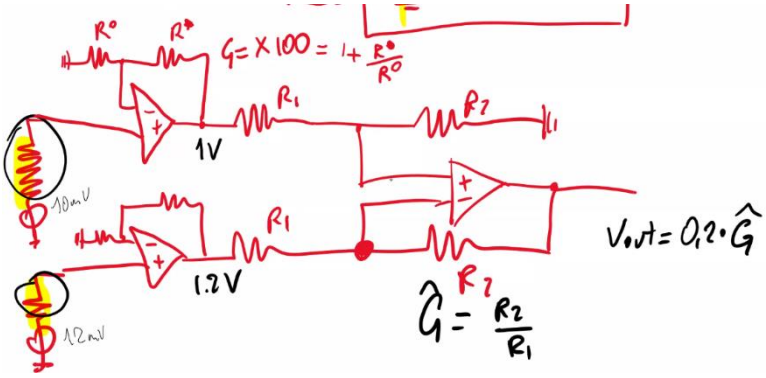
So the classical differential amplifier is not good.



If we have inputs with different source impedances, the problem is that even though the input voltages are the same, also because the input impedance of the amplifier are finite, the output is not 0.

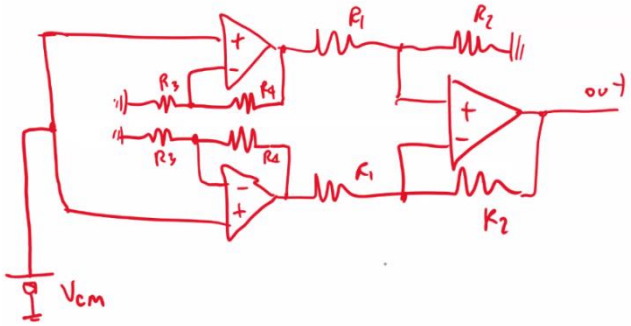
So we could in principle use a buffer between the source and the input terminals of the differential amplifier. However, we have another problem, because the buffer are noisy, so the same noise is also at the output at the buffer, whose gain is 1, and then will be amplified by the next stage.

So let's put an amplification instead of a simple buffer.



This is good because if R1 are noisy, thanks to the amplification, the input source noise gets so amplified by the first stage that the noise of R1 becomes negligible.
But then, there is another issue. If the signal has a common mode CM, typically we want to amplify the differential mode and not the common one, but with the upper circuit we amplify both, so if the input signal increases too much, we saturate the first stage.

Let's apply a common mode signal to the next circuit. The output of the opamps of the first stage will saturate, so the output signal will be random and any differential signal won't be seen because of the saturation.

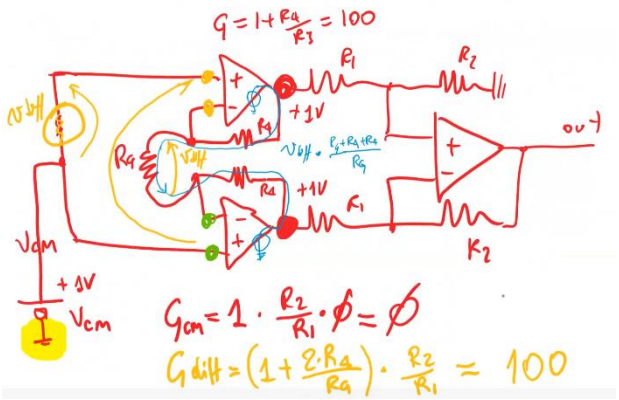


Hence in this case any differential input won't be seen to the output due to the fact that the Vcm saturates the opamp. To avoid the amplification of Vcm we connect the two R3 resistors.

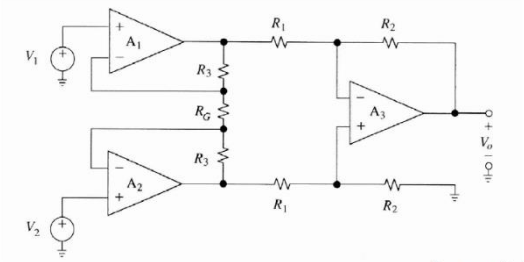
Let's connect the two resistors R3 to Vcm. Now the common mode gain of the first stage is equal to 1, so we don't have saturation anymore. Let's apply now the differential signal. If it is like in the image, also for this case the gain is 1, so it is not good.

Hence the solution is to connect the R3 resistances to the – terminals of the opposite opamp because I don't want to touch the input to preserve its quality (differential signal). This is the ultimate connection. By doing so, the R3 resistors are in parallel, and I can simply use one resistor Rg.
Thanks to this connection, the CM signal has a gain 1 in the first stage, and the second stage has a gain equal to R2/R1. But since the two nodes in output of the first stage are equal, now the common mode signal in output is 0.

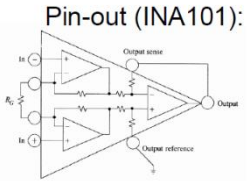
As for the differential signal, thanks to VG we are copying the values on the terminals of the opamps of the first stage, so we have a current in Rg that passes through R4 and goes to the second stage.



So we have a perfect rejection of the common mode.



Pin-out (INA101):



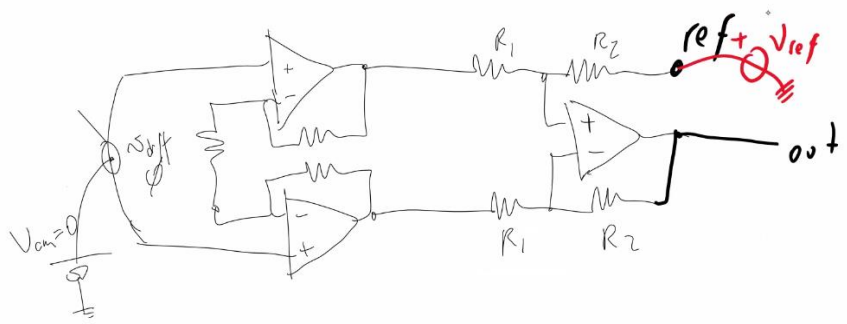
Differential gain: $A = A_I A_{II} = \left(1 + \frac{2 \cdot R_3}{R_G}\right) \cdot \left(\frac{R_2}{R_1}\right)$

Common-mode gain: $A_{cm} \approx 0$

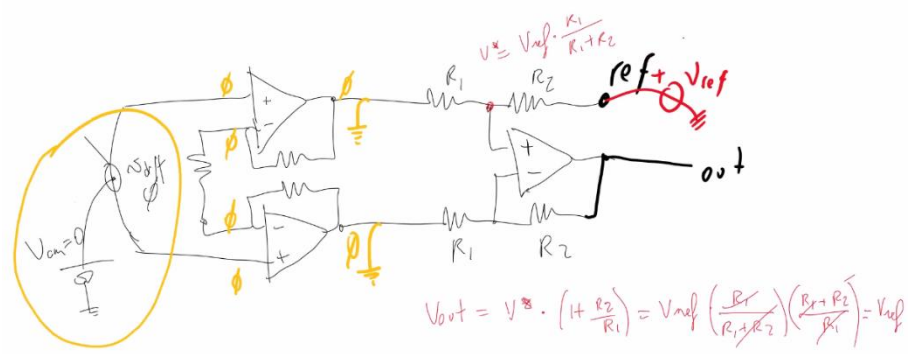
finite, accurate and reliable Gain	between 1 and 1'000
twin extremely high input impedances	> 10MΩ
extremely low output impedance	< 100Ω
extremely high CMRR	> 90dB

REFERENCE PIN

Let's see how to connect to an external load. Let's imagine the Vin at the input is zero and also Vcm, because we want to study the gain between Vref and Vout.

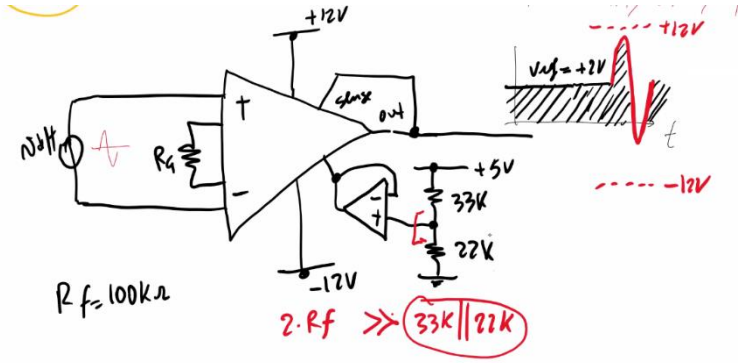


Since the input stage is off, we have everywhere 0, so the inputs to R1 are at ground. We can hence do the calculations. We end up having a gain 1 from Vref to Vout.



Whatever we connect to the reference pin, the effect is to have an offset in the output equal to the Vref. We can use the Vref to apply a constant DC value to the output of the INA. Then if Vdiff moves positive and negative, the amplified Vdiff in output will be superimposed to the Vref offset.

A minor issue is that the impedance that we see from the reference pin is R1 + R2, so usually R1 = R2 = Rf by the manufacturer, so the impedance we see is 2Rf. If Rf = 100k, the impedance is 200k, so we

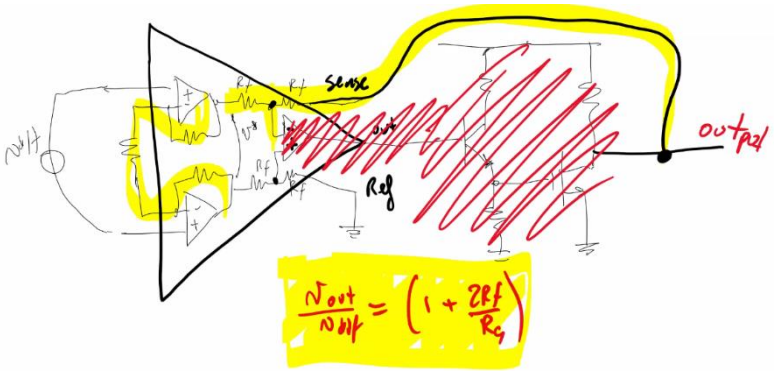


cannot attache a voltage partition network to the Vref pin without considering the resistance in input. If the impedance of the partition is smaller with respect to the 2Rf, it goes in series with the 2Rf and negligible, but if it is big, to avoid any voltage drop due to different impedance it is better to use another opamp (buffer) to connect.

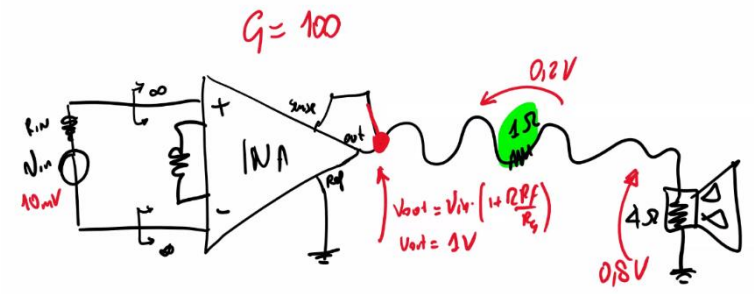
SENSE PIN

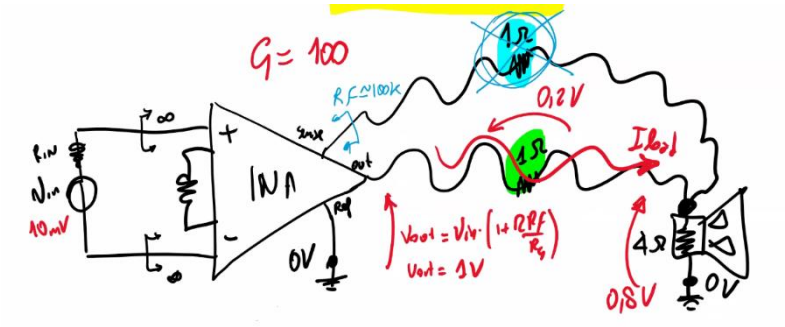
Whenever we apply an input signal Vdiff, we can compute the V* value and if sense is connected there, when Vref is grounded, Vout is Vdiff *G1*G2.

This equation holds whenever the sense pin is connected to the output. If the sense wire goes and touches another output, the gain of the stage is still independent on the A0 of the opamp and what we have after the output, because thanks to the feedback what sets the gain is REF.



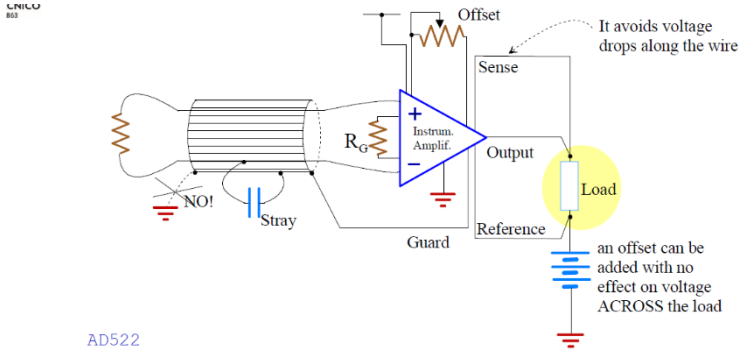
This is the reason why the sense pin and the output pin are kept separated. So it is important to connect the sense pin to the point of the network where we want our amplified output. Of course, also the sense wire has a certain resistance, but the difference is that in the output wire a huge current flows and the voltage drop across the 1ohm resistance of the output is high, but on the sense wire I have a small current because we have a 1ohm in series with the 2Rf resistance seen by the sense pin (hundreds of kOhms), so the wire resistance is negligible.





Moreover, if Vref pins are attached to two different ground, they are not exactly the same, so it is better to connect the two grounds also with the reference wire, because of the same problem of the voltage drop on the GND wire resistance.
We use the reference pin to go and touch the load as close as possible to solve this issue. In this way, we have something similar to the 4 wires connection.

To avoid further interferences, we can use also shielded cables.



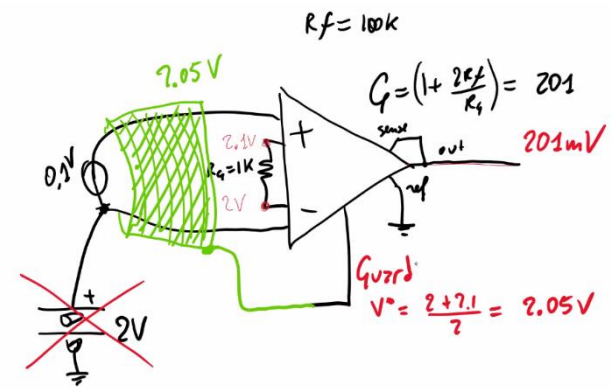
AD522

Vcc=10÷36V Isupply=10mA
Vos=1mV IB=25nA
Rin=1GΩ
GBWP=300kHz
$V_{noise\ in}$=15μV_{rms} (10÷15kHz)

GUARD PIN

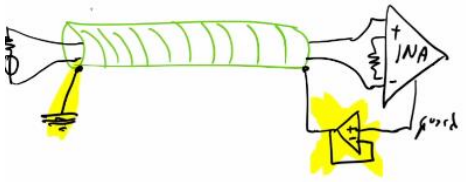
Used to provide a Vcm signal to the external world. This is done by adding two resistors in between stage one and stage 2, we take the average value, we buffer it an we provide it in the output with the Guard pin.

This is useful because if we have a small signal superimposed to a Vcm signal that is huge, the INA theoretically amplifies only the differential signal. So the Vcm signal can be used to drive the shield of the shielded cables that carry the signal.
If shield is at 2.05V, any possible parasitism between shield and wire will find the same voltages on the two sides, hence playing no effect.



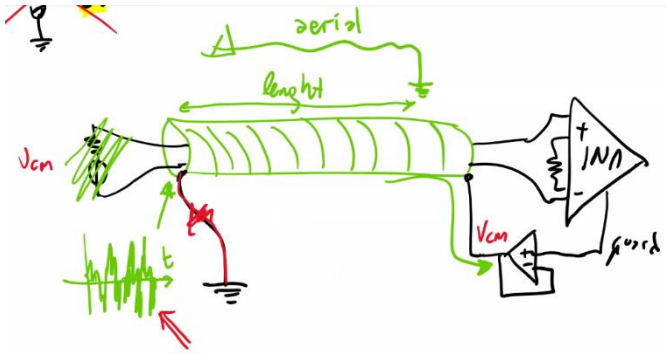
There is still an issue. Since one side of the shield is connected to Guard, it means that the shield is at a given voltage, so we cannot connect the shield to ground, otherwise we would destroy the INA. So ground pin of the shielded cable should be left unconnected. Eventually, we use an opamp to buffer the Guard signal so that if the GND pin is connected by error, this buffer burns but not the INA.

But if we leave the GND pin unconnected for the shield, since we are shielding for EM interference, if we drive the shield with the guard pin we cannot connect the other side of the shield to ground, otherwise we will cause shortcircuit. So what is recommended is to use another opamp in a buffer connection so that if by chance the shield is connected to ground, the buffer will destroy but the principal opamp is preserved.

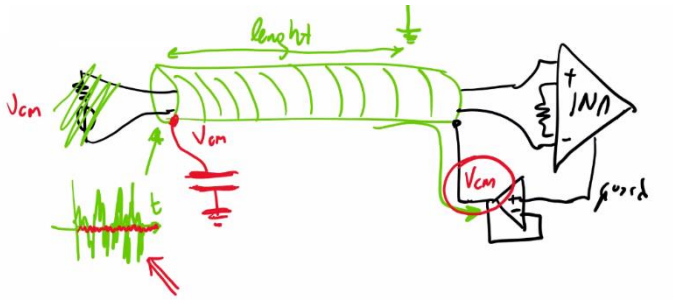


Moreover, if we use the shield to reject EMI, for the EMI the output of the buffer used to drive the shield is ground, so we should be ok. The problem is that every time we have a wire connected to ground, the wire acts as an aerial.

If the wire has a given length it receives all the EMI because it acts as an aerial, so it runs the risk of capturing EMI and the voltage of the source will be affected by disturbances. The best way to kill this disturbance is to connect to ground the source side of the cable but I cannot because I'm driving with Vcm from the Guard.



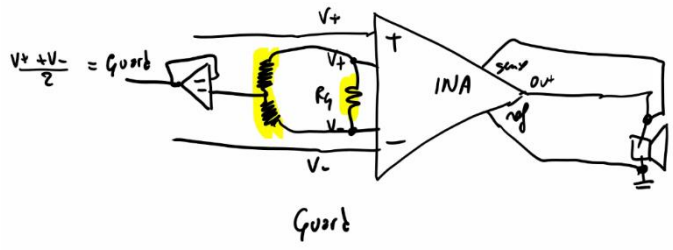
What we can do is to connect a capacitor to the other side of the shielded cable, so that the other side's HF component are killed by the capacitor.



Quality of the signal has to be preserved because the INA has a huge gain, for signal even in the order of 10uV, so any noise can have a huge effect on the output.

Missing guard pin

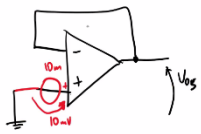
If however we have no Guard pin in the INA, we can do something different.
To create a guard pin we can take the V+ voltage and V- that are copied across Rg, so we use to high value resistors in parallel to Rg and take the value in the middle of them, that then is buffered. The resistance must be huge not to change the value of Rg and hence the gain of the INA.



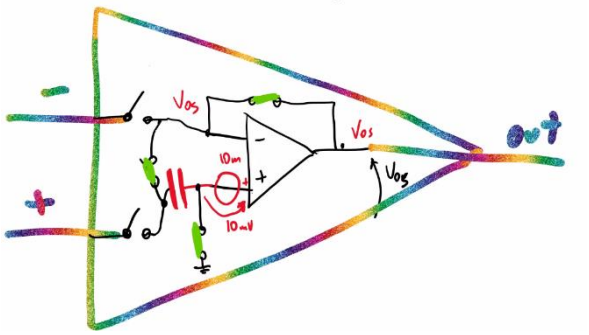
COMMUTATING AUTO-ZEROING

INA is usually used to achieve high gain, so I want the opamps in the INA to be very good opamps with low noise, offset and bias current. How can we buy an opamp with low offset?
There is a solution that is the auto zeroing.
We use a standard opamp even with a high offset voltage. We connect some components outside the opamp to measure the offset voltage and then subtract the offset voltage to the opamp itself.

If the opamp is buffered and it has an offset:



To subtract the offset I introduce switches and capacitors. I detach the input pins from the opamp itself. When they are open, I close the switch to ground and close the switch of the buffer. The output will reach the offset value. Then the Vos value in output that is buffered back is connected to a capacitor → I'm storing the Vos on the capacitor (I've closed the green switches).

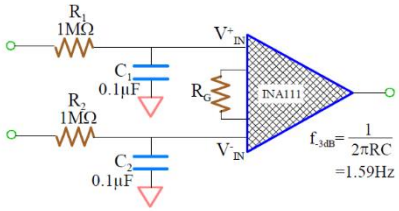


After this, I reopen the green switches and close the input path switches. Now I have a capacitor that charges to Vos in series (with opposite sign) with the offset itself and hence the two cancels out.
It is an approach used for the opamps in an INA.

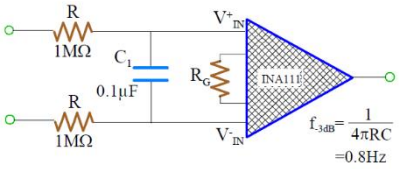
EFFECTS OF MISMATCHES IN INPUT PATHS

In the INA we should also preserve the symmetry of the stages, in particular of the first one. Let's imagine now we want to introduce a LP filter before the INA. The first solution is a bad solution because the circuit is no more perfectly symmetrical.

Tollerances of R and C degrade CMRR !!!



Clever solution (go fully differential):

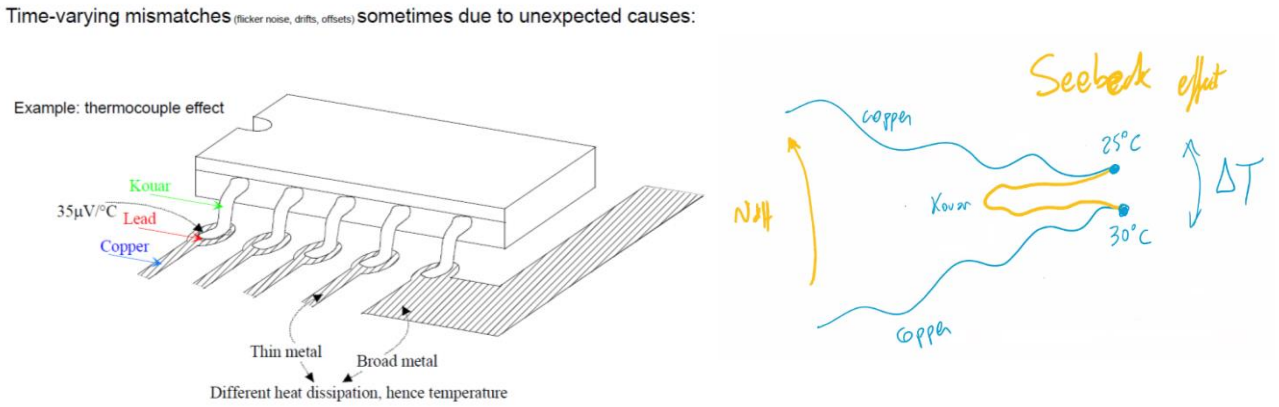


In fact the resistances R have their tolerances, and in the worst scenario they move in the opposite direction. Moreover, also the capacitors have tolerances and everything. So in the first case one path will have a certain gain, and the other path a different one due to the tolerances of components. Ideally the paths are equal but they have minor differences and the poles move.

If the common mode is not DC and it moves, it is rejected by the INA of course, but if we have mismatches in the case of input the filtering action will be different and so a Vcm signal eventually appears at the input of the INA as if it was a differential signal. But then the INA amplifies this and hence we would see in the output of the INA something related to the Vcm signal.
So either we use very precise components (too expensive) or we change solution, shifting to the second one. The resistors still have some mismatches, but the new situation can kill the signal.

TIME VARYING MISMATCHES

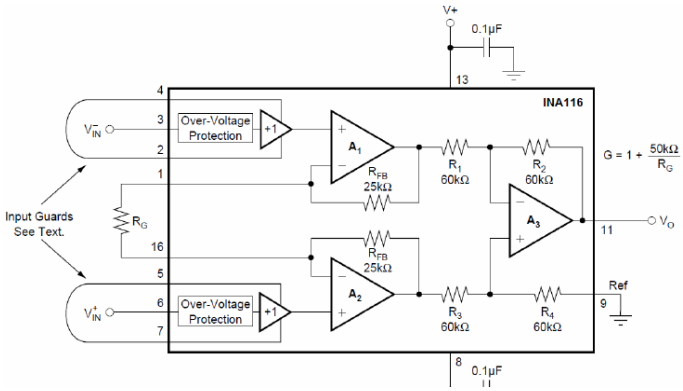
If we buy a very low noise INA, when we assemble it on a PCB, the traces are made of copper, the pins are made of Kouar and then we use lead to solder. Since we are soldering two metal wires, every time we have this, if there is a difference in temperature between one soldering and the other soldering, due to Seeback effect we generate a Vdiff → thermocouple.



If the temperature on the pins of the INA is different, we create a differential imbalance in terms of voltage and hence even a small change in temperature on the PCB can cause an output in the INA, because the INA is very sensible.

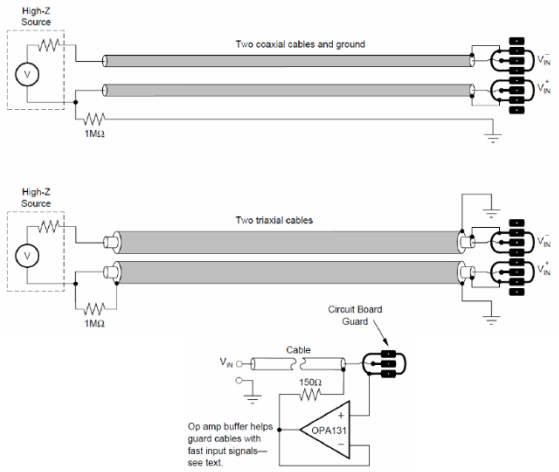
DATASHEET

We have name, image and suppliers name in it. Then there is a list of important features and a schematic of the inside of the component. Eventually, the manufacturer doesn't provide the sense pin, but sometimes we have an over-voltage protection.



We con for instance connect a source with two coaxial cables, one in the – pin and the other in the + pin of the INA, since the two input voltages could differ, the manufacturer can provide two different guards to drive the two different coaxial cables. To further protect the quality of the signal, we could use a triaxial cable (second image) with the final outer shield connected to ground.

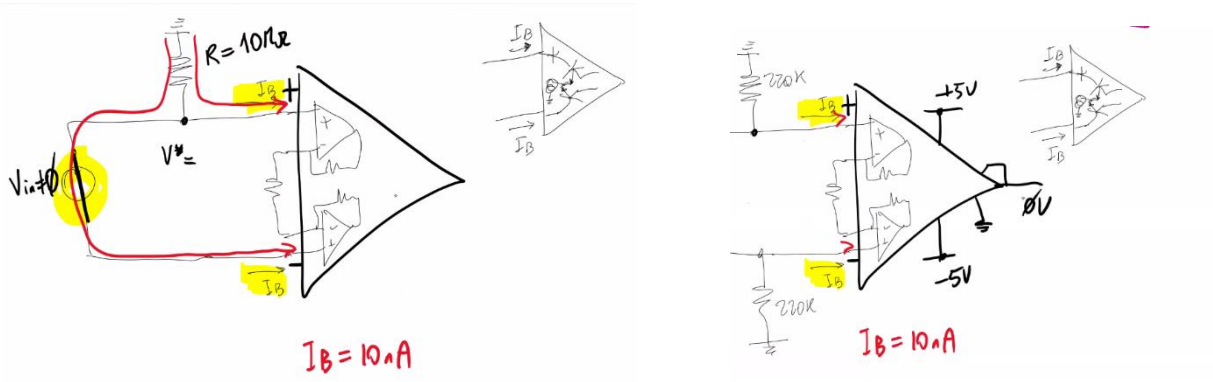
Shielding:



Voltage from a solution

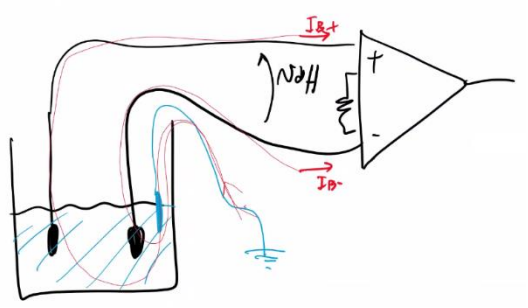
INA can be used often to measure voltage generated by two electrodes placed inside a solution. We have also an electrode that is the solution ground, without which nothing would work.
The opamps of the INA are mare by transistors, so they will require a bias current in input. This means that if we want the circuit to work, we need to go and touch the signal inputs with the bias. We have to introduce on resistor to ground because in this way if Vin is 0, bias currents come from ground and everything works (without the resistance it would be impossible). V* = -2V, which is a common mode voltage that can be removed by the INA, but to reduce it we could reduce the value of R to 100k, or improve the quality of the circuit by introducing symmetry.

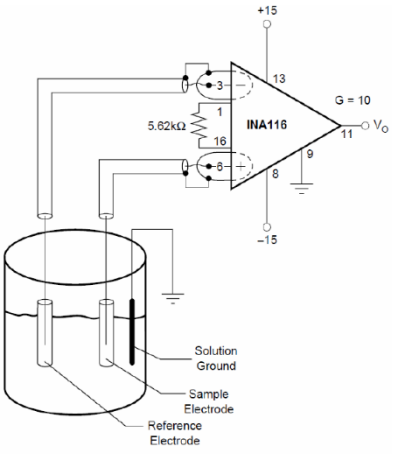
We have also to preserve symmetry, because if we detach the input and we use just a resistor, we provide Ib only on one pin but not on the other, and the output saturates to the value of the input (+ or -) where the output current enters. So we add another resistance to bias the – terminal.



Then, when we connect the source of the signal, the Rs should smaller than 2*220k, otherwise the voltage that we have in input to the INA will be affected by the resistive partition. So the resistances to be introduced should be big enough not to have voltage partitions.

Coming back to the solution, if I have a solution and give just the metal wires to the INA, the INA as this will never work because it doesn't have Ib. So we need another electrode that touches the water and goes to ground (blue one).

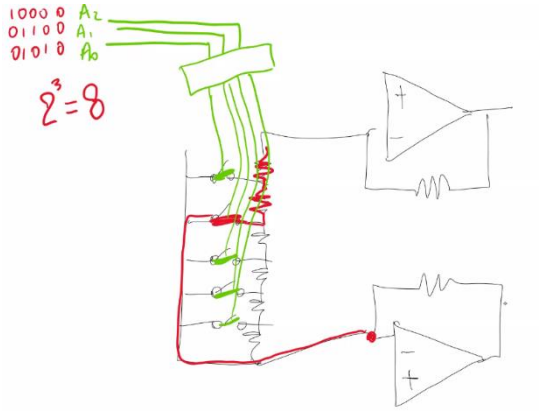


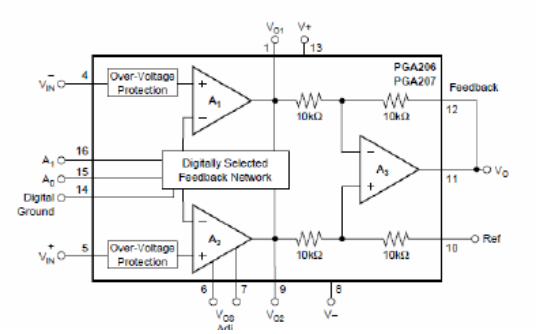


Thus the two Ibias can come and flow from it.

PROGRAMMABLE GAIN AMPLIFIER (PGA)

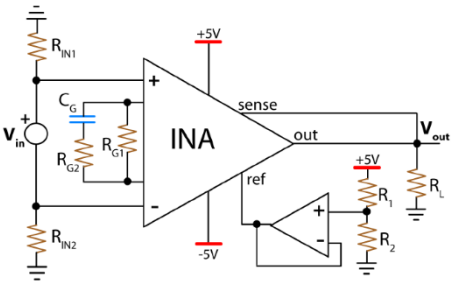
It's not just an INA because we have a selectable net of resistance inside. Basically, in place of a single Rg we place a parallel of resistances with switches, we use a decoder with digital inputs and we are done.





We could have used a classical INA and a series of selectable resistances depending on the closure or not of mosfet transistors. However, the problem is that if we have a Vcm signal, eventually if we use a uC to control the gate of it, we have to grant on the gate Vg = Vs + Vt to have it closed, but the problem is that in this case we are not sure, because if Vcm varies too much, it could be closed anyway.

Example 1



R_{IN1}=100kΩ R_{IN2}=100kΩ R_{G1}=10kΩ R_{G2}=1kΩ C_G=1.6nF R_1=3.3kΩ
R_2= 2.2kΩ R_L=1kΩ all resistors inside INA be 100kΩ

a) Compute V_{out} for V_{in} = 0V

b) Compute the ideal gain v_{out}/v_{in} at low and high frequency

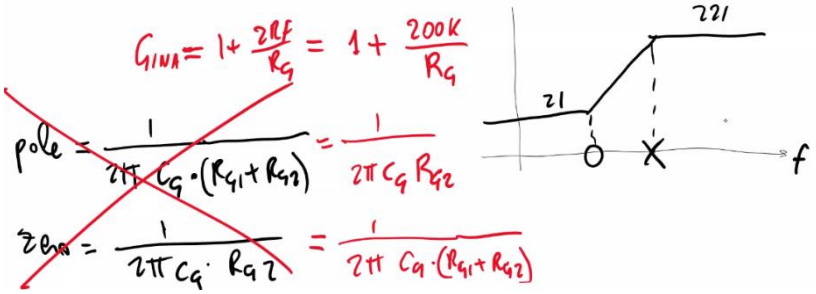
c) Compute poles and zeros

Resolution

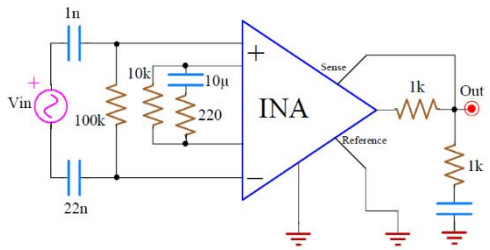
a) At DC output will be 2V because of the Vref signal.

b) G = 1 + 2Rf/Rg but Rg depends on the frequency because we have the capacitor Cg. At DC, Rg = Rg1 and G = 21, while in AC Rg = Rg1||Rg2 and G = 221. This is the gain up to where the sense pin touches. Now I have to compute the gain between Vin and Vdiff (input of INA), but since this gain is 1, Gina = Goverall.

Let's plot the Bode diagram and compute poles and zeros. As for the pole, the total R is NOT Rg1 + Rg2, because inside the INA we have opamps. So terminals of Rg1 are at 0 thanks to virtual grounds of the inside opamps (input has to be switched off when we study poles).
As for the zero, we do it with the GBWP.



Example 2



INA with R_f=100kΩ

a) Plot the Bode diagram for the $v_{out}(f)/v_{in}(f)$ **ideal** gain

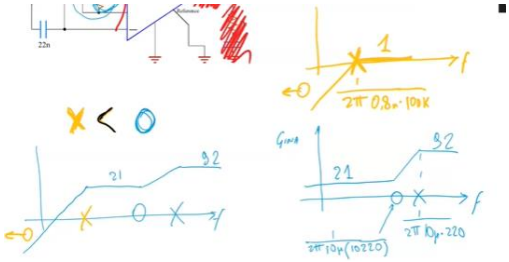
b) Modify the circuit to provide a DC gain of 1000 and a low-pass filtering action with 2 coincident poles at 100kHz

Resolution

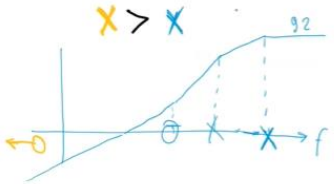
a) Reference is to ground, so Vout is 0 in DC. Let's see where the sense pin touches, and then we have Gina. I don't care about the presence of the 1k resistor. Moreover, since I touch the output there the impedance I see looking from the output is 0 because the output is a voltage source, so the output capacitor is not introducing a pole because the output is driven by a voltage generator whose output impedance is 0, so the output voltage is independent on the 1k resistance and the capacitor.

Now let's compute Vdiff with respect to Vin. 1n and 22n are in series → 0.8nF. Then we have a HP filter with the 100k resistance, with the pole that will be 1/(2*pi*0.8n*100k).
Then the INA has a certain network as Rg, so again gain will be low at DC and high at AC, so we will have a pole and a zero related to the Rg. Gina|DC = 21, Gina|AC = 92.
Pole is 1/(2*pi*10u*220) and zero is 1/(2*pi*10u*10220).

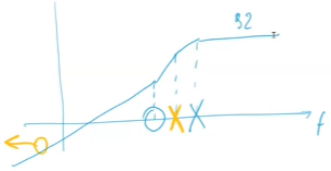
Now I have to combine the three gains.
If the pole is at a lower frequency than the zero, the Bode diagram will be the one on the left.



If instead the pole of the input capacitor is higher than the pole of the other capacitor, we have the following.



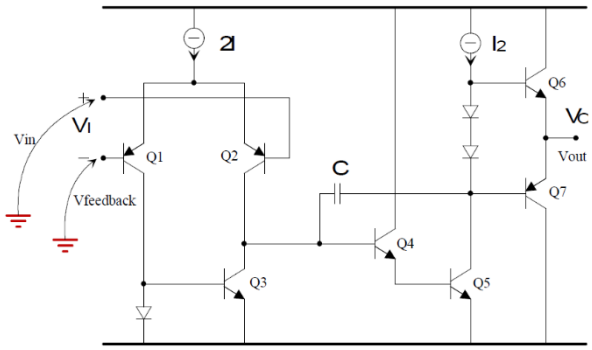
Last case, if the pole is in the middle:



CURRENT FEEDBACK AMPLIFIERS

VOLTAGE MODE CONFIGURATION

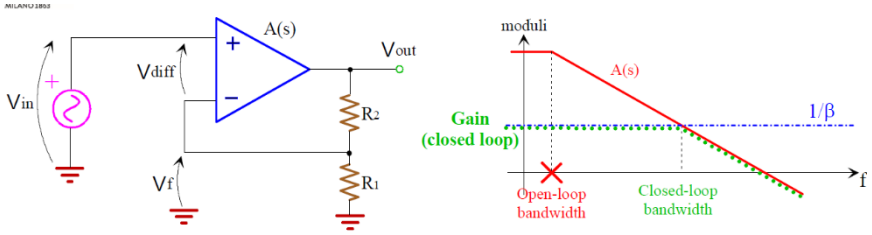
The basic idea is to use voltage mode opamps with two high impedance inputs (+ and -) and then the output was a voltage.



Requirements:

voltage-driven inputs	→	high impedance inputs
voltage output	→	low-impedance output

In this configuration we can introduce a feedback. If the Vin is applied to the +, on the other pin we apply a voltage with the feedback and thanks to the high feedback gain the error becomes 0 and we have Vin also on the other pin, and eventually the ideal gain is the one of a non-inverting configuration. Then we can study the frequency response of this structure.



Performances:

-20dB/dec slope

constant GBWP (hence trade-off between gain and bandwidth)

limited SR (tens of V/μs)

strict relationships among f_0 I_{tail} C_{comp} SR (again trade-off)

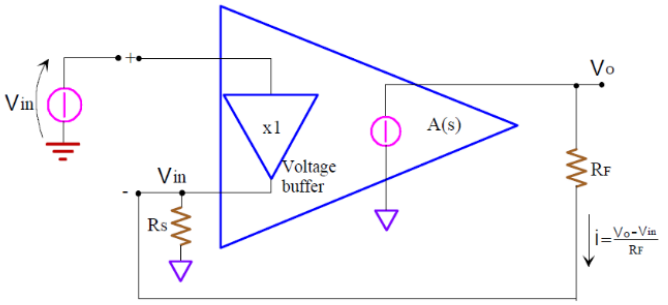
Thanks to the feedback, the pole of the opamp is pushed to HF, at f* = f0*(A0/(1+R2/R1)). Then G*pole = f0*A(0) = GBWP (gain is the one of a non-inverting configuration).

In the voltage mode configuration, the position of the pole depends on the gain. Than gain multiplied by pole is a constant number called GBWP. If we want a high gain the BW will be low, due to trade-off gain-bandwidth.
But we can change the way to introduce a feedback to reduce this trade off.

CURRENT FEEDBACK (CFA)

We design a brand new opamp where we have a positive input pin, **inside the opamp we have a buffer that reads the voltage on the + and provides it to the other pin, the – pin**. The voltage at this node is equal to Vin by design, independently on the feedback.

If then we connect a resistor Rs, the buffer should provide a current through the resistor Rs. Now **the second trick is to provide a voltage at the output at the opamp that is not proportional to the error voltage but to the current that the buffer has to provide on Rs**.



Requirements:

one high-impedance input	(hence voltage-driven input)
one low-impedance input	(hence input-current, i.e. Acting as an output)
voltage output	(proportional to the input current)

$$G_c = 1 + \frac{R_F}{R_S}$$

It looks identical to VOA, but completely different feedback action

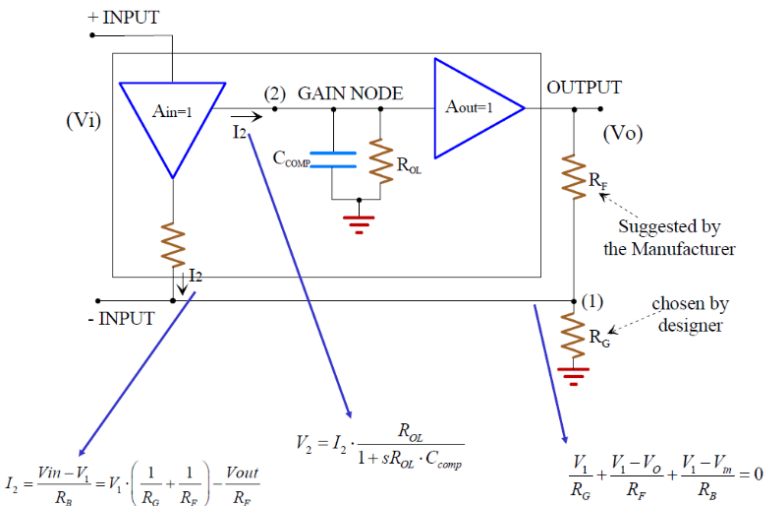
The opamp is no more the voltage mode opamp. If the gain is very high, even if the current is 1mA, the output voltage is high. But if the voltage in out is high, thanks to Rf, there will be a current flowing through it provided by the opamp because the – terminal has to remain to Vin. So now the current that should flow through Rs is not just provided by the input buffer but also by the output voltage generator. Eventually, the input buffer has to provide a small current if the output voltage is high.

Eventually, if the gain of the voltage generator that provides a current proportional to the current the buffer should provide is so high, the buffer has to provide a very small current. So in the end in output we will have a reasonable voltage that is the one needed to flow across Rs without the need of any current from the buffer. So it's the output voltage generator of the opamp that provides the current, not the buffer.

Vout is given by Vin/Rs * (Rs + Rf), the current that must be provided because the buffer sees Vin at the input (the current could be provided by the buffer but it is provided by the feedback).
In the end, the ideal gain is the same gain we get with the voltage mode non-inverting opamp.

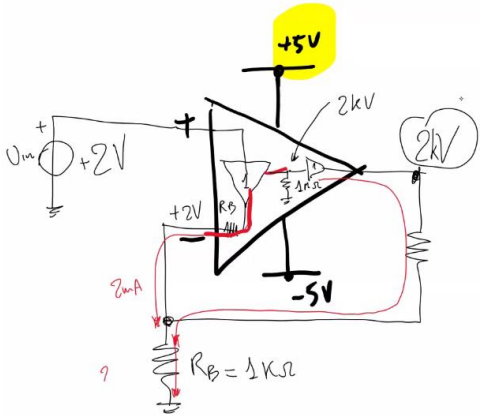
Let's design such a circuit.

CFA DESIGN



The buffer has his own output impedance Rb (that ideally is 0). Then let's provide a copy of the current that the buffer provides in output. Then i2 is converted into voltage using a high value Rol. Then this voltage is buffered and the voltage is provided to the output.

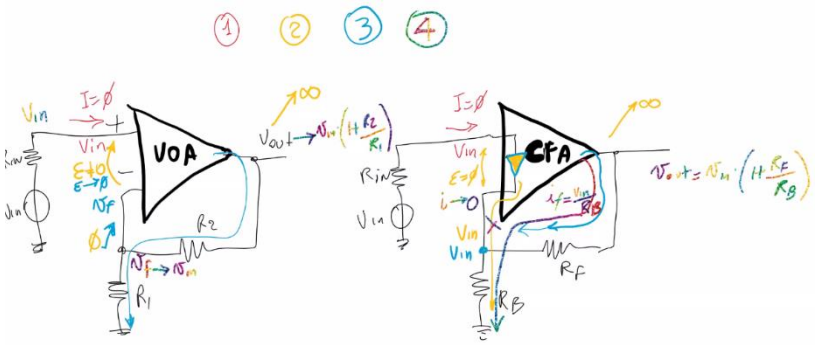
If this is our CFA, we have our opamp and if in input Vin = 2V, on the – we get 2V. If now Rb = 1k, the current of the buffer should be 2mA. But the current should not come from the buffer, but from the opamp power supply. Now we copy this current and feed it to a resistor and we buffer the voltage to the output. If the resistor is huge, e.g. 1M, the output will be 2kV, which is impossible because it will saturate to PS. Now I take this output voltage and use it to help provide the current in RB through Rg. Current in RB must be 2mA, but now I have two currents, one from the input buffer and one from the output buffer. So the input buffer can reduce its current provided. Going on, the output voltage reduces because the current in output to the input buffer reduces, up to the point the input buffer is not required to provide any current, because Rf will provide the current thanks to the internal gain of the amplifier.



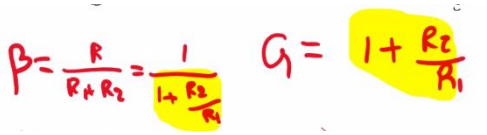
What is the right value of Vout? It's the one such that Vout/(Rf+Rb) = Vin/Rb (current we need).
Rf and Rb are not in series, but since there is the CFA and the opamp that has a smart gain inside such that Vout = iin*Rol, where iin is the current the buffer input has to provide, then if Rol is infinite is enough for iin to be 0 to have a value.

So iin (current provided by the input buffer) could be something but in reality is 0.
Let's make a comparison between the voltage mode amplifier and the CFA. In the first one the current in input to + is 0, but also in the CFA it is. Then it means that there is no voltage drop and + is Vin for both.
Now, for VOA, node – is 0 at the beginning, so there is an epsilon different from 0, so Vout increases a lot towards infinity, so also – increases to a given value Vf and epsilon decreases towards 0. Eventually, Vf goes to become equal to Vin so that Vout = Vin(1+R2/R1).

As for CFA, since there is a buffer, we find Vin at -. Epsilon is already zero. But then we must provide a current. This current is provided by the input buffer, but this causes the output voltage to increase to very high values. – remains fixed at Vin, but the current through Rf (provided by the opamp) increases. Finally, due to the gain, the current through Rf increases so much that I no longer need to have a current fed by the input buffer and if = Vin/Rb.



From an external point of view, they both provide the same gain. If we study the beta in the VOA, it is as below.

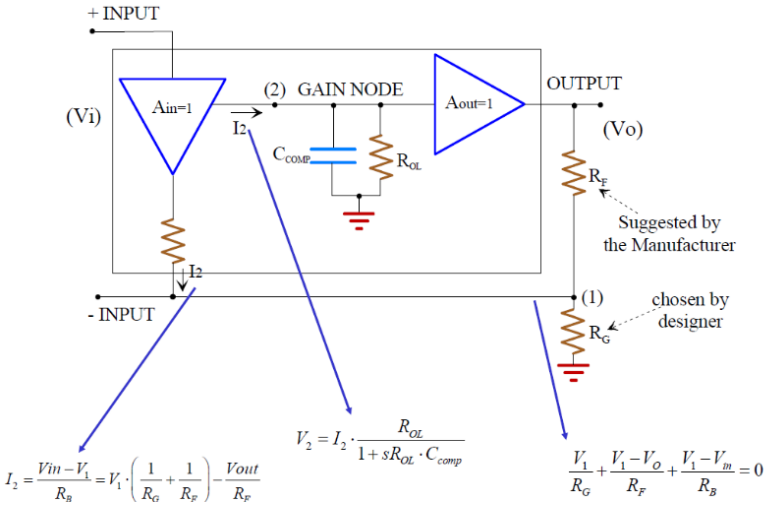


Hence in VOA the beta is very much related to the gain. If we want to change the gain (by changing R1 or R2) we change also the beta and hence we change also the pole.

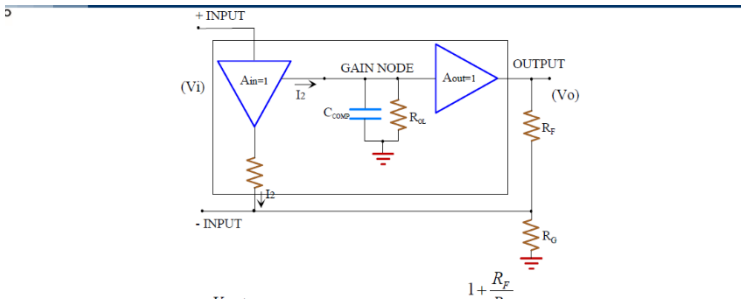
As for the CFA, the beta depends only on Rf and not on Rb. **Hence if we change the gain by changing Rb, Gloop doesn't change and the position of the pole remains the same**.

CFA MODELLING

Let's try to solve the KVL.



Ccomp introduces a pole, then the voltage on Rol (V2) is buffered at the output. Once we have Vout, (Vout-Vin)/Rf is the current in Rf. Then we can find the current in Rg and the one in Rb.
The final equation Vout/Vin is really a mess. But we can understand it. It says that the gain is still 1 + Rf/Rg, but this quantity is divided by a correcting factor. However, if Gloop is very big (hence Rol very high), ideally the denominator is 1. Then there is a pole of the Gloop, due to Ccomp and Rol.



Closed-loop Gain: $\frac{Vout}{Vin} = \frac{1+\frac{R_F}{R_G}}{\left(1+\frac{R_F+\left(1+\frac{R_F}{R_G}\right)\cdot R_B}{R_{OL}\cdot A_{out}}\right)\cdot\left[1+s\frac{\left[R_F+\left(1+\frac{R_F}{R_G}\right)\cdot R_B\right]}{A_{out}+\frac{R_F+\left(1+\frac{R_F}{R_G}\right)\cdot R_B}{R_{OL}}}\cdot C_{comp}\right]}$

The new pole is with a new tau that is not Rol*Ccomp.
If we look at the tau, we realize the pole, if Rol is very big, what remains is Aout, that however is a buffer so it is one. So if Rol is big, the real gain is the ideal one and the new pole has a certain dependency on resistances, as below.

Bandwidth: $f_{pole} \cong \frac{A_{out}}{2\pi\left[R_F + \left(1+\frac{R_F}{R_G}\right)\cdot R_B\right]\cdot C_{comp}}$

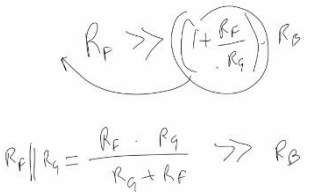
For low gain ($R_B << \frac{R_F \cdot R_G}{R_F + R_G} = R_F \| R_G$) we get $f_{pole} \cong \frac{1}{2\pi \cdot R_F \cdot C_{comp}}$

Bandwidth depends only on R_F, not on R_G, hence *NOT ON GAIN* !

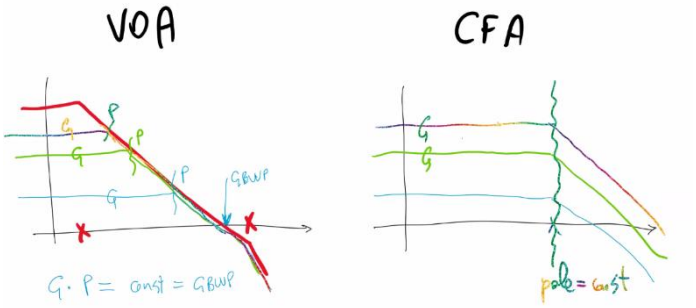
Instead for high gain (>50) we get $GBWP = \frac{A_{out}}{2\pi R_B \cdot C_{comp}}$

As for VOA, again trade-off Gain &Bandwidth (i.e. constant GBWP)

If Rf >> than the rest inside the [] brackets:



Rf||Rg should be in the range of 100 ohms, so we design an input buffer with Rb << 100 Ohm. If so, what remains in the [] is just Rf. So the pole depends on Rf and Ccomp. Ccomp is inside the opamp. If Rf = 1k and Ccomp = 1p, fp = 160 MHz.
Thanks to CFA the bandwidth is huge and set by Rf, and the gain is changed by changing Rg.
Let's make a Bode plots comparison (colored ones are the gains). In VOA, the higher the pole gets, the smaller the pole, because GBWP = const.



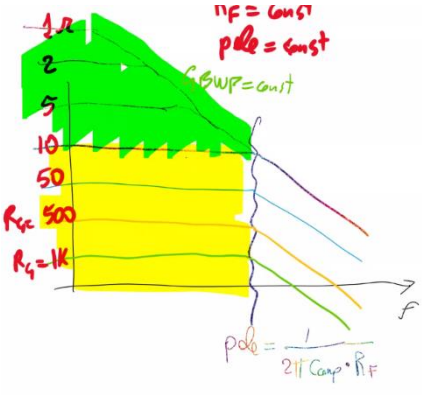
NB: In CFA, we must not change the gain with Rf. However this holds if Rb << Rg||Rf. But if we want to go to high gains, Rg becomes so low that eventually it is not granted that the previous inequality holds → valid for not to high gains.

If the gain is high, then it is not true that Rf prevails in the pole equation, but it is the opposite, so Rf is negligible with respect to the rest.
To know the value of the GBWP, we get (the only parameter is Rb, the others are constant and inside the opamp):

Instead for high gain (>50) we get $GBWP = \frac{A_{out}}{2\pi R_B \cdot C_{comp}}$

As for VOA, again trade-off Gain &Bandwidth (i.e. constant GBWP)

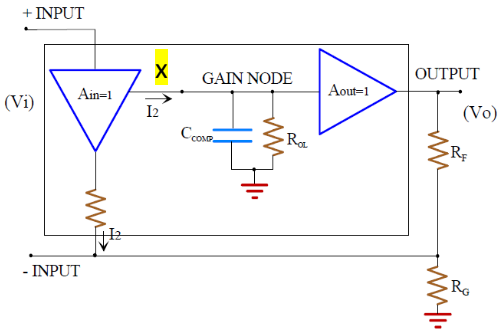
At high gain, the CFA configuration becomes equal to the VOA, so we are in a condition where the GBWP remains constant as the VOA, but I want to remain in the region where I can change the gain while the pole remains constant.



CFA BANDWIDTH

We can also study the CFA in a standard way we have always studied them.

Let's study the feedback in a smarted way...
...same result at the end!



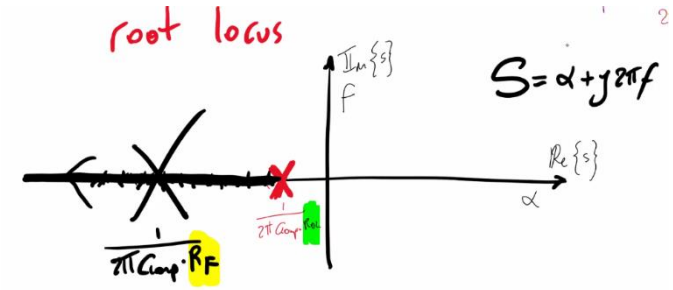
With open-loop: $pole_{openloop} \approx \frac{-1}{C_{comp} \cdot R_{OL}}$ $\quad G_{loop} \approx -R_{OL} \cdot A_{out} \cdot \frac{1}{R_B + R_F} \approx -\frac{R_{OL}}{R_F}$

… hence, when loop is closed:

$$pole_{closedloop} = pole_{openloop} \cdot \left(1 - G_{loop}\right) \approx \frac{-1}{C_{comp} \cdot R_{OL}} \cdot \frac{R_{OL}}{R_F} = \frac{-1}{C_{comp} \cdot R_F}$$

NB: Vout must not to be partitioned on Rf and Rg because it is not Vout that dominates and determines the Vf to be fed back. Here the middle of the voltage partition is already set to Vin.

Let's study the circuit and get the CL pole of the configuration. The OL pole is Ccomp*Rol (on the two links connected we have high impedances).
The root locus is a plot where we plot where the poles are in the S plane, being S the complex frequency and real part.
If we know where the OL pole is, if we close the loop the CL gain pole can be somewhere on the left. In fact, the possible poles on the CL config lie down on the left. The higher Gloop gets, the far away the CL pole is. The CL pole will depend on Rf.

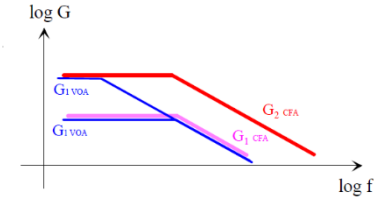


As for Gloop, let's cut at the output i2 (x). I pump i_test and we are in DC, so Ccomp is open. we can get the voltage Vout that is i_test*Rol*Aout. If Rg is larger than Rb, when we study Gloop, we have

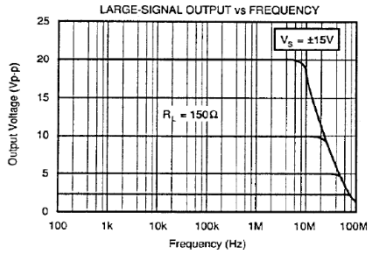
Rg||Rb, but Rb wins the parallel and so we can say Rf + Rb (Rb wins in the parallel with Rg). Then we have that the current generated on this series of resistance is equal to i2. So we can make the KCL and get the Gloop. Then Rb + Rf is almost Rf because Rb is very small.

So the CL pole is equal to OL pole multiplied by 1 – Gloop. **Pole_CL = Pole_OL*(1 – Gloop)**. Gloop is negative in general.
So the closed loop pole is given by Ccomp*Rf, when the loop is active (not when the loop is open). so the previous result is valid, always remembering the assumption on Rb.

In conclusion, at low gains...

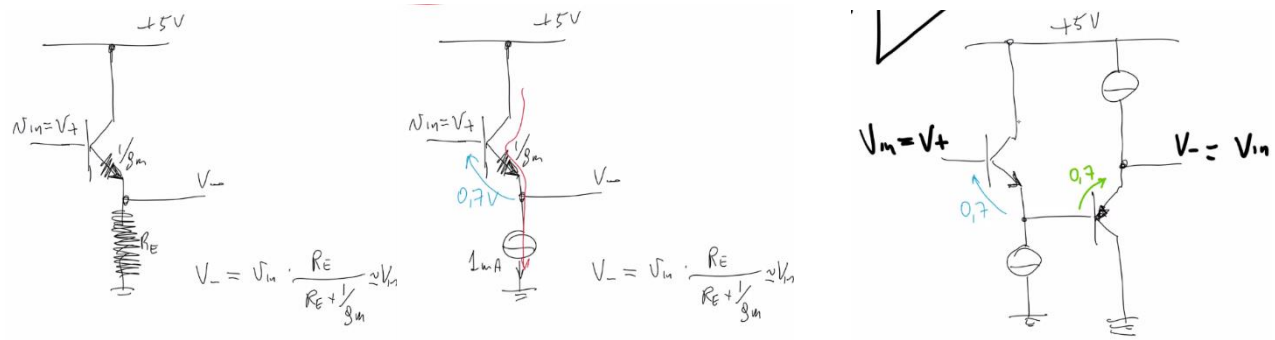


Instead at high gains...



REAL CFA ARCHITECTURE (to be known)

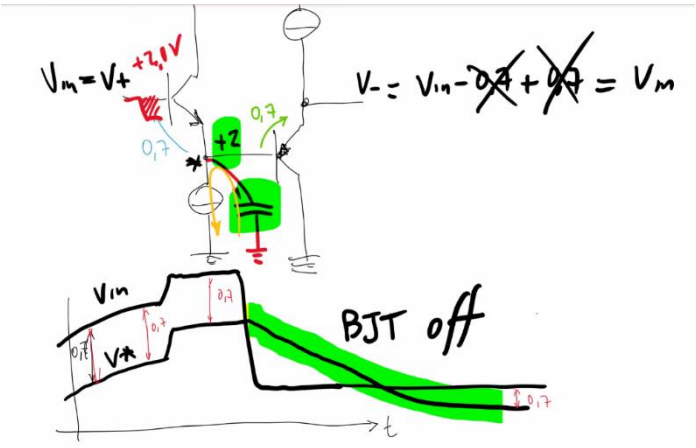
I want a buffer with a gain 1, and I take a BJT where I enter with Vin at the base and I take the signal on the emitter. If Re is very high, much higher than the 1/gm of the BJT, gain is 1.



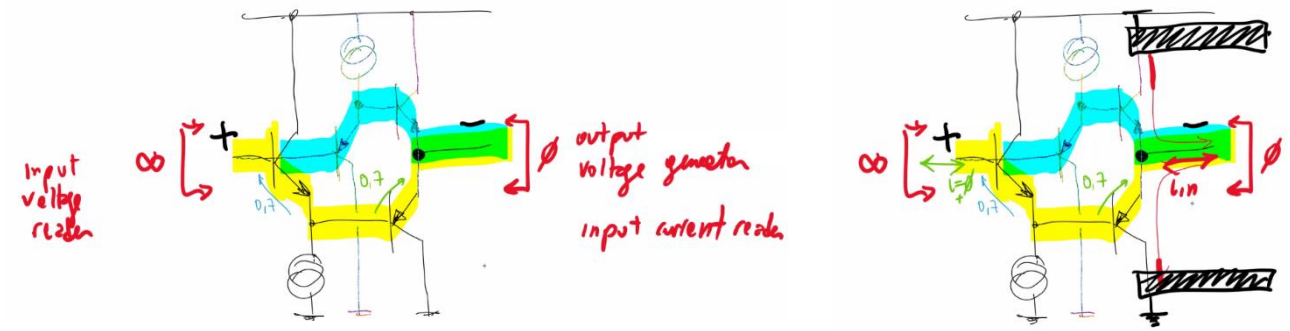
But if Re is excessively high, we have not to use a resistance, otherwise even with a small current we get too much voltage drop on Re. So we use as a resistor a current generator. In small signal analysis it's a buffer, but in big signal analysis, Vout = Vin – 0.7V, but I need a buffer, I cannot loose 0.7V.
To regain the 0.7V we use another buffer but we loose 0.7V in the opposite way.

But there is an issue, if Vin goes positive, Vin increases in output, but if Vin decreases the output decreases but a NPN configuration is happy if the base increase, because the emitter increase (first buffer), since we have a parasitic capacitance that takes the charge. However, if we decrease the base of the first buffer, e.g. we go down from Vin = 2V, the capacitor is still at 2V and the capacitor discharges slowly thanks to the current generator that drinks the current. So if Vin drastically increase, V* copies, but if Vin drastically decreases, V* goes down and reaches the 0.7V gap after a long time because there is a slow discharge of the capacitor because ethe BJT goes off.

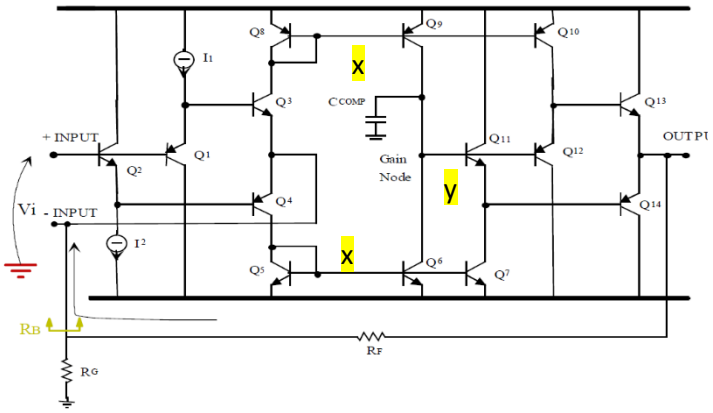
Hence the transistor is smart in going up put not in going down. To have symmetry in the circuit, let's duplicate the input stage.



Let's introduce a P-channel with a current generator and another buffer.



So input required current is 0, and the output current is the one that flows out in Rg. Then I need a copy of this current to be provided to Rol. So the i_in current comes either from the top or bottom branch of the output BJT couples. So we place current mirrors (black, x) to copy the currents.

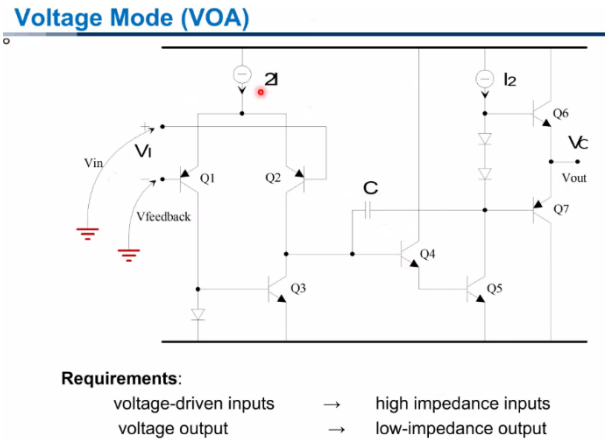


Complementar npn+pnp darlington configurations, hence better SR...
but problems of matching, hence higher offset voltage

Q2 and Q4 is the first couple of followers, the other one is Q1 and Q3. The collector of Q3 should go to PS, and the collector of Q4 to ground, but instead I use current mirrors.
So the first input buffer is created. Now I have to add Rol and the buffer output. As an output buffer I copy the output buffer, but there is a difference. I use a current that is the current proportional to the one that we have in the mirror. The impedance we see from y is beta*1/gm parallel to something both in Q11 and Q12. Since we want the highest possible Rol, the only Rol that suits the work is the parasitic Rol. r09||r06||(r_in of the output buffer) is the Rol. Rol will be in the range of kOhms at the most. To increase it I can increase the mirroring factor if the x mirrors, increasing the area of the Q9 transistor. So we will have a lower Rol but a higher current that is like having a higher Rol in the end.

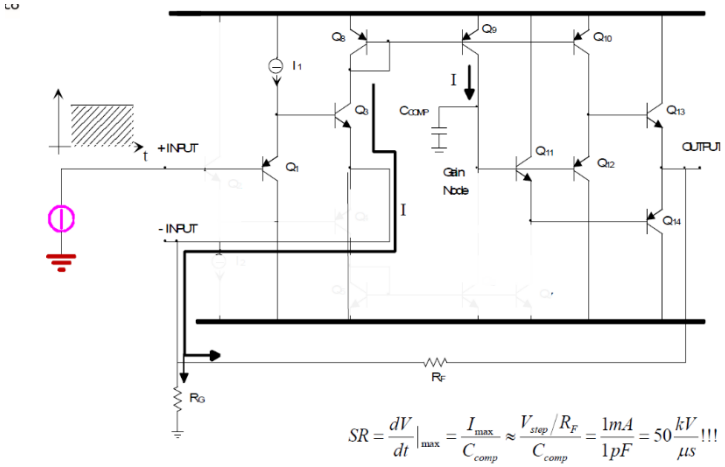
CFA SLEW RATE

In the differential pair VOA we pump a constant current (2I) that is splitted evenly but then we have the capacitor C and so the pole of the configuration is set by the Miller effect with the capacitor C. The SR of a VOA depends on the current 2I, and also the bandwidth depends on it.



Instead, in a CFA imagine we apply a big step in input. This step is applied to the other pin; since I apply a big step in voltage, the current that flows is the step we gave divided by Rf. But this current is suddenly mirrored by the mirror and flows in Ccomp giving a dV/dt = SR.

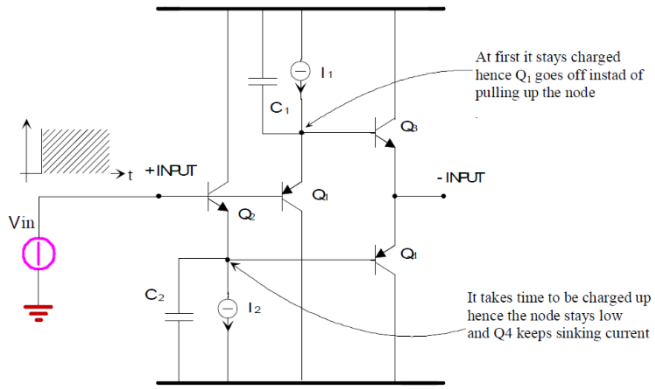
So the Slew Rate that is dVout/dt depends on a current that is not constant, because it gets higher if it's high the step we apply, it is not constant as it was in the VOA that is constant.
So the CFA is good because it has a high slew rate.



$$SR = \frac{dV}{dt}\Big|_{max} = \frac{I_{max}}{C_{comp}} \approx \frac{V_{step}/R_F}{C_{comp}} = \frac{1mA}{1pF} = 50\frac{kV}{\mu s}!!!$$

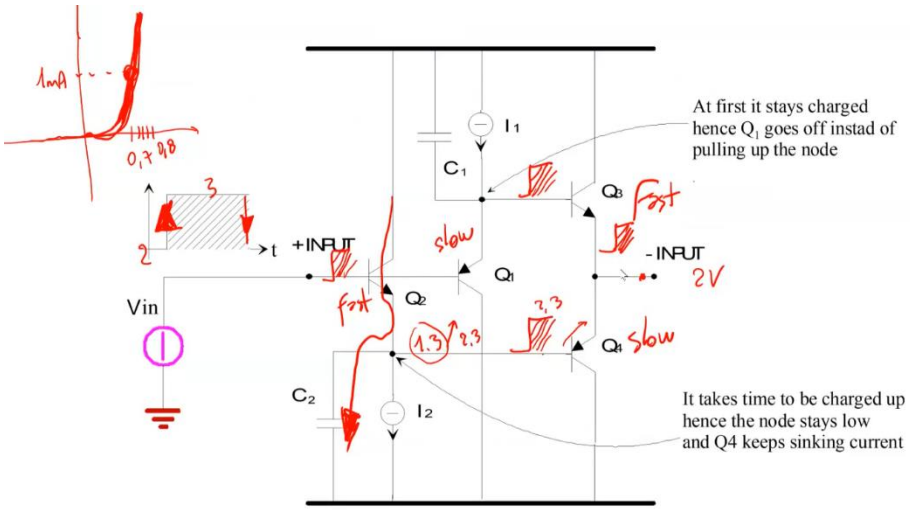
CFA TRANSIENT RESPONSE

Indeed there are second-order effects, which degrade SR



This is the real circuit. We apply a big step (e.g. from 0 to 2V) and we have a npn transistor, which is good, because if we increase the base, we have 0.7V across base and emitter. Then if we increase further the voltage (up to 3V) at the base, also the 0.7V wants to increase, and the current in the transistor Q1 increases too much, because we have the characteristic of a diode, and above 0.7V if we increase the voltage a bit, we increase also the current but a lot more, so we quickly charge the capacitor C2, so base of Q4 rises rapidly and the emitter goes to 2.3V (from 1.3V previously). Unfortunately, now Q4 is slow, because we have -0.3V across Q4, because the output previously was 2V, so Q4 is off. So to move up the output, we have the Q3, which is fast. In fact, npn transistors are fast in the rising edge transition, pnp not (due to parasitisms), but they are in the falling edge.

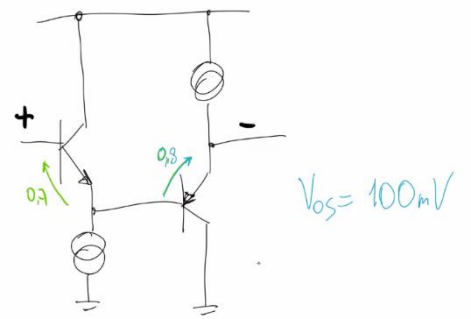
In this case we have the parallel combinations of two paths, fast and slow and slow and fast. So I don't have a fast-fast path to provide a SR.



ISSUES

The input buffer is ideal, but this is not perfectly true.

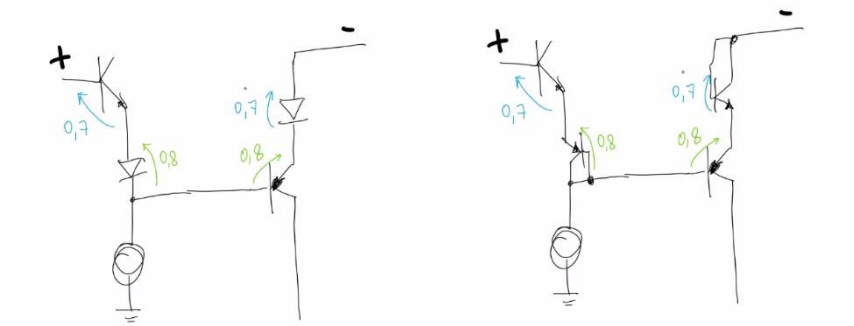
If we consider just one path, the one below, we have 0.7V on both. But this is not true, because pnp and npn transistors are made with different doping, so ther might be a mismatch between the two values. If so, there is an offset equal to 100mV between input and output. I can compensate this mismatch by changing the current in the two BJTs, so we reduce the current in the second BJT.



But again, this is not true because the I-V characteristic for a pnp or npn transistor changes if the temperature varies. So even if I find the perfect matching with currents to cancel out the offset, temperature changes create a problem.

We have to avoid this offset due to the mismatch between npn and pnp transistors. So we can try to improve by using not just two transistors but more.

The idea is that, since one transistor needs 0.7V and the other one needs 0.8V (in our example), instead of using just a current generator let's use a diode, and for the second stage we use the pnp transistor and we recompensate the drop we introduce with the diode. So the npn needs a voltage of 0.7V, the pnp of 0.8V, why not using diodes to compensate? Yes, I design a diode made by transistors of opposite polarity.

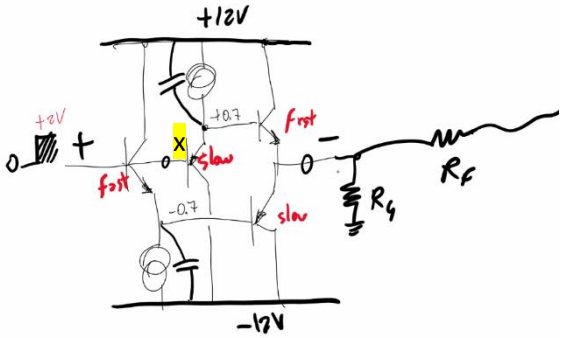


Now Vos is ideally 0, the transistors are all bad but they compensate each other. So both on the bottom side and top side we use a diode. The red improvement of the next image is very good. And we added it also at the bottom.

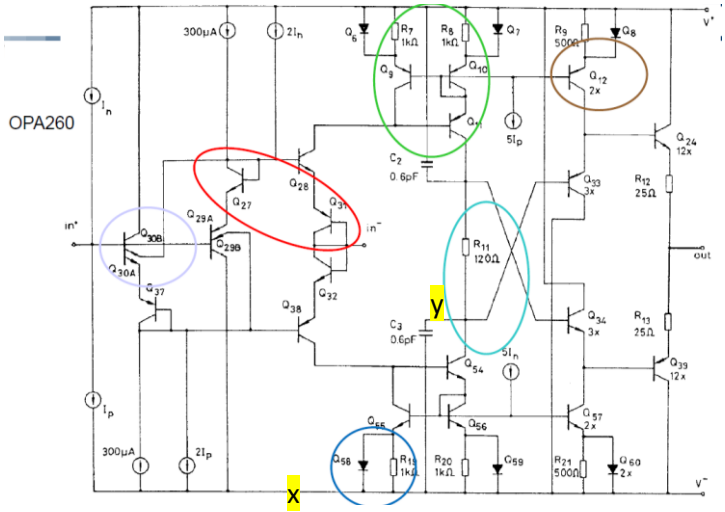
Then we have another issue, let's forget about diodes for now. Imagine now I increase the + voltage node (we have also the twin path). Usual the PS is dual, so at the bottom top we have +- 12V. Externally then we connect Rg and Rf.

When input is 0 (black values), we have some parasitic capacitances, but no current at all. If we increase to +2V, some BJTs are fast, and some are slow.

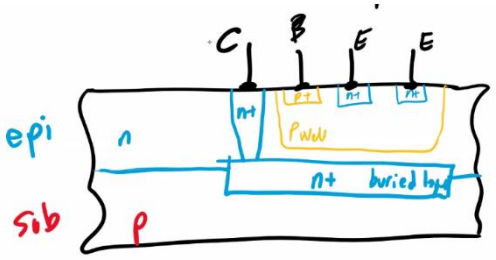
When we move from 0 to 2V, the x BJT goes off and the upper capacitor is discharged through the current generator.



To speed up the circuit, if we go to 2V suddenly, we face immediately a slow transistor after the fast one. So we can use a double emitter transistor.

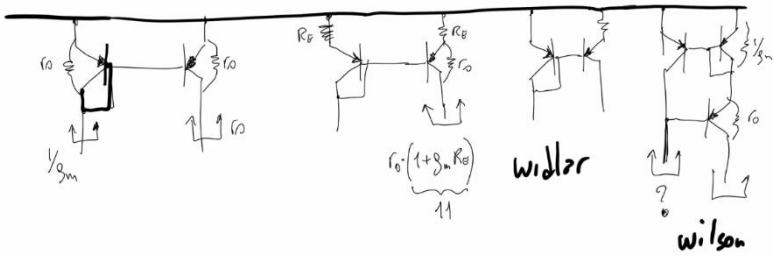


Double emitter transistor (npn)



We use the second emitter and we connect it to the emitter of transistor x of the previous image. In the schematic is the purple circle. If the input voltage increases too much, the second emitter activates and we pump current into the node of Q28 speeding up the circuit. Thus we drastically increase the SR.

Now I want to mirror, in the previous configuration we relied on it, and we want also to improve the value of Rol that was in the order of kOhms.
We can use different mirrors, like the classical one or an improved one placing a Re.



By improving the output impedance we improve Rol, the impedance of the gain node. In the schematic it is the green circle (Wilson), where the diodes further improve performances. In fact, if we increase the voltage at the base, we have the 1/gm and a resistor, and if the base is increased to very high values, some portions of the voltage drops across the base-emitter and a portion across the resistor. But if the base is increased very much, soon or later we would love the voltage we apply to fall totally as Vbe. For this reason we introduce a clipping diode. Hence for low values of Vbe (0.7 – 0.8V) we are ok, but then when the voltage is high the diode clips the voltage at the maximum value of 0.7V, so the emitter is clipped at 0.7V, the Vbe has a maximum value of 0.7V and hence the vase cannot go above 1.4V compared to the bottom level x, otherwise the transistor starts drinking a huge amount of current that it is okay that because it helps in pumping a huge amount of current from the gain node and then get the amplification.

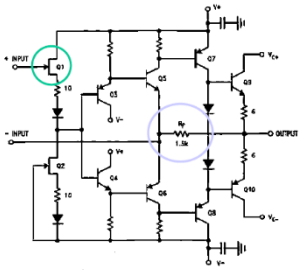
Finally, the light blue circle. We would like to have the transistor always on, no output current but on to have a low parasitism. To achieve so, the pnp has to be connected to a slightly lower voltage (y) and the npn at a slightly higher one. To do the splitting I use a low value of resistance, R11.

CFA OTHER ARCHITECTURES

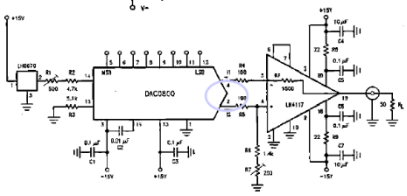
It is the same configuration as before. The buffer with Q3,4,5,6 is the same, then we have a buffer Q9,10 and they introduce Rf inside the opamp, we cannot choose it. This is ok because Rf has to be kept constant to get the pole we want, then we change Rg.

In our initial modelling we considered just Ccomp, so there will be for sure some parasitism due to the buffers, and because of this there will be other poles that limit the bandwidth and the manufacturer knows that eventually if we shift the pole due to Rf, eventually the BW is limited by the other poles due to parasitics. So the manufacturer directly places Rf in the CFA.

Data-sheet LH4117:

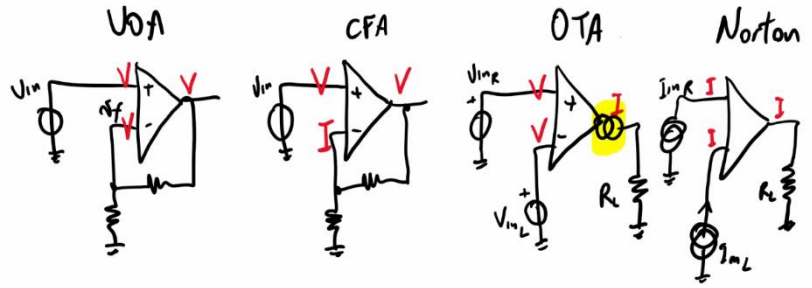


Application-note:

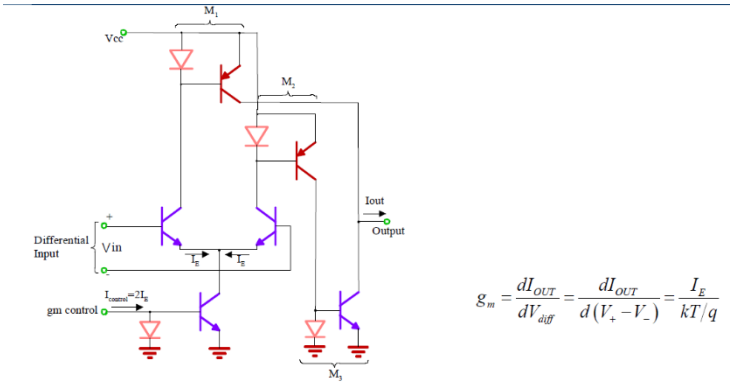


OPERATIONAL TRANSCONDUCTANCE AMPLIFIER – OTA

It is neither the VOA, nor the CFA. The VOA has volts in input and provides volts at the output. The CFA has voltage in input an output is again a V. Now we apply a Vin in two positions and we want a current in output. Then there will be the amplifier where the inputs are currents and the output is a current as well.

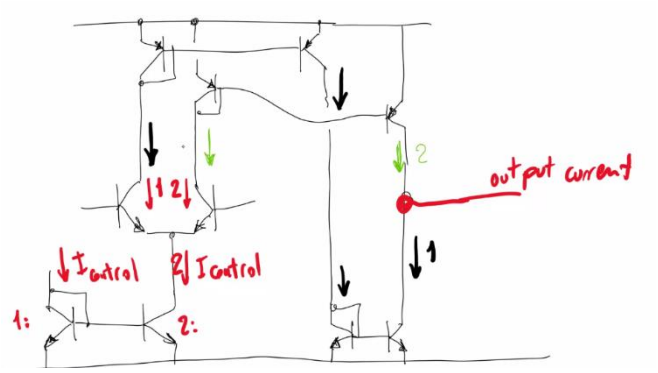


OTA is a differential pair with a tail current that is provided with a current mirror, and the diode is the current mirror.



$$g_m = \frac{dI_{OUT}}{dV_{diff}} = \frac{dI_{OUT}}{d(V_+ - V_-)} = \frac{I_E}{kT/q}$$

both inputs be high impedance (like VOA)
output be current (like Norton)
transconductance gain (I_{out}/V_{diff}) adjustable through a pin!

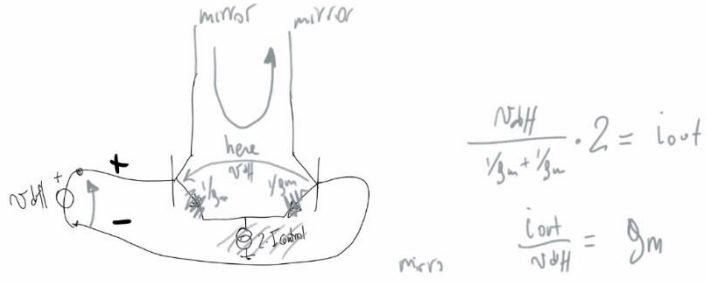


We can redraw the circuit as above (using active loads). We need to properly subtract the black and green currents. If I pump I_control, then due to mirroring ratio I have 2I_control, and the output current should be I1 (black) – I2 (green). To compute the difference we use another current source mirror at the bottom.

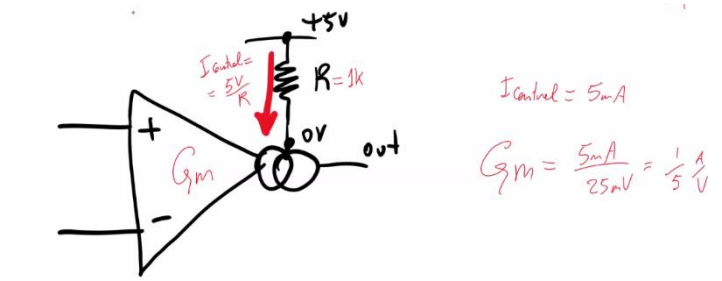
If no input is applied, the output current is 0, but if we have an imbalance in input we have output current. To compute the value of the output current, we shift to small signal analysis and the input pair will have a 1/gm = Vth/I_control (all in DC the I_control) and Vth = kT/q.

SMALL SIGNAL ANALYSIS

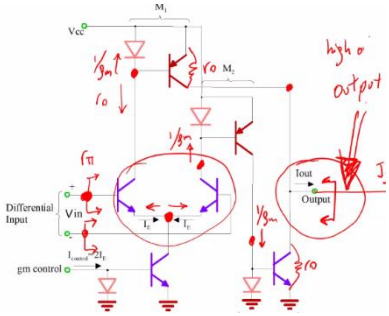
Current generator 2I_controll will be open in AC signal analysis.
The signal current that we have, thanks to the mirror, is doubled and we have the output.



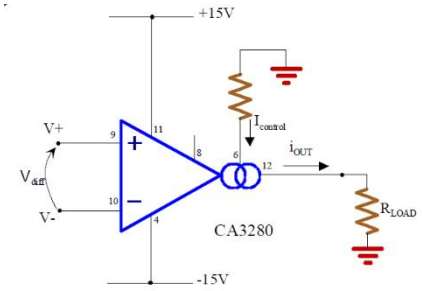
gm is set by the I_control, so we can change the gain by changing the I_control. The output pin is the one with which I select I_control.



Ideally, if we enter in the circuit above with 2V on the + and 2.3V on the minus, the voltage difference is 0.3V that, multiplied by that value of Gm will give a i_out = Gm*v_diff.
It is like a voltage mode amplifier cut in half, because a voltage mode amplifier has another stage after that converts the current into a voltage.



All the internal nodes of an OTA are low impedance, the only high impedance node is the output, that shows an impedance that is the r0 of the two transistors in parallel.
Hence in an ota we have a very high input and ouput impedances.
We have also another pin, the control pin (I_control, also called I_ABC).



$$i_{out} = G_m(I_{control}) \cdot v_{diff} = \frac{I_{control}}{kT/q} \cdot v_{diff} = \frac{I_{control}}{25mV} \cdot v_{diff}$$

Advantages:
all low-impedance nodes
wide bandwidth
$f_{pole} = 1/2\pi C_{LOAD} R_{LOAD}$

Disadvantages:
no infinite gain
no "VIRTUAL GROUND"
it is used open-loop

Gain of the OTA

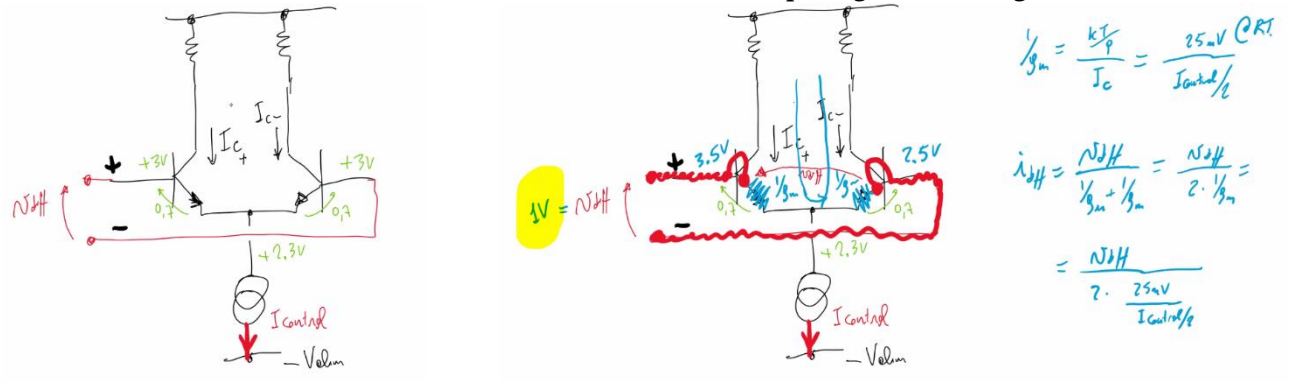
It is based on the differential pair. We apply a differential signal between the two input transistors, and since they are biased with certain currents, if we measure the two collectors' currents, I+ and I-, if v_diff is 0, the two collector currents are just the biasing currents, so the tail current I_control is split in half in the two sides (left).

If we apply a v_diff, the two transitors are slightly modified, we bias them with different voltages. We can study the circuit by linearizing the behaviour of the transistors, using the small signal equivalent. But this is a wrong assumption with a v_diff of 1V. anyhow, let's try to do this. The transistors are modelled with the 1/gm. A differential current will be generated with the value in the following image.

However, in an OTA we don't simply use a resistive load but a current mirror, so that the current is mirrored and the i_out = 2*i_diff = Gm*v_diff.
In the OTA, the Gm = gm = (I_control/2)/25mV.

The I_control/2 depends on the design of the OTA. If the mirroring factor of the mirror is 1:1, if the tail current is I_control we will have half it in each branch.

In the linearized analysis we found that the output current is proportional to v_diff and the transconductance of the OTA. However this is not true if the input signal is too high.

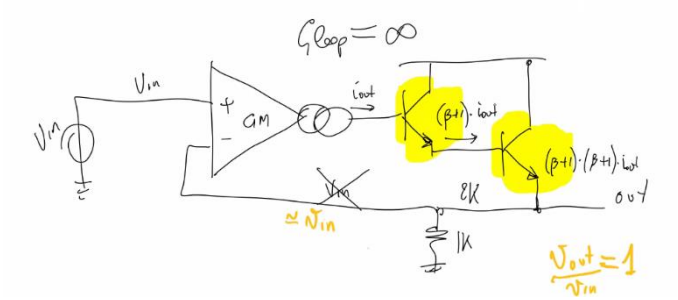


Advantages and disadvantages

The OTA has all the input nodes at low impedance, so even if we have parasitic capacitances, the poles they introduce are at very HF, and this is an advantage. Who sets the pole in the OTA is the capacitor Cl at the output stage.

The disadvantage is that the voltage mode opamp has a A0 very huge, whereas the Gm (tens of mA/V) here has a very low value, so the gain of the OTA is limited, so we cannot use it to perform a feedback in this configuration, because the Gloop of the stage is not infinite and very very limited.
If we want to use it like this, we can but we have to increase the gain of the OTA for example with a Darlington couple to improve the gain of the overall stage as below. i_out gets amplified by the beta + 1 of the two components.

If then we place another resistance we can have a gain we want. But the message is that with just the OTA we cannot achieve a feedback loop.



However the big issue is the **strong nonlinearity**. The I_control that we pump is constant, and IC- + IC+ must be equal to I_control. Then we know the transistor has 0.7V across its Vbe, so Vin- - Vin+ cannot be high as 1V, because it is impossible to have 0.7V on one Vbe and on the other. In fact, the v_diff should be 0.7V – 0.7V, so it should be limited to few hundreds of mV at the most.

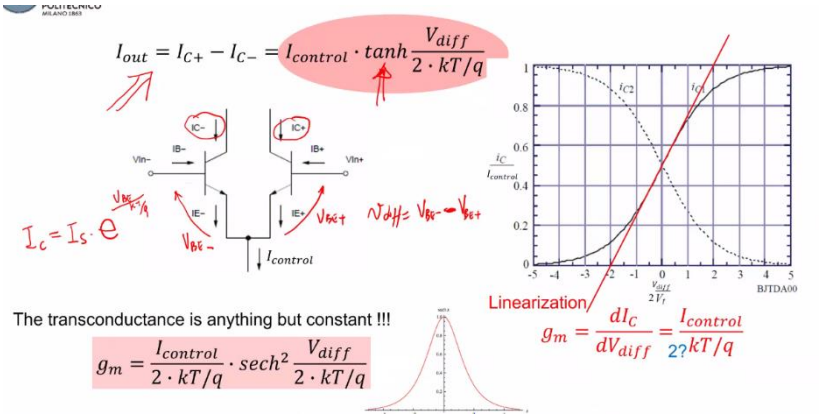
Real gain

We have to study the circuit using the correct equation for the BJT transistor, the exponential one for Ic. We wrote the value of v_diff and we compute I+ and I-. We end up with a I_out that doesn't depend linearly on v_diff. There is an hyperbolic tangent that links the v_diff with the I_out.
So the real input output characteristic of a differential pair has an hyperbolic tangent behavoiur, meaning that if v_diff = 2*Vth, so 50mV, then one input transistor is carrying more current than the other, but

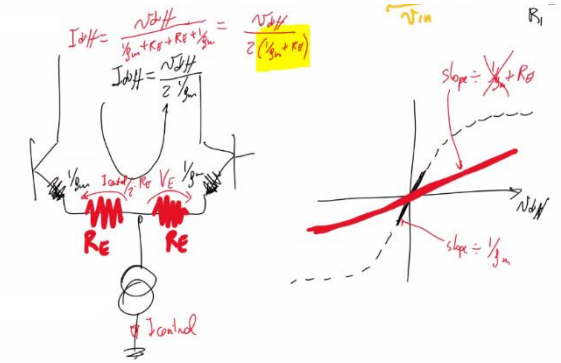
since their sum must be equal to I_control it means that, compared to I_control/2, one transistor is carrying more current and the other one less.

Eventually, in order to operate as a differential pair the differential signal must be in the range of few tens of mV, because the trend is linear only in this case. If we apply e.g. 1V, we are outside the linear range and one transistor carries the full I_control and the other is completely off. Hence we have a saturation of the differential pair when the full I_control flows in one transistor or the other.

If we compute the slope of the curve, it is maximal for small v_diff. dIc/dv_diff = gm that is the slope of the curve and the transconductance of the stage. So we can use this differential pair in the range of few thermal voltages.



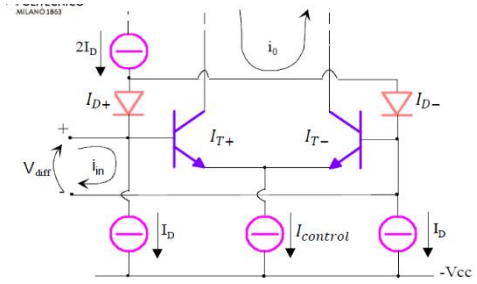
To increase the linearity region, we can start from the differential pair (with the small signal 1/gm for each input transistor, which is very steep when v_diff = 0 but then saturates) is to add a degeneration resistance Re. Given then I_control, we have voltages drop on the Re (Ve) and the differential signal is now given by the following.



With Re the gain drastically reduces (slope is less pendent) but we have a more extended linear region. So we can apply even higher differential voltages in input, but soon or later we will reach anyhow the saturation. So we have a trade-off between dynamics and linearity.

IMPROVED DYNAMICS AND LINEARITY

A much better approach is by adding two diodes and forcing a current through the diodes, and then we remove the current (Id). If so, the two current compensate. But we introduce a close loop.



$$I_{in} = I_D - I_{D+} = -(I_D - I_{D-})$$

Now there is an input current !!!

$$I_{in} = I_{D-} - I_{D+} = I_+ - I_-$$

But the out/in relationship is linear!!!

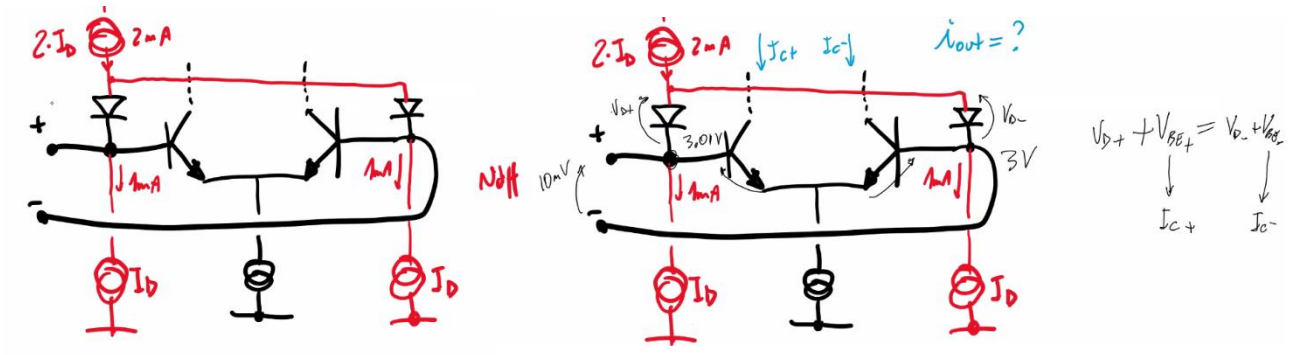
$$I_{out} = I_{C+} - I_{C-} = \frac{I_{control}}{I_d} \cdot I_{in} \quad \text{where } |I_{in}| < I_d$$

Translinear principle: $I_{D+} \cdot I_{T+} = I_{D-} \cdot I_{T-}$

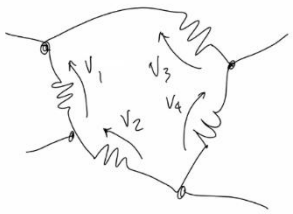
Node currents: $I_{D+} = I_D - I_{in} \qquad I_{D-} = I_D + I_{in}$

Output signal current: $I_{out} = I_{T+} - I_{control} = I_{control} - I_{T-}$

We take the differential pair, we put diodes and we use current generators to remove the currents. Then the two diodes are connected together and we pump 2*Id and we remove them separately. If v_diff = 0, the stage is fully symmetric but if we apply some v_diff, one input will be higher than the other. But at the same time also the voltage across the two diodes will be different. We can write a KVL. Vbe+ causes Ic+ and Vbe- causes Ic-.



Every time we have a circuit that has a diode, pn junction, diode, pn junction and so on and then is closed, every we have a loop where there are just pn junctions connected in series, we can apply the translinear principle. It applies every time we have a loop of transistors or diodes.
According to Kirchhoff, the sum of the voltages in a resistive network should be equal to 0.



$$\Sigma V_i = 0$$

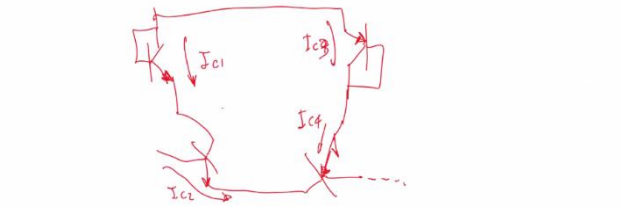
$$V_1 + V_2 - V_3 - V_4 = 0$$

$$V_1 + V_2 = V_3 + V_4$$

If now we consider diodes or transistors, given the V across the transistor, which is the current that flows? It is more or less the I_collector.

Now we can find out the value of V. Hence the voltage across a transistor is proportional to I_collector (Ic). So the voltage across a pn junction is given by a constant coefficient (K = kT/q) multiplied by the logarithmic.
Now, let's apply KVL but with this expression for voltages.

$$k \cdot \ln\left(\frac{I_{c1}}{I_{s1}}\right) + k \ln\left(\frac{I_{c2}}{I_{s2}}\right) = k \ln\left(\frac{I_{c3}}{I_{s3}}\right) + k \ln\left(\frac{I_{c4}}{I_{s4}}\right)$$



If all transistors are equal and belong to the same Si chip and are equal in terms of area and processing, all Is1, Is2, Is3, Is4 are equal. Then, if the transistors are at the same temperature, also the K coefficient is the same. So we can write the following.

$$\ln I_{c1} + \ln I_{c2} = \ln I_{c3} + \ln I_{c4}$$

$$\ln\left(I_{c1} \cdot I_{c2}\right) = \ln\left(I_{c3} \cdot I_{c4}\right)$$

$$\prod_i I_{cc} = \prod_i I_c$$

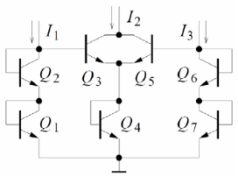
Hence the product of all current in the clockwise direction must be equal to the one in the counterclockwise direction (it applies just for only BJTs or diodes, not mixed).

Given a loop of closed-connected branches, Kirchoff's law says: $0 = \sum_{j \in \{CW\}} V_{cj} - \sum_{l \in \{CCW\}} V_{cl}$

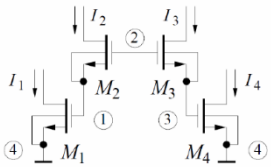
$$V_c = K \ln\left(\frac{I}{I_d}\right)$$

$$\prod_{j \in \{CW\}} I_j = \prod_{l \in \{CCW\}} I_l$$

For loops with just BJTs, K=kT/q

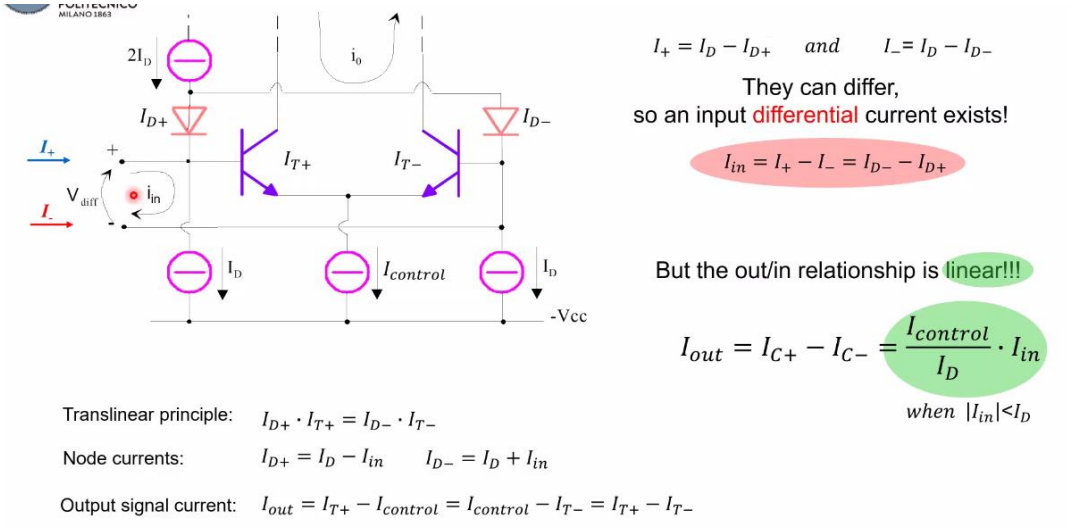


For loops with just MOSFETs, $K=nV_T$



Every time we find a loop where we have just base-emitter, base-emitter junctions, if we know the currents flowing in each transistor, then the Kirchhoff law, that is the sum of all the Vbe that must be equal to 0 results in the equation with the products of the currents.

Coming back to the circuit, we have introduced diodes and current generators. We have Id+ and then we remove Id. If Id+ is different from Id, the difference must flow out through the + pin, and the same for the minus terminal (I'm assuming the base currents to be negligible). So we have an input differential current i_in.

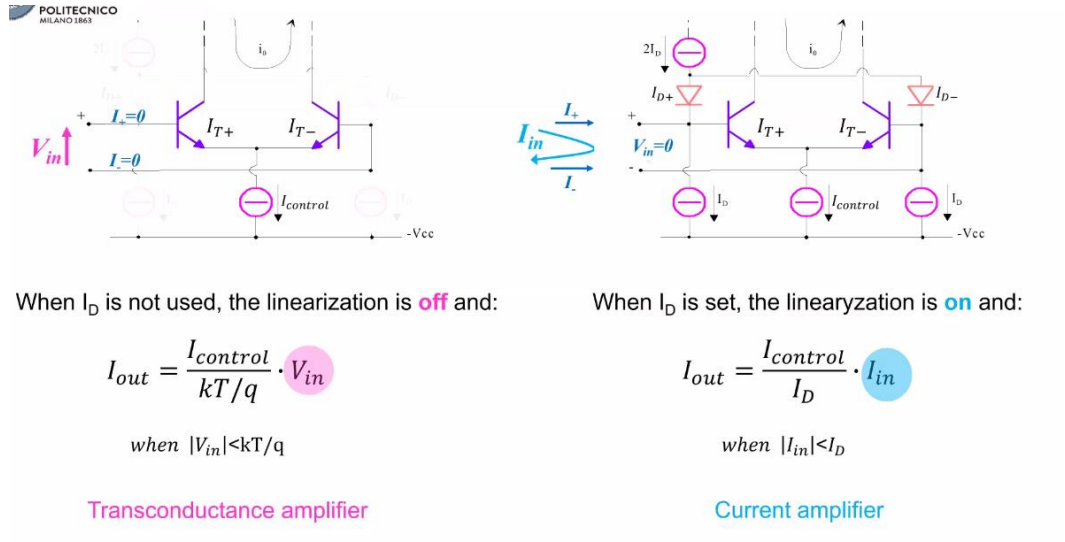


Let's now use the translinear principle (bottom equations). I'm interested in i_out, that is It+ - It-.
We end up with an equation that tells us that the current output is proportional to the input differential current multiplied by a gain that is I_control/Id. With a minor v_diff, Id+ and Id- differ but we remove the same Id, so i_out is created as a consequence of i_in (with no input signal there is complete balance and i_in = 0). The gain of the stage is now I_control/Id.

NB: i_in must be smaller than Id, because otherwise it means that one diode is completely off. But I want the translinear principle to operate, so I want some current Id+ and some current Id- and some It+ and It-.

We started from a circuit with infinite input impedance in input on + and – terminals, so we can apply a Vdiff but input current is 0, but we ended up with a configuration with a low Vdiff but a i_in.

We ended up with a configuration with a differential input current, so it seems that we have a brand new configuration.
In the left mode we have a differential pair and I_control and that's all, input current is 0 at + and – terminal. We just can apply Vin and we get a certain Iout. However, this solution is strongly nonlinear, so the OTA can be used only with a Vin within few Vth.

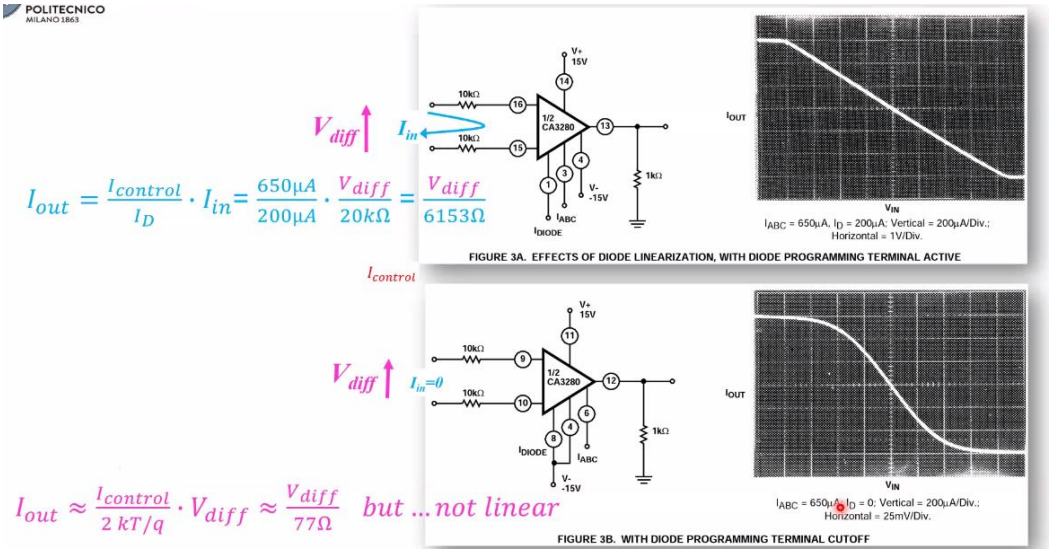


If we use the right new solution, now Vin (of the amplifier, not v_diff) is almost equal to 0 (not equal to 0 otherwise the input transistors won't be on) and there exist a Iin differential input current. Now the

output current is proportional to the input current through a gain set by I_control/Id, and this holds just if i_in < Id, because I want the input transistors to be on.

So the OTA can be use in the standard left mode or as a current amplifier with a linearization network, if we can pump current into the stage.

OTA DATASHEET



With the standard mode we have a high gain but a strong distortion and saturation (bottom). In the top improved situation, we have a smaller gain and thanks to Idiode, 16 and 17 pins become a shortcircuit almost, because Vin at the input of the OTA is almost 0 due to the translinear principle (v_diff is not 0, we have the 10 kOhm resistances to separate the two). So the classical OTA can be used with small v_diff.

In the brand new configuration there may be saturation if Vin gets so high that Iin = Id and one of the two transistors or diodes are off, so the stage is no more operating. However, the linear range is still larger than in the standard case.

Moreover, the transconductance changes with I_control, since it is I_control/25mV. So the higher I_control the higher the Gm.
The frequency of the OTA is very high, so Gm remains constant for frequencies up to hundreds of MHz.

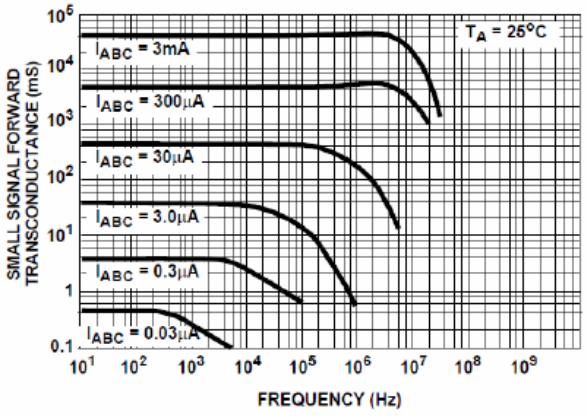


FIGURE 11. AMPLIFIER GAIN vs FREQUENCY

Why if we increase the I_control there is also an increase in the pole?

It's a secondary effect of the differential stage. Inside the OTA we have some parasitic capacitances, the one of the base-emitter junction of the transistors, which changes with the transconductance gm. gm changes with current and so if we increase the current, then 1/gm increases and also the parasitism decrease and the pole improves and gets better.

Function generator

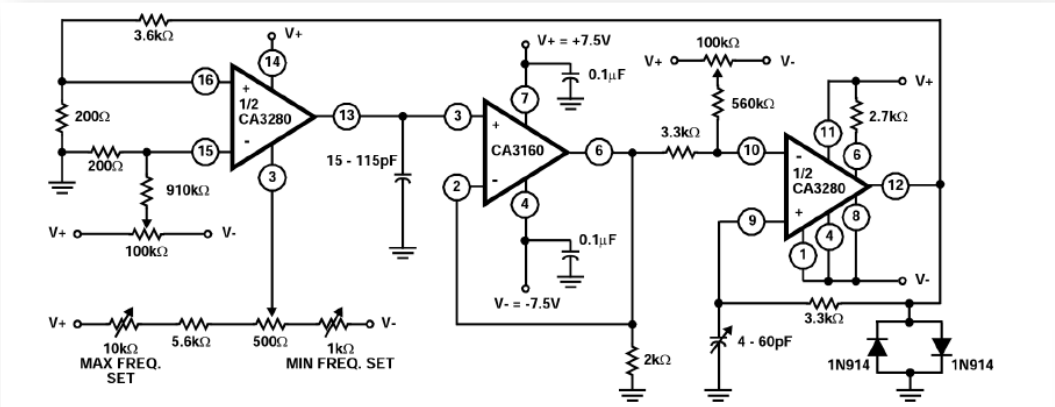


FIGURE 9. CA3280 USED IN CONJUNCTION WITH A CA3160 TO PROVIDE A FUNCTION GENERATOR WITH A TUNABLE RANGE OF 2Hz TO 1MHz

We have a CFA at the beginning whose output current depends on the differential input signal (16 and 15) multiplied by the Gm of the CFA. But the Gm depends on the I_control current. So in output of the CFA I have a constant current proportional to the potentiometers positions. The current flows in the capacitor, the voltage on the capacitor increases linearly and then we go to another amplifier and a positive feedback opamp, that a Schmitt trigger. So we have a constant current in output of the CFA flowing into a capacitor, the voltage gets amplified, then I reach a Schmitt trigger and when I cross the threshold the trigger commutes. Once the commutation is done, the voltage in input to the CFA reverts, so the current direction in the CFA reverts and so also the ramp on the capacitor goes down and so on.

Triangle wave-to-sine converter

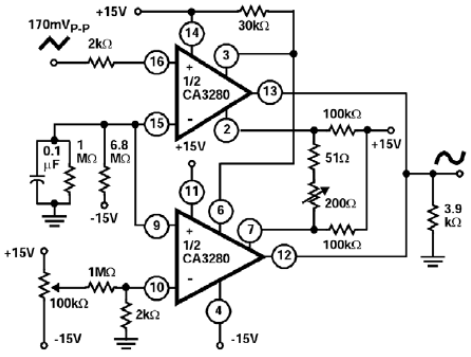


FIGURE 10. TRIANGLE WAVE-TO-SINE WAVE CONVERTER

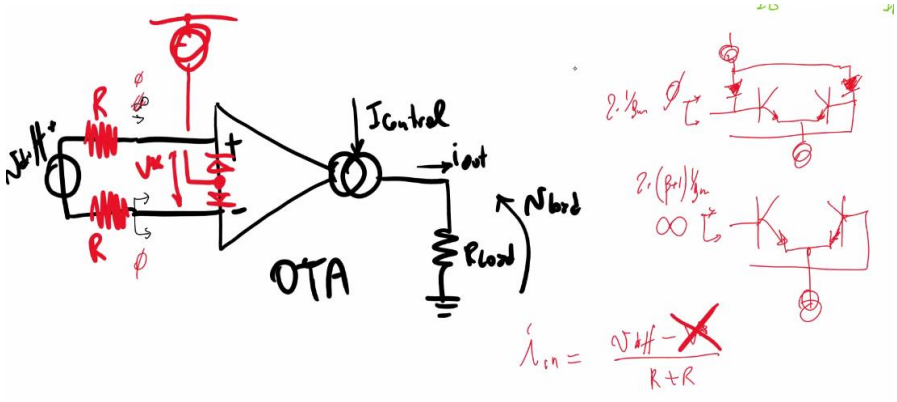
We enter with a triangular signal and the output is a sinusoid. So this stage performs a sort of non-linear input-output characteristic.

Moreover, sometimes the OTA is represented with the input diodes displayed, to highlight the input diodes used for linearization.

Advantages of the OTA

The OTA is important because it can be used as a voltage reader (infinite input impedance) and the output is a current or we can exploit the linearization current, pump it through a pin in diodes and now the input

impedance is almost zero because the Vin at the input of the amplifier is 0. With the diodes that are permanently on I see the impedance 1/gm of the two diodes, no more (beta+1)*(1/gm). Hence to use this circuit I must introduce two resistances otherwise I cannot apply a v_diff.

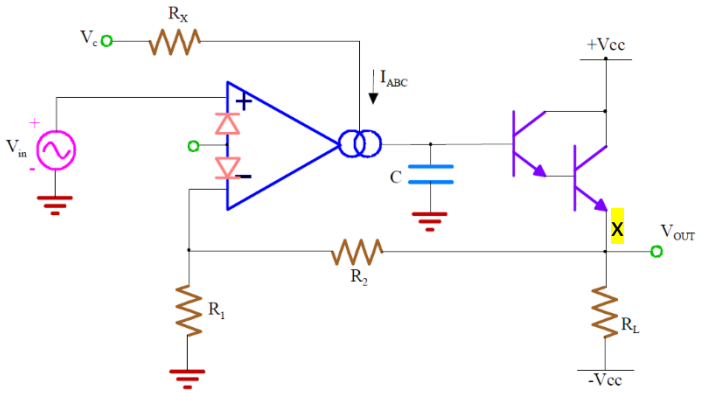


The advantage of an OTA with a linearization network that is actually behaving as a Norton amplifier (we will see this) is that the OTA has I_control, so we can change the gain of the OTA by changing it, in both the standard case or in the linear one. Conversely, in the Norton is fixed.

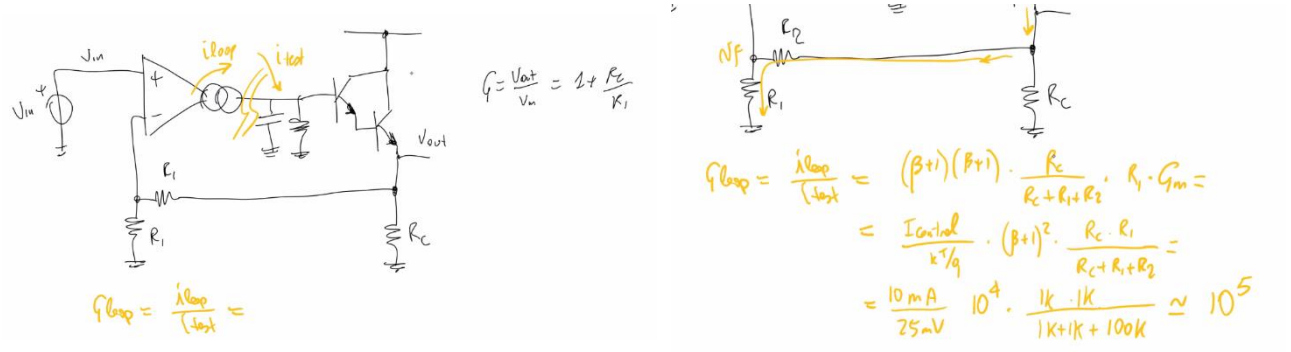
Example – Voltage controlled LP filter

We have a component at the output to increase the gain. If Gloop is very high and we apply something positive at +, a current is pumped out and it causes a voltage over the Darlington network (which has a very high impedance). The Darlington is a follower, so Vout increases and also the – termina, where we will get Vin. But if I have Vin at – termina, what is Vout? It is not Vin * R1/(R1+R2), no!
The signal propagates in the clockwise direction, it is the voltage at x that gives the votlage at the – terminal, so Vin = Vout* R1/(R1+R2).

i.e. $f_{pole}(V_c)$



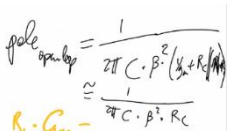
If we change I_control the gain is always 1 + R1/R2. so what is I_control doing? Let's compute the Gloop by cutting the circuit.



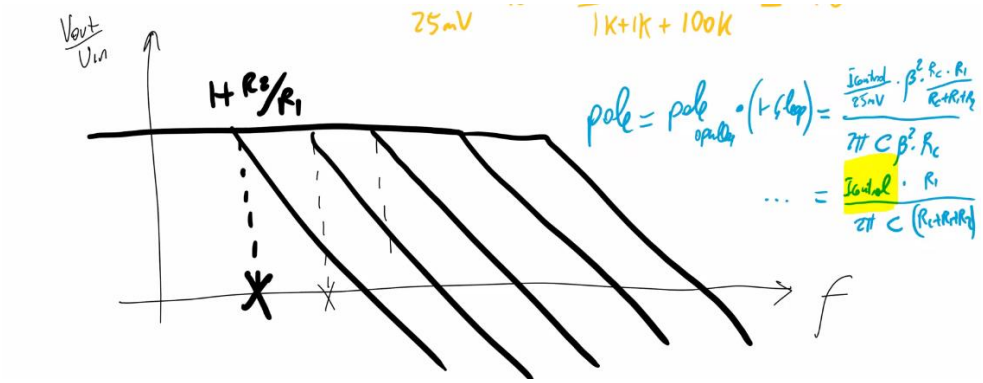
If we pump i_test, Rol (in parallel with the output capacitor) is so high that the transistor is pumped in the Darlington couple. Then we have the current that flows in R2 and R1. Then I get v_f, but then since I'm studying Gloop I multiply by Gm (open loop gain of the OTA). Since Gloop is very high, we can assume the stage to be ideal.

Real gain

Let's now study the real gain of the circuit. The open loop pole is due to C multiplied by the total equivalent resistance.



We can see that the pole is constant. Let's now write the Bode plot. The gain is constant and given by 1 + R2/R1, and we have a pole in the circuit (no zeros) that is the open loop pole multiplied by 1 – Gloop. We can see that the pole varies with I_control.



In this configuration the gain is fixed but we can change the pole of the circuit.

ISOLATION AMPLIFIER – ISO

We want the power supply of the output stage to be separated wrt the one of the input stage, so that the we don't have damages on both sides.

Requirements:

isolated input and output stages
double power supply
galvanic isolation

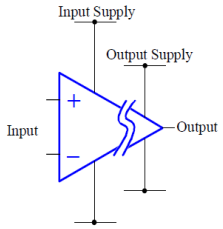
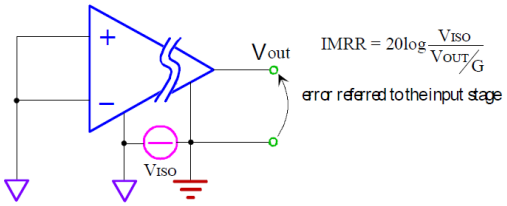
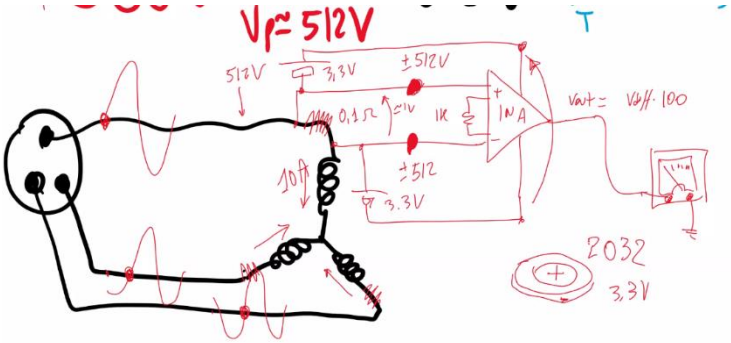


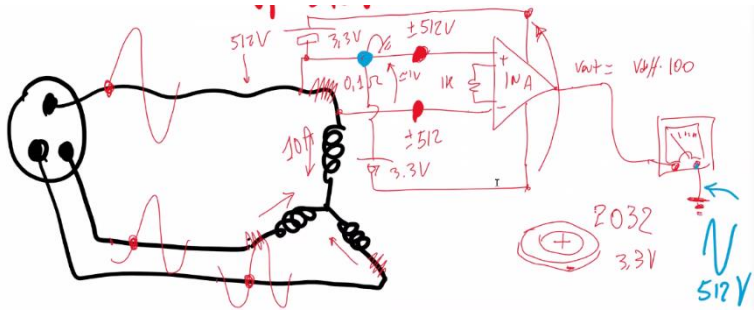
Figure of Merit: **Isolation Mode Rejection Ratio**



A possible application is the following, where we don't have defined a ground.



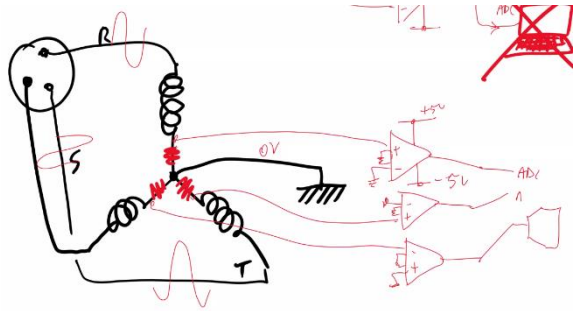
So we need to define a ground, that is not 0V but it moves up and down by 500V. So it is not easy to connect a low voltage opamp with a high voltage source.



To connect the electronics to a tri-phase power supply we put a resistor ad the center of the tri-phase, not on the outside. Thanks to this the center is at 0, and to be sure I can connect it to earth (literally ground). Now we can attach the amplifiers to the small power supply and the electronics work.

But if the ground disconnects, we broke everything and we kill all, destroying the opamps, having hence shock hazards.

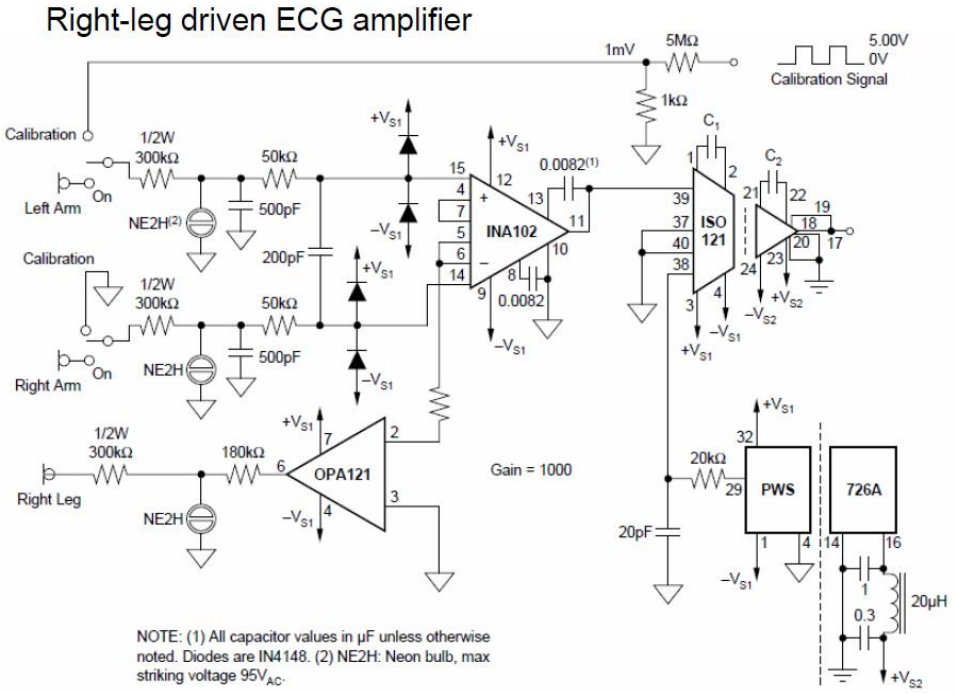
So I want an amplifier that is capable to introduce galvanic isolation, because I don't want the shock hazard to propagate.



Galvanic isolation means that we can bias the two stages independently inside the same amplifier. We can define the Isolation Mode Rejection Ratio (IMRR). An ideal ISO has output 0 if input is 0 whatever is the difference between the two power supply. The better the isolation, the higher the IMRR.

EXAMPLE OF APPLICATION

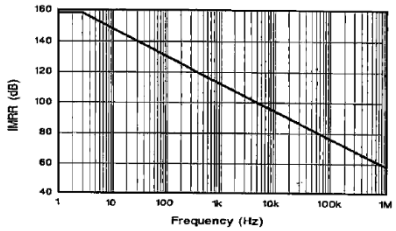
Not just to measure high current and voltage, but also low voltage if we want for instance to measure the ECG of a patient and we want to prevent any short in the equipment to kill the patient.



Example of IMRR response

The IMRR can be very high, over 60dB, but if we increase the frequency at which the two grounds move, the isolation mode becomes less effective.

IMRR frequency dependance

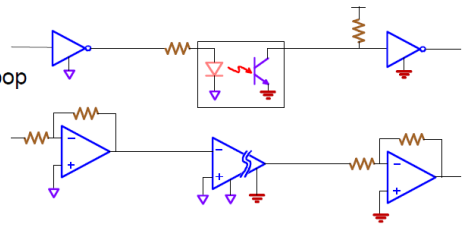


Transient Immunity:	ISO122	TI < 1000V/ms
Isolation techniques:	Optical Magnetic Capacitive	

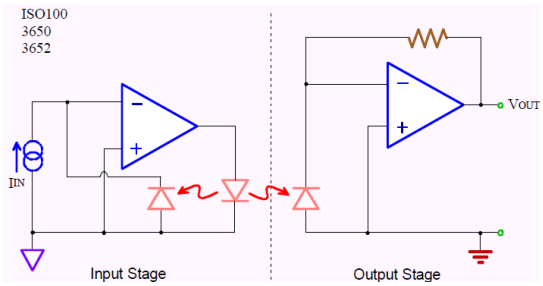
OPTICAL COUPLING FOR THE ISOLATION

Based between an emitter and a receiver. We can use a LED and a transistor; if we have current through the LED we will also have it through the transistor, so it is for instance for digital circuits (upper circuit, optocoupler). Typically LEDs require 1V across them to be on. Thanks to the pull up resistance, if there is a current flowing there is a voltage drop across the resistance of the second stage and the output is high.

Advantages:
"light" breaks the "electrical" loop
high EMI
trivial for digital signals
problematic for analog signals



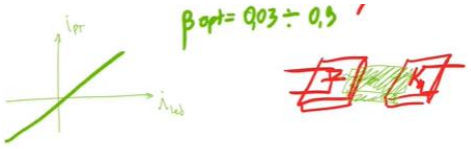
Example of optical coupling for analog signals:



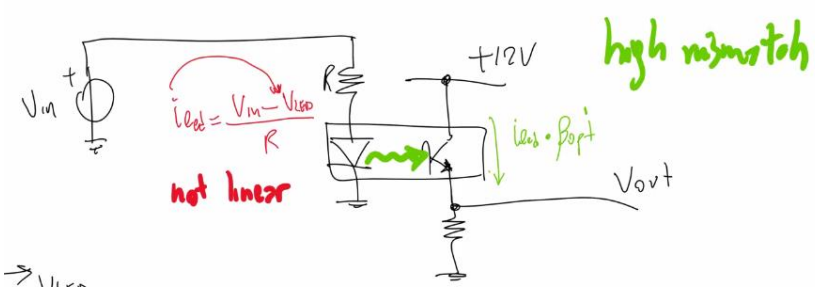
For an analog information we need something smarter, because we have a certain signal and we want an analog output.

Optocoupler for analog signals

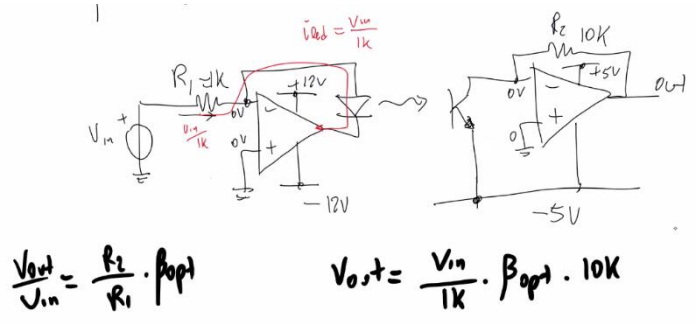
LED and phototransistor are not made on the same substrate because if something happens on one side we destroy both sides of the chip, so the real optocouplers are made with two different chips.
Then somehow they are coupled in someway optically. Usually, given i_led and i_ph, the relationship between the two can be linear. So if we pump a given current in the LED, the current in the PD will be of a certain value depending on the coupling. If the coupling is not perfect, we run the risk that there are some mismatches, because if we pump a given input current, we don't know the value of the output current.



So we cannot use an optocoupler applying a Vin and using a resistor and an optocoupler to transfer an analog signal, because it doesn't work. In fact, the input current i_led is as below, but the relationship between i_led and v_led is not constant, but exponential. So the light we create generates a current in the phototransistor depending on beta_optical, which is eventually varies too much.
So in the end v_out = i_led*beta_opt*Rout is not linearly related with v_in.



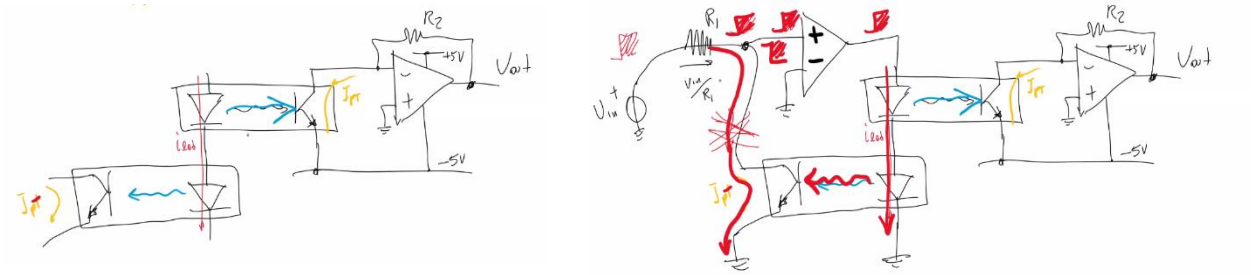
A drastic improvement is the following. Now i_led is v_in/1k. For the second stage we use another bias voltage but instead of using just a buffer as in the circuit above, let's use a transimpedance amplifier and the transistor is biased at high reverse voltage. Now Vout = Vin/1k * beta_optical * 10k.



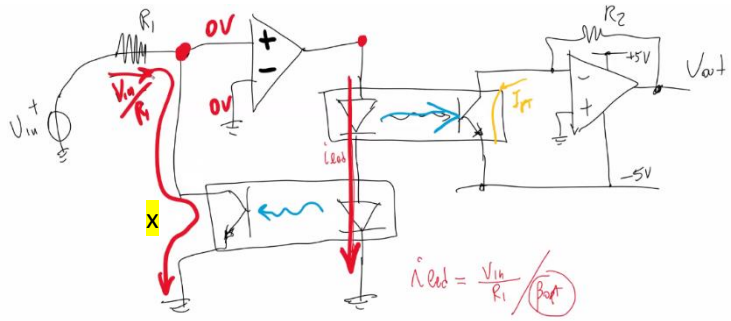
$$\frac{V_{out}}{V_{in}} = \frac{R_2}{R_1} \cdot \beta_{opt} \qquad V_{out} = \frac{V_{in}}{1K} \cdot \beta_{opt} \cdot 10K$$

The drawback of this configuration is the beta_optical, which is a factor that changes a lot. It is not fixed, but it has a very high variability.
So if we have a component that is really bad because it has a high tolerance, instead of putting something that has a low tolerance, let's use this component (the optocoupler) twice to compensate the error introduced by one.

Since the beta varies, one optocoupler will be used to move the signal in one direction, and the other to check what the cross-signal is. In this way Ipt will be the same on the two branches. This current Ipt can be used to properly drive the LED.

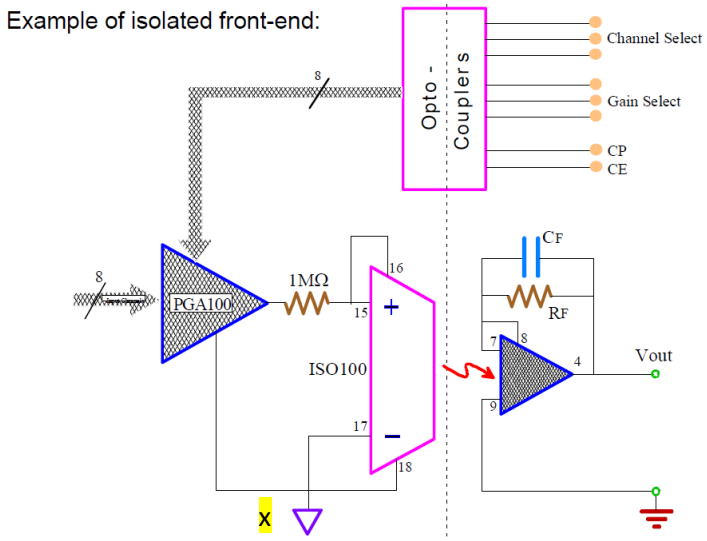


We must use the input signal Vin to drive the LED, but then I cross-check with the phototransistor to regulate the i_led. We can hence let the current Vin/R1 to flow through the transistor of the second optocoupler (but we cannot do this because we have the collector of a transistor), so I use an opamp to drive the LED, and the opamp is happy if the voltage it sees is 0 at the + terminal. So we have the negative feedback loop, because the current in the optotransistor increases if the light from the LED increases. If the loop is strong, we have the following. The current flows through the transistor because we are pumping it.



Since beta optical is very much variable, also i_led will be but I don't care because what I care is the current Ipt, which is a twin brother of the current x. But **we must assume that both optocouplers are identical**, they cannot mismatch, but they can be both 'bad'. Thus we can conclude that Vout = Vin*(R2/R1).

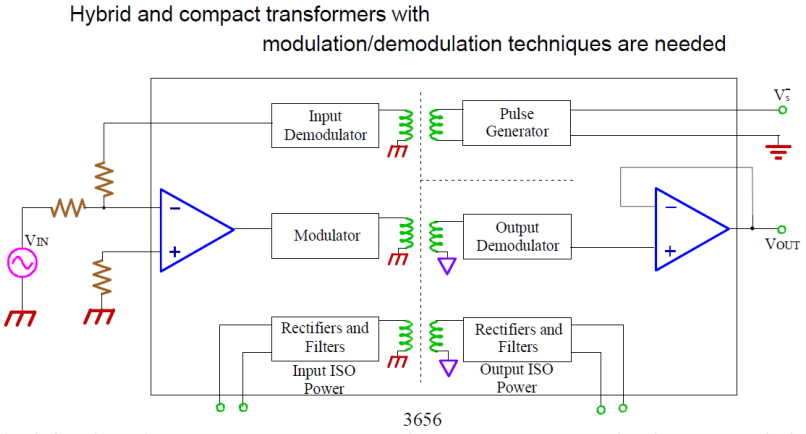
Example of an optocoupler



One pin is to ground (x), in another pin we enter with a resistor (1 MOhm) and it is used for the feedback and the output need an external Rf and capacitor Cf to introduce the low pass filtering action.
Then we can also use individual optocouplers (top of the image) to apply the signal from the microcontroller back to the PGA. Thus the isolation amplifier guarantees the isolation of the analog part and the 8 optocouplers guarantee the isolation in the digital connection between input and output.

MAGNETIC COUPLING

Another way to isolate input and output is by using transformers, but they are not very much used. But if we apply a constant DC value in input, the output current will be 0 because we are applying a constant value to a transformer. In order to transfer a DC signal and have an output DC signal we have to modulate that value and demodulate it.



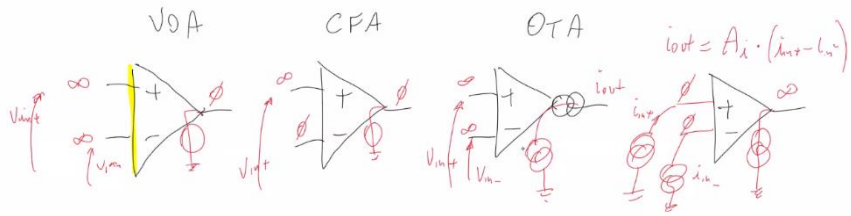
To introduce a global feedback we cannot use a resistor to connect the input and the output, we need to use another modulator and demodulator.

CAPACITIVE COUPLING

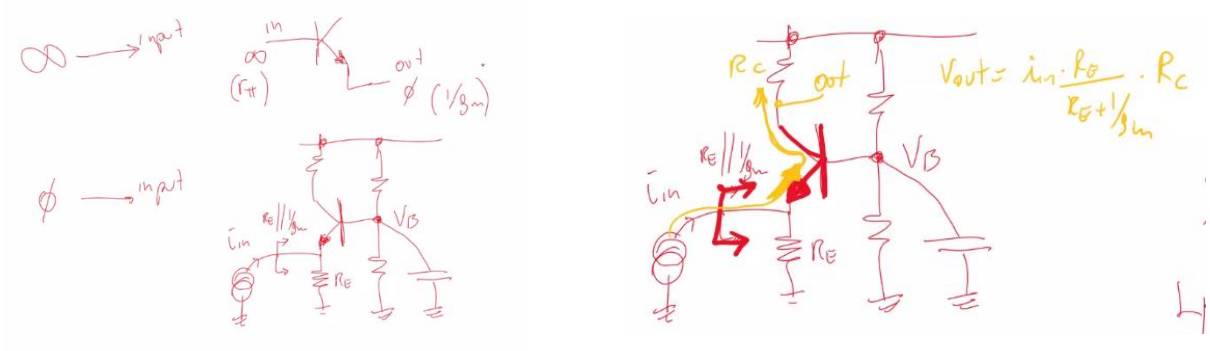
Another possibility is to use them. We can use and ADC modulator, we enter with a voltage, convert it into a digital information, we charge the capacitor and on the other side we use a DAC that reads the information on the capacitor and converts back to an analog information. Nevertheless, the optical one is the better.

NORTON AMPLIFIER

Amplifier that reads current and provides a current. So low impedance in input and infinite impedance in output.

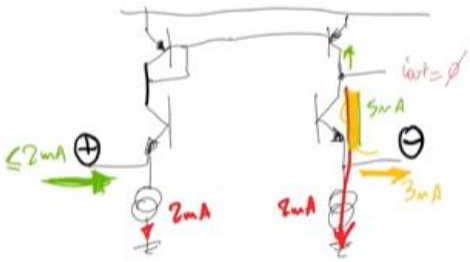


To have an infinite input impedance we have used a transistor (BJT, or MOSFET, even better). If we want to have a low input impedance we need to enter into the emitter of a transistor, after the transistor is properly biased. The impedance we see is Re||1/gm. The advantage of BJT transistor is that 1/gm is Vth/Ic, usually in the order of ohms, while in the MOSFET transistor 1/gm is 1/(2kVov).



With the Norton amplifier we want an input differential pair, so let's add two inputs. We want the inputs to be very low impedance. The current that is useful is the one that enters into the emitter and exits from the collector, to provide the output signal. If we want that all the current enters in the collector, Re must be maximized → better to use a current generator in place of Re so that I'm sure that I will see 1/gm and the current will all flow in the transistor.

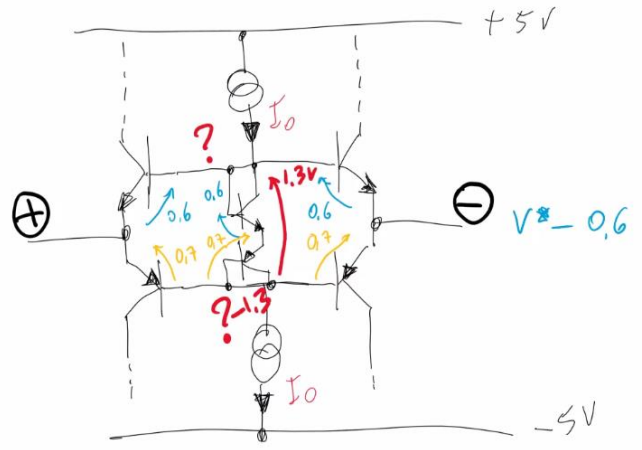
To then perform the difference between the two currents I need to introduce a current mirror. If the two currents are equal, then the output current is 0 due to the mirror.
So if we read no current from the output pin, the bias currents flow in the transistor. We can also have currents moving out from the transistor but I cannot pump to much current in this configuration because it is at the most equal to 2mA, the current of the current generator, otherwise I push the current in the current generator and not in the transistor and the transistor is off. So this configuration is good if I want to sink current but not if I want to pump it.



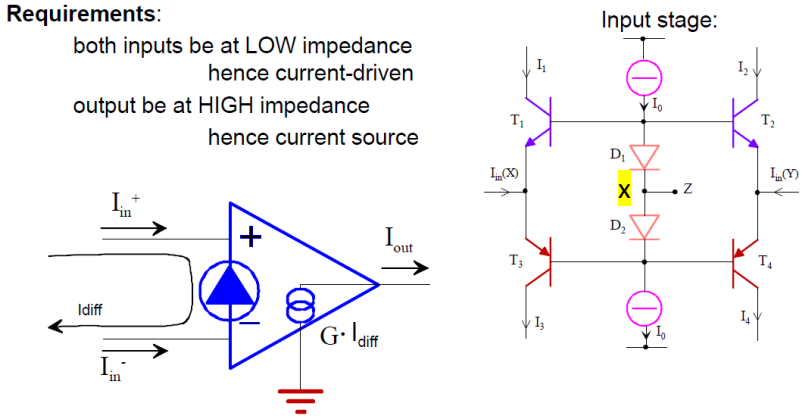
So I need to add other two BJTs. Of course I need to bias the two branches. A possible solution is to apply the same bias on the two bases of the two couples of BJTs.

But it is better to use, instead of resistances, diodes, which will be the exact copy of the pnp and npn transistors, so I connect them in transdiode configuration. I need to keep those transistors on so I pump current and also sink it. Hence I use two current generators I0.
If I pump I0, the two diodes force the voltages across them to be 1.3V, but we don't know the voltages at the sides.

To fix the voltage, I can use the intermediate point, put it outside the circuit and depending on what I attach to that pin I will have the upper and lower voltages.



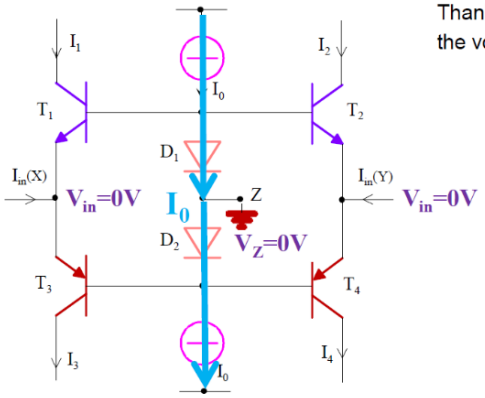
So the Norton amplifier has a Z pin that determines the voltage I want to apply to the Norton amplifier. Once the voltage applied to the intermediate point is selected, the two terminals will be at a fixed voltage.



We can use the intermediate point x, bring it outside the circuit and, depending on what we connect at that point, we are fixing the voltage at the input. If for instance we connect it at ground, the two inputs will be at 0V. Hence the Norton amplifier has another input pin that is used to set the voltage we want to apply on the two inputs of the opamp. Because of this we cannot apply a voltage on the terminals of the Norton, it is set by the amplifier, we apply currents.

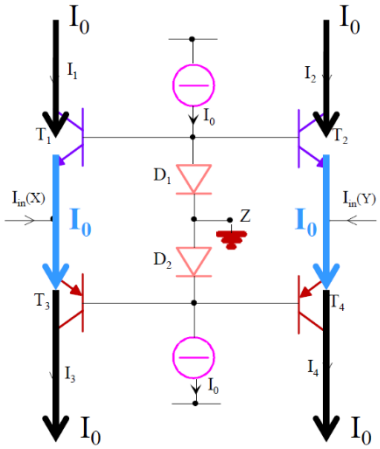
Then I want the output to be a current proportional to the current difference at the input and so I need current mirrors to subtract signals in input in the proper way.

So we are forcing current in the middle branch, so the two diodes are on (light blue path). We set node Z to 0V. If the purple npn transistors are matched, and if the pnp red ones are matched, if the input current is 0, the output current is 0.



Thanks to pin Z
the voltage at the inputs is set

Now I want to compute which is the current that flows through the right branch. We can use the translinear principle. Given a network of transistors, the sum of the Vbe voltages in clockwise and counterclockwise directions must be 0. Then according to the translinear principle this concept extends to the currents, but in terms of product.



Translinear principle:

$$I_1 \cdot I_3 = I_2 \cdot I_4 = I_0^2$$

With identical transistors...

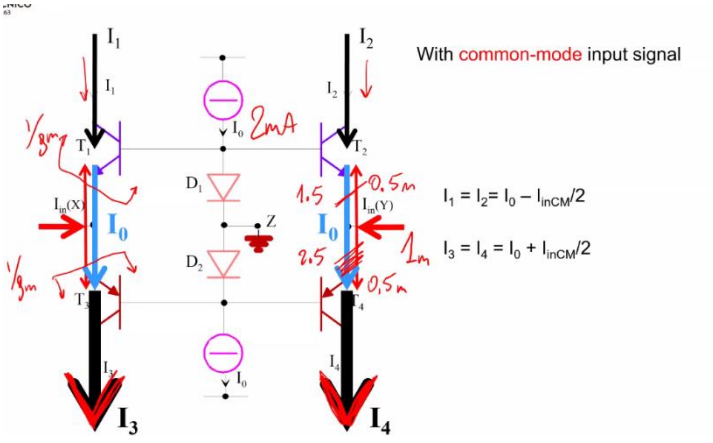
... and with no input signal

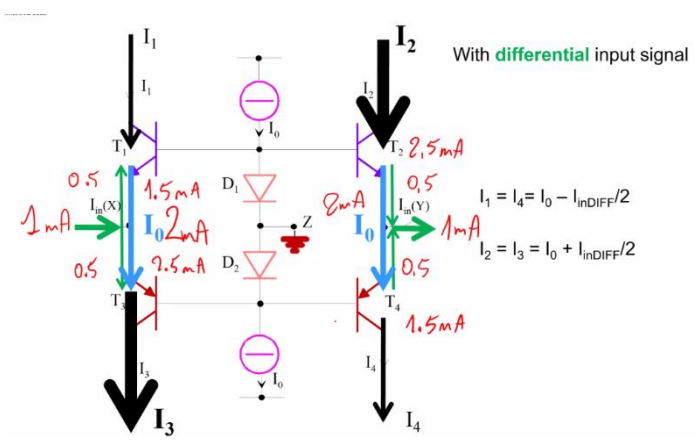
if $I_{in}(X) = I_{in}(Y) = 0$ we get

$I_1=I_2=I_0$ and $I_3=I_4=I_0$

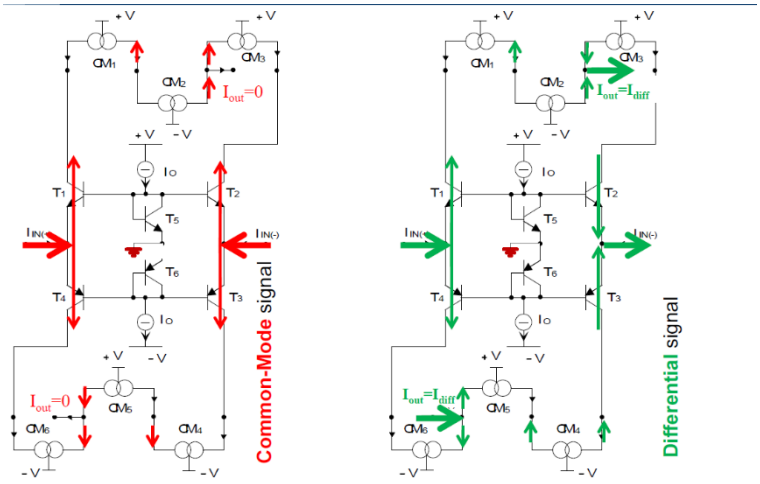
If the pnp transistors and npn transistors have the same area, the I0 are the same.

If then I apply a common mode input current, e.g. 1mA, so the currents are the same on the two input terminals, I'm pumping a current that is adding to the biasing current. I0 is 2mA if the diodes are biased with 2mA, so the current will be 2.5mA below and 1.5mA above to maintain the KCL ad x and y nodes. So the common mode signal splits in two if the two 1/gm of the transistors in input are the same.
So the CM signal increases the two downward currents and decreases the top currents if the signals are inward and coherent. Instead, if we consider a differential signal (right), we have a different situation.



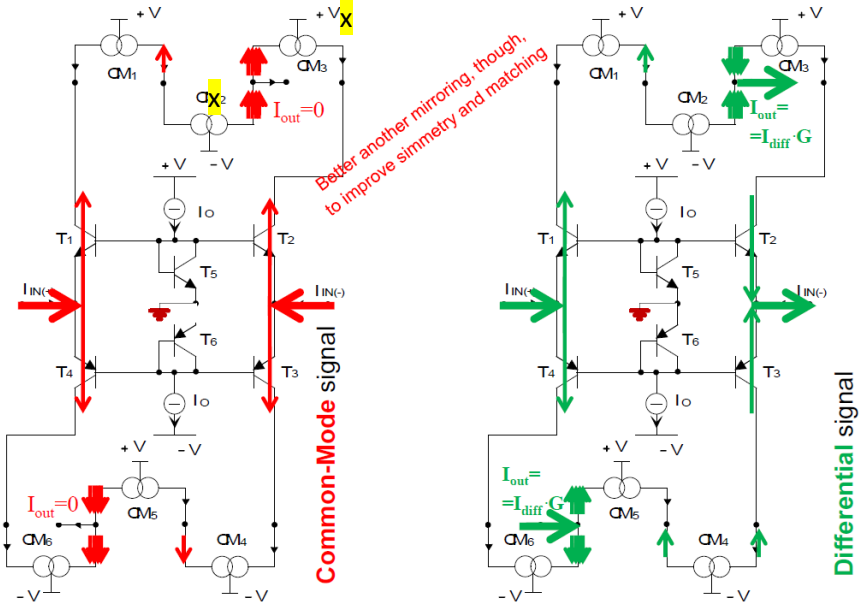


Now we can use current mirrors.

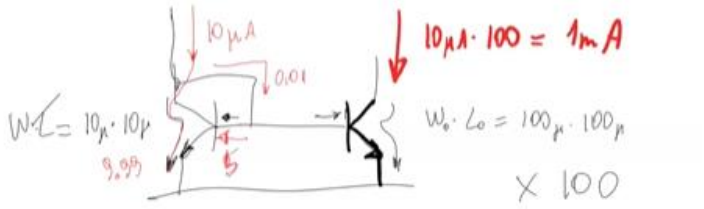


The output is in the branch of the mirror where the transdiode part of the mirror is not present (we have two high-impedances nodes touching together, so two collectors). This applies both got the top and bottom sides.
We notice that in the case of CM input signal the output is 0. In the case of a differential signal the output is actually a current.

So this is the final Norton configuration we are looking for. The two inputs are low impedance and the only high impedance node is the output, which provides a signal that is the differential signal at the input. The output is half because we have also the bottom part. So in a Norton, the highest gain we can get is 1. If we want a higher amplification we can add amplification to current mirrors x. So let's use a current mirror that is not symmetric.

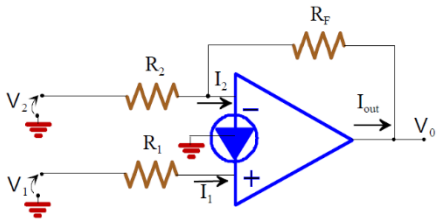


However, in this analysis we are not considering the base current, in fact the 10uA in the left transistor has a portion that goes in the base of the two transistors. But if the area of one of the two transistors becomes larger, to increase the mirroring ration, then also the base current increases, and hence the current that is mirrored is reduced, so the mirroring factor is impaired.



So the Norton amplifier has a Ai amplification but this amplification is limited.

POSSIBLE CONFIGURATION



$$I_1 = \frac{V_1}{R_1}$$

$$I_2 = \frac{V_2}{R_2} + I_0$$

$$I_0 = \frac{V_0}{R_F}$$

$$I_0 = A_i \cdot I_{diff}$$

$$V_O = \frac{A_i}{1 + A_i} \cdot \left(V_1 \cdot \frac{R_F}{R_1} - V_2 \cdot \frac{R_F}{R_2} \right)$$

Note that A_i is NOT infinite but just few tens !

... hence **NO** "IDEAL GAIN" behaviour

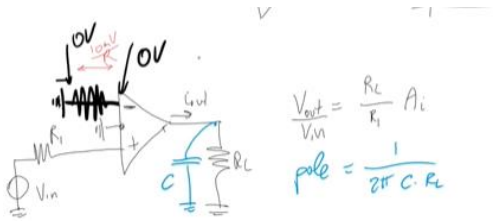
We have the Z pin. If we apply GND there, by design the input terminals of the Norton will be at 0. So if I want to apply a voltage in input, I need to place a resistor to prevent shortcircuits. So the current in R2 is V2/R2 and then it splits in the feedback and in I2, and the same for I1 = V1/R1.
Iout flowing through Rf generates a Vout. Moreover, Iout = Ai*(I1 - I2).

So we have a voltage difference amplifier. If Ai is infinite, much higher than 1, Ai/(Ai+1) is almost 1. So this configuration cannot be considered an ideal negative feedback configuration because the gain is not infinite and so we have a real gain of the configuration.

NB: Rf must be present, we cannot use a shortcircuit, because otherwise Vout will be always equal to 0.

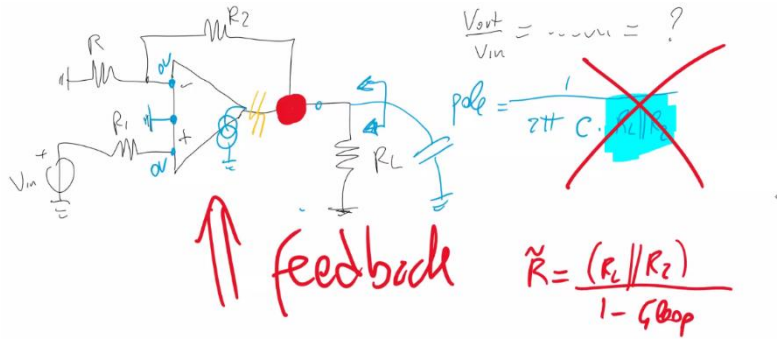
Inside the Norton amplifier, all input nodes are low impedance, so even if there are parasitic capacitances I don't care, because all the internal nodes are 1/gm of the transistors.
If we want to decrease the bandwidth of the amplifier, I can do this externally.

NB: never connect a pin of the Norton amplifier to ground, because it is a voltage source so it wants to stay at the voltage set by the Z, but even if the Z voltage is 0, the voltage on the terminal should be few mV due to mismatches, so if we connect ground to the pin we are creating a short. So we always need to introduce a resistor in the middle. Of course we will have a current across the resistor, but if R is very high, the current will be negligible.
To set the pole, we put a capacitor in parallel with the output load resistance, so we can set it wherever we want.

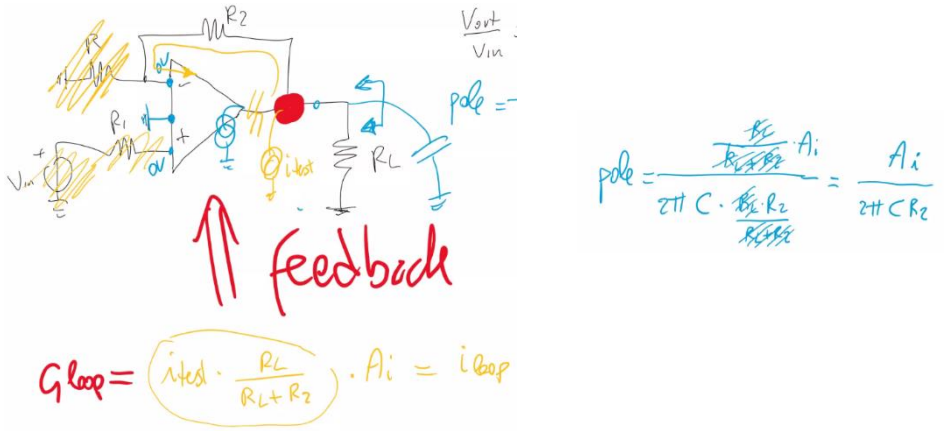


If we use the configuration with R2 in feedback (below), and I put C in parallel to Rl, now the capacitor sees R1||R1, because – terminal of the Norton is grounded. NO! This reasoning is wrong because we have a feedback circuit and every time we have a node of the feedback, and we try to move that node,

we are the feedback actiong. So the impedance we have is the 'stupid one', that is R1||R2, but then divided by 1 – Gloop, so we need to compute Gloop.



To compute Gloop we cut the loop and pump Itest. The current in feedback is Itest*(R1/(R1+R2)), that then enters in the ground, there is no need to perform current splitting. So Gloop = Ai*(R1/(R1+R2)).



So we have changed the position of the pole with this value of equivalent resistance affected by the loop.

ADVANTAGES AND DISADVANTAGE

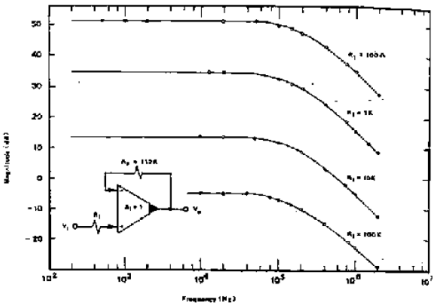
We may change the gain and the pole is not changing, because all the internal nodes are low impedance. So the pole of the Norton amplifier as it is is at very high frequency. If we want to shift it to other frequencies we need to add an external capacitor. In particular, is Cload equal to Cf since they are at the output and attached to ground? They could differ, but we need to plot the Bode diagram if one of the two is connected.

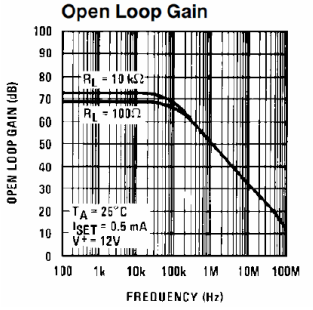
Advantages:

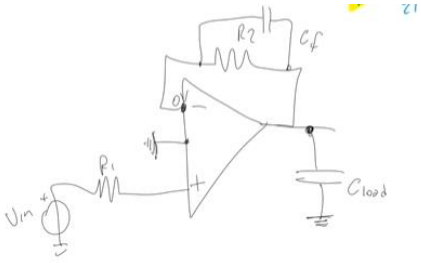
all nodes have low-impedance ($1/g_m$) hence very low parasitic time constants

very wide bandwidth

bandwidth independent of closed-loop gain



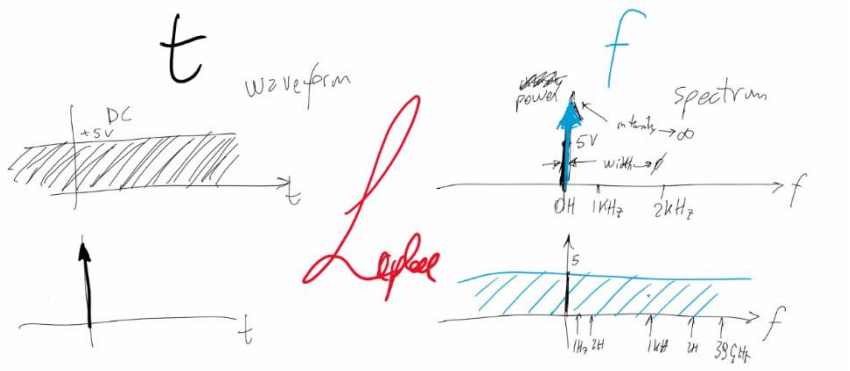




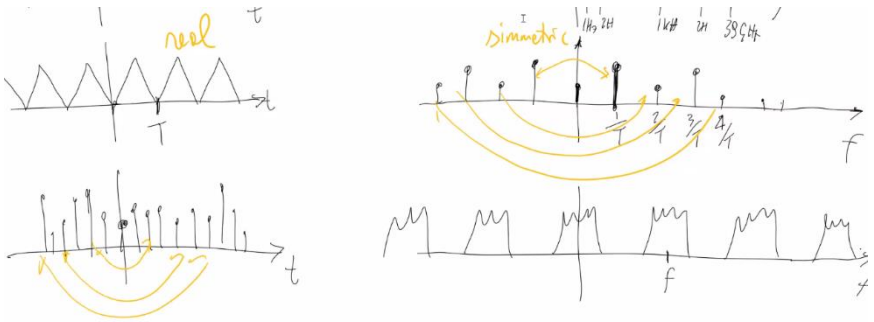
SAMPLING

We have to differentiate the time domain and the frequency domain. If we have a DC signal in the time domain and we plot its spectrum, it will be just at 0Hz, where I put its intensity. Sometimes is better not to draw the intensity but the total power. The power of a constant signal from -inf to + inf is infinite, so I will have a Dirac delta with infinite height and 0 width.

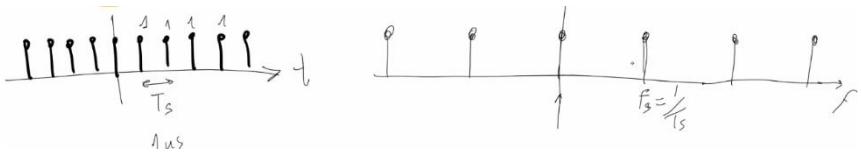
We can then refer to the Laplace transform. A constant spectrum in the frequency domain returns a delta in the time domain. In fact, infinite spectrum it means that the intensity is the same along all the frequencies. If in time domain we sum all the possible frequency, all the sum at a given t collapse at 0, only at the origin t = 0 they are summed without disappearing.



Moreover, a periodic signal in the time domain, thanks to the Laplace transform, in frequency it will be the sum of harmonics at replicas of the period of the signal. Not only exists positive frequencies, but also negative ones. If the signal is real, then the Laplace transform is symmetrical. Of course, it is true the vice versa. If we have something periodic in the frequency domain, then for sure in the time domain it should be a stack of deltas.



If in time domain I have a comb of deltas, a set of Dirac's delta with height 1 and period T (called Ts, sampling time), since it is periodic, the Laplace transform should be a set of delta, and since it is a set of delta, the Laplace transform should be periodic → we still have a comb of deltas.

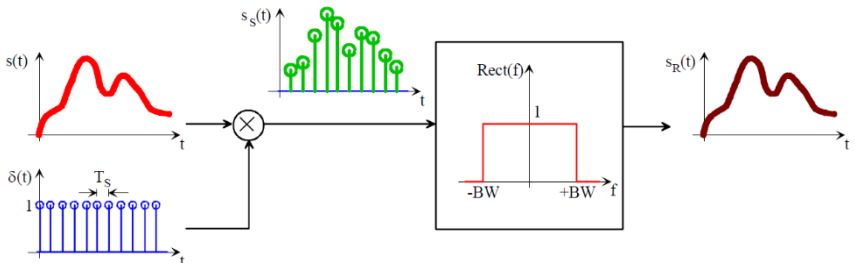


If we push Ts towards 0, it means that the set of deltas becomes very close one to the other, in the frequency domain the deltas are instead widely spread. If Ts = 0, we get a DC signal in the time domain and so in the frequency domain we have a delta at f = 0 and the other deltas are at infinity, so they disappear.

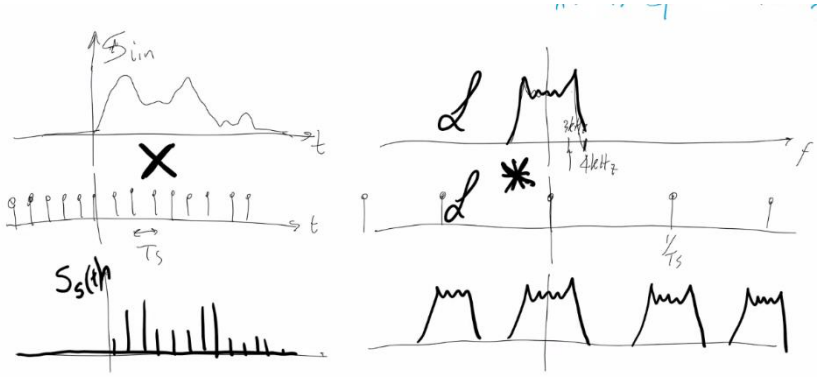
SHANNON THEOREM

Shannon theorem, 1949:

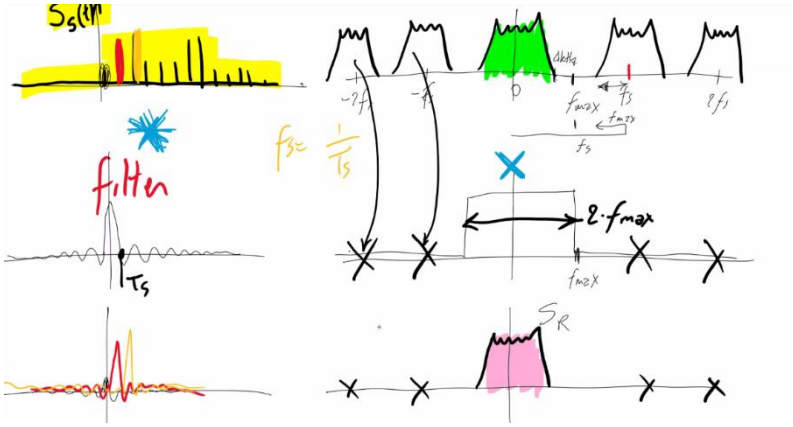
"if a ***function s(t)*** *has no frequency components above BW Hz, then it is fully* ***determined*** *by its* ***values****, sampled every* $T_S = 1/2BW$"



The minimum distance between two samples must be at least 1 divided two times the bandwidth of the signal. To reconstruct the signal we can use a LP filter with a width as large as the bandwidth of the original signal.



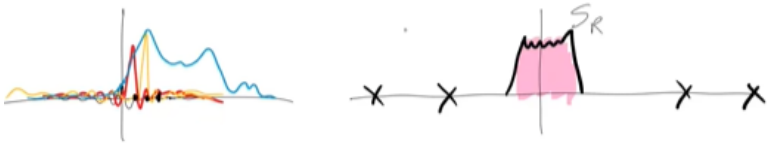
We have a replica of the original spectrum at multiple of fs = 1/Ts. Then to reconstruct the signal I will use a rectangular filter with a fmax slightly higher than the fmax of the original signal, but it has to be lower than fs – fmax. Thus, when we perform the multiplication in the frequency domain (filtering in time domain equals the multiplication in the time domain) we get in the end a convolution of my samples with the Laplace transform of the filter we use. And we know that the Laplace transform of a pulse is a sinh. So if we multiply the signal spectrum and the filter spectrum we get the original spectra with all the other replicas killed.



Eventually the reconstructed spectrum is equal to the original one.

In the time domain we have to convolve the deltas with the sinh of the filter. So we have to draw a sinh for each delta and, given the width 2*fmax, the sinh cancel each other apart from the positions where we have the zero crossings.

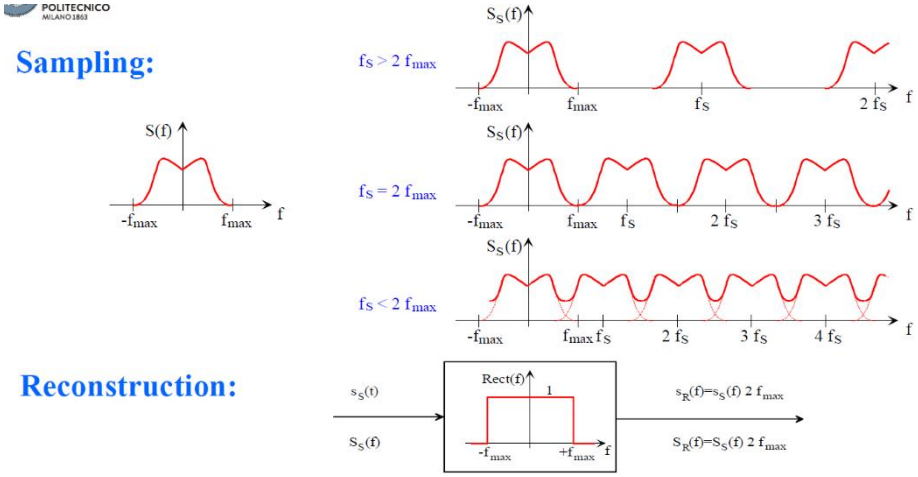
In the end we have an analog signal equal to the original one thanks to the Shannon theorem.



This of course happens only if the replicas of the signal in the frequency domain are far from each other, because in this way the filter can kill the replicas.

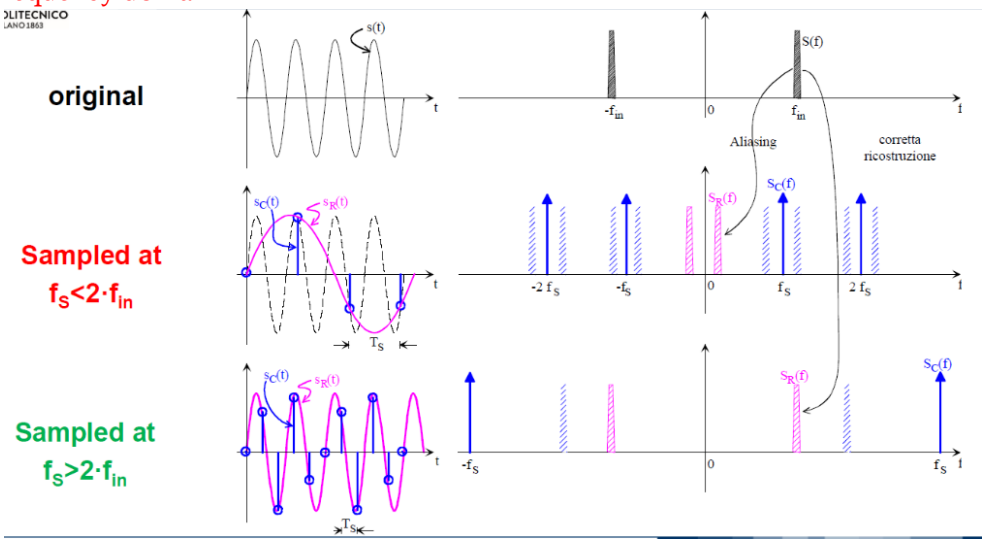
ALIASING

It happens when we don't choose the proper sampling frequency.



In the limit case, if fs = 2*fmax, then the replicas touch. If the spectra overlap, we cannot properly reconstruct the signal.

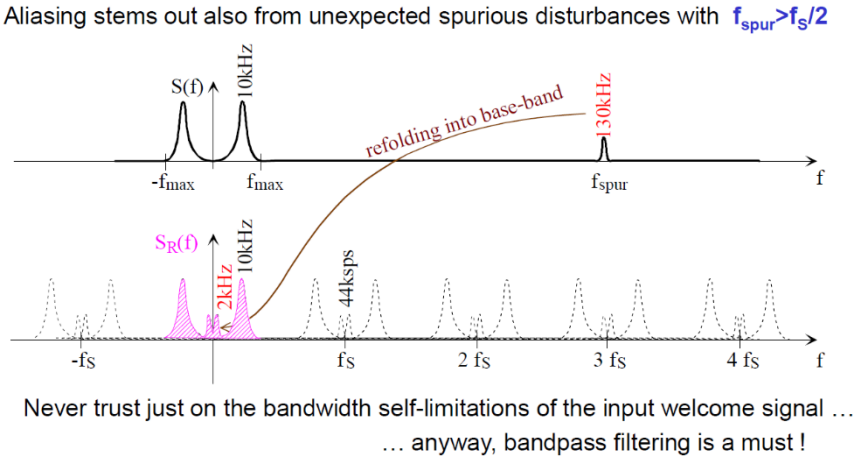
Aliasing in frequency domain



We have our signal, a t_sampling and in the last row we are obeying the Shannon theorem, so in the frequency domain we will have the original S(f) and some headroom between the replicas.

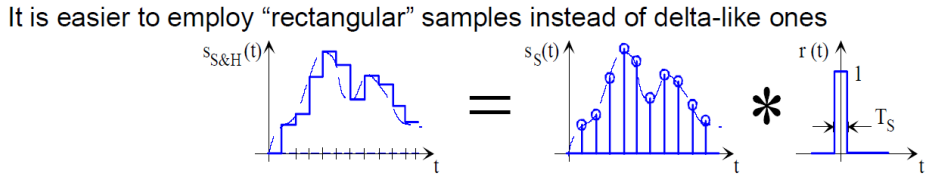
If Shannon is not respected, either if we filter at f_in, we don't simply get our signal but also three replicas, so we have a mis-reconstruction of the original signal.

Anti-aliasing filtering

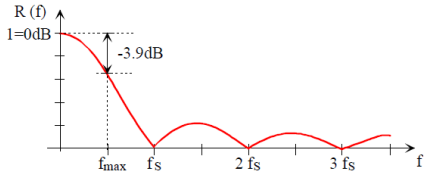


If we know where the frequency content of the signal is, and then eventually we have also spurious disturbances, better not to perform the sampling of the overall signal as it is, but to remove the disturbances.

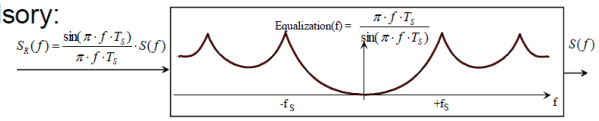
SPECTRAL DISTORTION AND EQUALIZATION



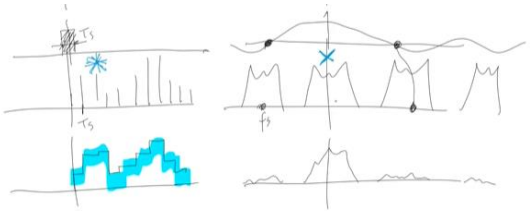
… but this causes spectral distortion:



Therefore **equalization** is compulsory:



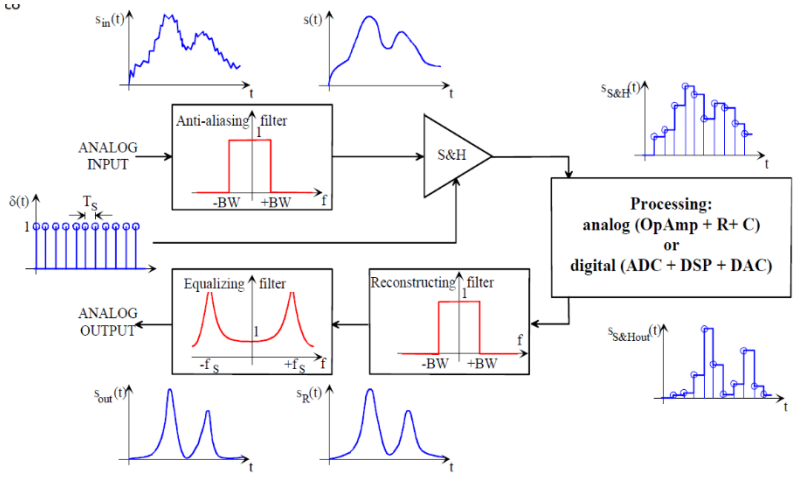
Delta-like samples are difficult to be provided with respect to rectangular samples. Mathematically, the two signals are not the same, but we can say that the rectangular samples is a convolution of the deltas with a rectangular shape whose width is the sampling period.
So we have our original signal made by deltas whose spectrum is composed by replicas of the main spectrum, but know we use a rectangular signal and the spectrum won't be anymore the same as before, because the spectrum of a pulse is a sinh whose zero crossing is equal to fs = 1/Ts. Since in the time domain the rectangular samples are obtained by convolution, it means that in the frequency domain we have to perform the multiplication.



The problem is that if now I apply the filter, the spectrum is distorted due to the sinh behaviour.
Hence we need to use an equalization filter, which is a filter that amplifies the signal with a reconstructing filter that has a gain of 1 in DC and a gain of 1.57 at f_max to recover the attenuation.

SIGNAL PROCESSING CHAIN

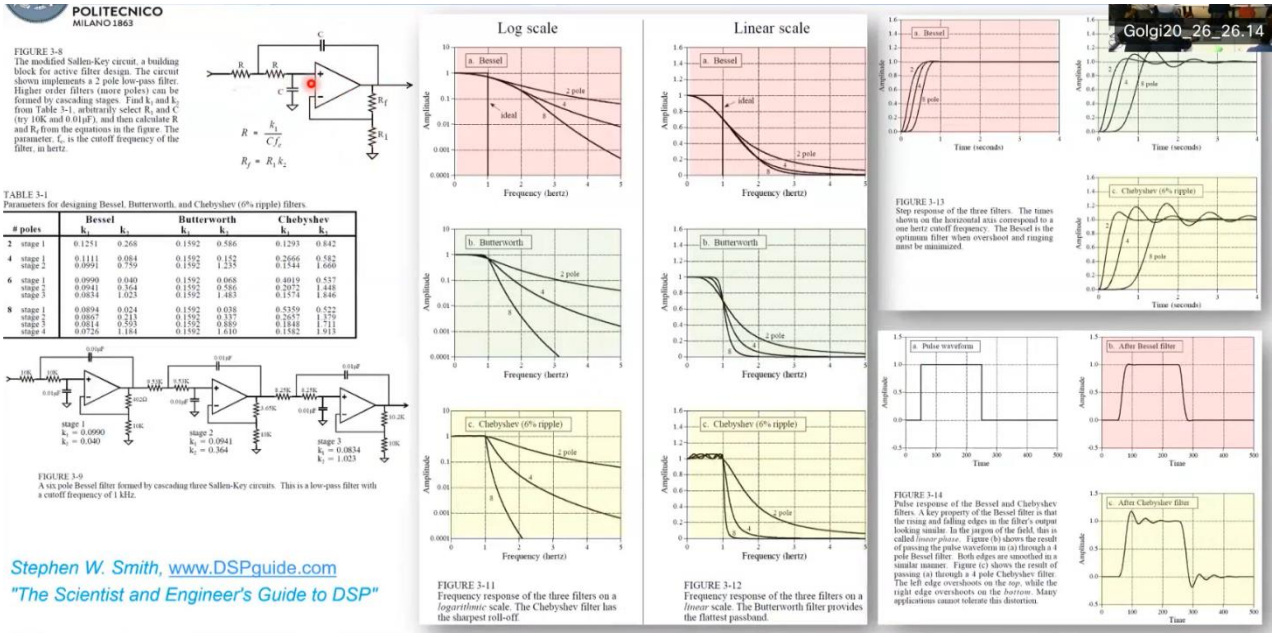
In a global signal processing chain we enter with the signal that can be noisy, we use an anti-aliasing filter and we sample it with a proper Ts.



The reconstruction filter can be identical to the anti-aliasing filter and then we should amplify with an equalizing filter due to the distortion introduced by the S&H.

ANALOG FILTERING

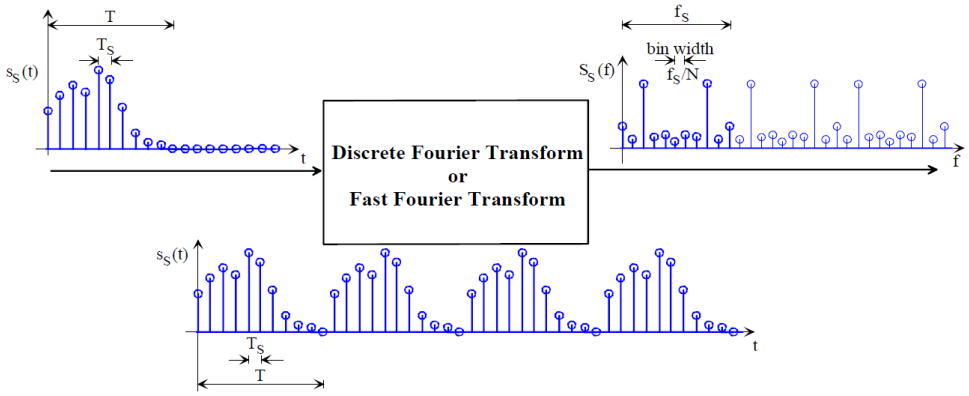
To properly reconstruct a sample signal we need filters, either anti-aliasing or reconstructing or equalization ones. We can have different filters, as below.



The disadvantage of Chebyshev is that if we look at the time response for a step, the Bessel is smooth but slow. The Butterworth can be faster, but we run the risk of overoscillations, that can also be present in the Chebyshev.

HOW TO COMPUTE THE SPECTRUM (FFT)

If we have a signal and the signal lasts for a time T. Then we choose a Ts and we have the samples. If we give the samples to the DFT, the DFT will give us the same number of samples we fed in input, but in the frequency domain. So the number of samples will be T/Ts, and it will be the same number in the frequency spectrum, where the first sample is at f = 0, and the last at fs = 1/Ts.
Moreover, the DFT (and also the FFT) assumes that the signal is periodic. Since after T the DFT doesn't know what there is, it assumes the signal as periodic, retaking the initial value.



The FFT algorithm trusts on the periodicity of the input sequence
... hence FFT spectrum is for the periodic sequence and not the original one!

Since the DFT assumes that the signal is periodic, then a periodic signal has a spectrum made by deltas. But since the spectrum is periodic itself, then it means that the original signal was made by deltas.
Whenever we have a signal, we compute the samples and we apply a DFT or FFT, we get a number of samples in the frequency domain equal to the samples we gave and the spectrum is periodic.

If the original signal was real, then the spectrum is periodic around fs, and last sample is equal to the first one, so its symmetric. Since the maximum frequency is fs and we have N samples, the distance between the samples is fs/N in the frequency domain.

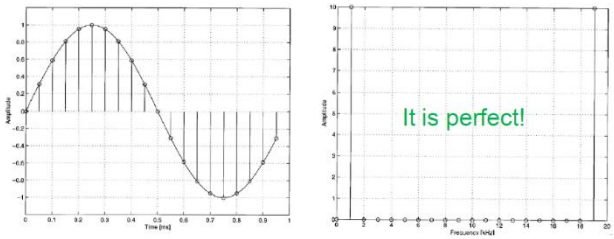
If we sample every Ts, the maximum frequency is fs = 1/Ts. If we sample for a period T of time, the bin width is 1/T.
Let's imagine we have a sinusoid that is continuous, but we just listen to it for a period T and we take 5 samples every Ts. The DFT will give in output 5 samples with a maximum frequency fs, and each sample will be separated by the other by 1/T. However, if we do the same thing but we increase T, the spectrum will be better. fs will be the same, but we have more samples in the frequency spectrum, because the bin width decreases.

To understand the frequency of a certain sinusoid in time, we compute the spectrum and the smaller the bin width the higher the accuracy with which we detect the frequency.
Moreover, the spectrum may belong to a periodic signal, and this may cause issues, as in the example below.

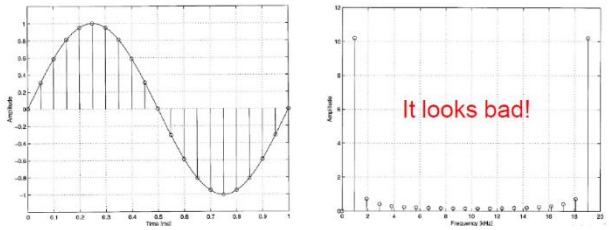
In the perfect case I see that the spectrum is 0 a part from the 1kHz value, that is the frequency of the sinusoid. But if we add a sample, the last one, that is the same as the first one, know the spectrum is bad, because it is != 0 in the intermediate points.

In the first case we have stopped before, and since the algorithm for the DFT and FFT considers the signal periodic, we are doing ok. In the other case it's like modifying the signal.

Correct choice:

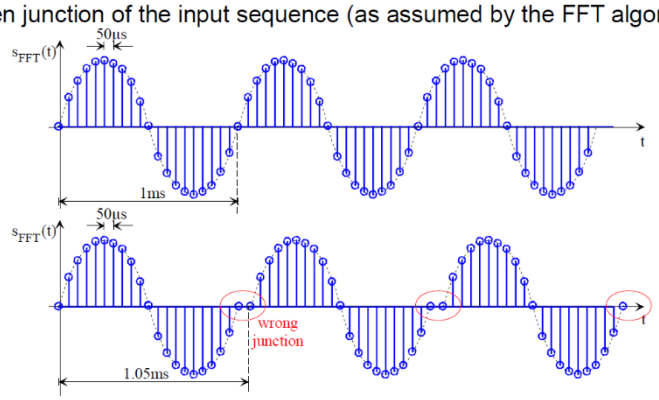


Wrong (unlucky) choice:



So increasing the number of samples is something that must be done thoughtfully.

Due to uneven junction of the input sequence (as assumed by the FFT algorithm)

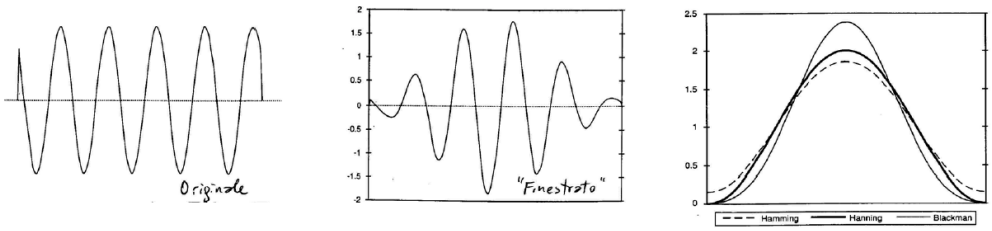


Therefore, let become independent of the junction…

To solve this problem, we can do windowing.

WINDOWING BEFORE DFT

The oddness at the junction can be smoothed out through **windowing**



Rectangular (no windowing): $window(n) = 1$ for $-M \le n \le M$ and 0 elsewhere

Triangular: $window(n) = 1 - \frac{|n|}{M}$ for $-M \le n \le M$ and 0 elsewhere

Hanning (rised cosin): $window(n) = \frac{1}{2} \cdot \left(1 + \cos\frac{2\pi \cdot n}{N}\right)$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

Hamming: $window(n) = 0.54 + 0.46 \cdot \cos\frac{2\pi \cdot n}{N}$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

Blackman: $window(n) = 0.42 + 0.5 \cdot \cos\frac{2\pi \cdot n}{N} + 0.08 \cdot \cos\frac{4\pi \cdot n}{N}$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

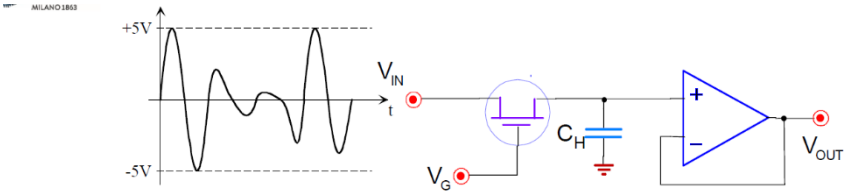
Given our signal, we window it in the time domain. I can use a rectangular one, that however is like no windowing. After windowing, the signal gets well smooth at the end and it is possible to perform a DFT. The drawback is that the deltas in the frequency domain widens a little, but it is not a huge problem.

SAMPLE AND HOLD CIRCUIT

The idea is that we want to acquire a signal that can be modelled with a voltage generator and it's time and amplitude continuous (fully analog). So we need to reach an ADC and then a microcontroller or microprocessor. Between the signal and the ADC we need to introduce a S&H circuit, which needs to freeze the input signal at specific time instant, so we need to choose a sampling frequency fs and, once we have selected the sample, since the ADC requires some time for the conversion, we need to keep the signal constant for a certain amount of time, not just providing deltas of signals.

So while the data is constant the microcontroller gives a pulse to perform the conversion. At the end of the conversion the ADC provides the output data and the microcontroller can read it with its input data buffer. We want to design the S&H.

It can be modelled as a circuit that, when is open, provides a constant information thanks to a charged capacitor. The buffer provides the current without affecting the voltage on the capacitor.



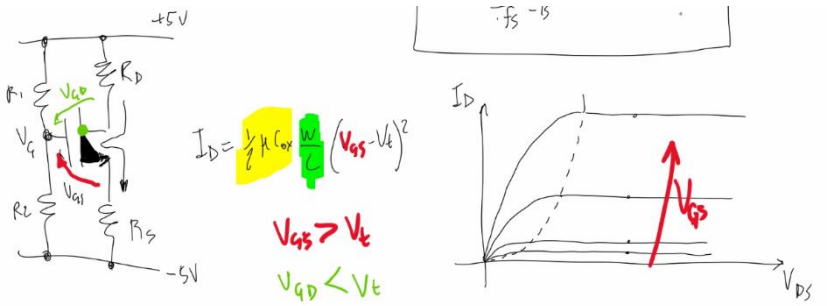
Specs:

Input signal:	-5V÷+5V	20kHz bandwidth
maximum admitted error:	< 0.01% FSR = 1mV	
sampling frequency:	100kHz	

Components: OpAmp: A_O=110dB I_{bias}=50pA
MOSFET: V_t=2V R_{on}=50Ω C_{gs}=0.5pF C_{ds}=0.1pF

SWITCH

The switch can be a MOSFET with a proper Vgs so that the current that flows through it is given by the following.

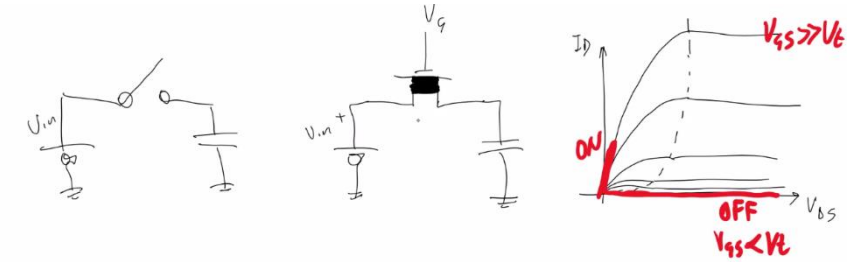


The other parameters are depending on the technology of fabrication, except for Vgs. We want to provide channel at the source and we don't want it at the drain. So Vgs must be higher than the threshold voltage, and Vgd should be lower than Vt, not to have channel at the drain. This if we plot Id vs Vds we have the characteristic on the right. The characteristic increases quadratically with Vgs. If Vgs < Vt the transistor is off, otherwise it starts to increase the current.

So the transistor can be used as an amplifier if we operate in the saturation region, where the current is proportional just with the input signal and not with the output. This is good because Vgs varies depending on Vin and Vout is not varying depending on itself. So the red and green are the required biasing conditions to have a MOSFET as an amplifier.

Moreover, these two conditions can be put together as Vds > Vov.

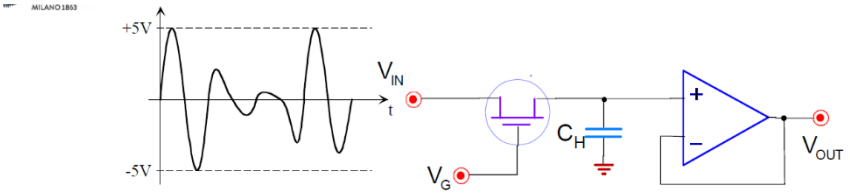
Instead, in the S&H we need a switch, not an amplifier, we need to connect an input voltage to an output voltage. So the transistor will be used in another configuration. We want to have either no channel both on the source and drain (open switch) or to have channel both in the drain and source. So compared to the amplifier case, we want it to operate in another region, in the ohmic region with sufficiently high Vgs or in the completely off condition when Vgs < Vt.



If Vgs >> Vt we have a very low r_on = dVds/dId, and I want the r_on to be very small so that Vds is the smallest possible.
In the amplifier we had to indicate which was the source and the drain, whereas here I have the same behaviour on the source and drain, so I can connect it as I want.
If e.g. I have a nMOS and I want to operate it in triode regime (so channel on both source and drain), given Vin and Vout (that should be equal to Vin), then Vg > Vin_max + Vt to have the channel, and if Vin moves +-5V, I must be sure that Vg is higher than the highest Vin value and Vt. So I have to apply at least 5.8V to have the transistor on. Instead, if I want it off, the worst scenario is when Vin reaches the lowest value, so Vg < Vin_min + Vt not to have the channel.

CIRCUITAL IMPLEMENTATION

So the analog signal is applied to the source or drain of the mosfet, accumulated on a capacitor and given to the buffer. The value must be stored and computed without committing an error, so we need to understand which the sources of errors are. The following are the specifications we are considering.



Specs:

Input signal:	-5V÷+5V	20kHz bandwidth
maximum admitted error:	< 0.01% FSR = 1mV	
sampling frequency:	100kHz	

Components:

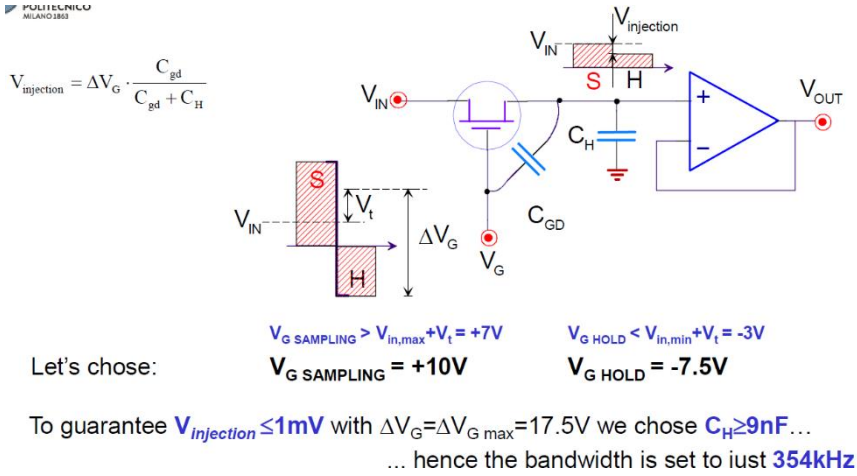
OpAmp: A_O=110dB I_{bias}=50pA
MOSFET: V_t=2V R_{on}=50Ω C_{gs}=0.5pF C_{ds}=0.1pF

I'm considering a transistor whose on resistance is limited to 50ohm.

ERRORS

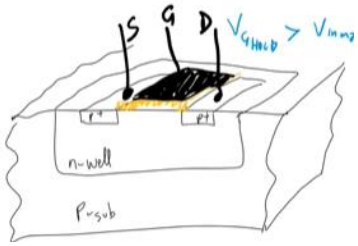
CHARGE INJECTION

Let's compute the Vg value that keeps the MOSFET close. I'm considering Vin moving from +5V to –5V, if we apply a Vg = 10V to close the switch and Vg = -7.5V to close it.
Having chosen a n-channel transistor we need a Vg high to keep it close and low to keep it open. It's the opposite if we would use a p-channel, so to have it on Vg < Vin_min – Vt and in the off configuration Vg > Vin_max – Vt.



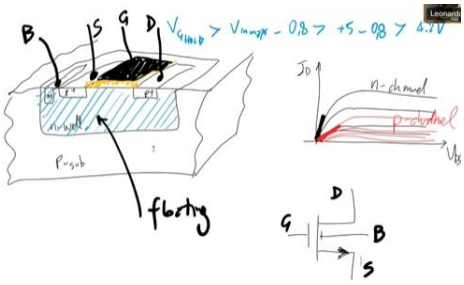
If the MOS transistor is fully symmetric, it's difficult to understand which is the source and which is the drain.

If the MOS transistor has the following cross section, then it is symmetrical (pMOS).

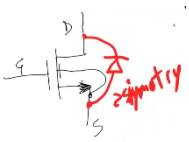


nMOSFETs carry higher current with respect to pMOSFET because of the higher mobility. But then the type of transistor to be chosen depends also on the voltages we are able to apply to the gate of it, if positive or negative.

In reality, a transistor is not fully symmetric because we have, in the previous example, a floating n-well not connected to anything, and to avoid this (that could change the threshold due to the body effect) another n+ region is created in contact with the well and it is the bulk terminal. So we have 4 terminals in the transistor.



If, instead of 4 pins, the transistor has 3 pins, it means that internally the manufacturer has connected the bulk to the source. The fact that we connect the source with the bulk causes, however, an asymmetry. Thus we create a reverse bias diode.

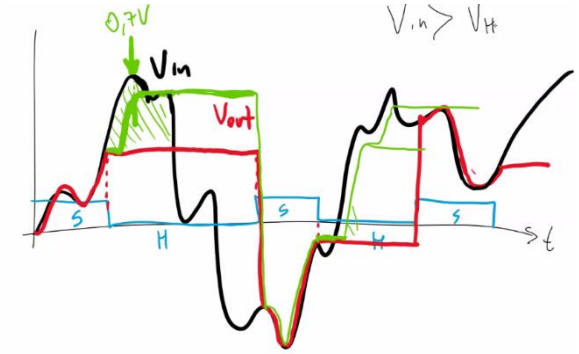


The diode has no effect if I use the transistor as an amplifier, because the drain is usually at higher voltage than the source for a nMOSFET, but it is relevant if I use it as a switch for S&H.

If we have the Vin and we buy a discrete component transistor (circled) with 3 pins, if Vin is lower than the voltage stored on the capacitor we are ok, the transistor is off and to turn it on we have to apply high Vg. But even if Vg is in the hold phase, so sufficiently low to have the MOS off, still Vin moves and if it

happens that Vin > Vh the diode is on, so the transistor is off but the current flows anyhow and we charge the capacitor.

So we have our Vin, the S&H signal, and when we sample the transistor is close, so Vout = Vin and when we open we would like Vout to be in hold phase, and when we reclose the switch we want to quickly track Vin again. The red one is the ideal operation. But if Vin > Vhold, then the parasitic diode turns on and it causes the transistor to conduct a current, so Vout follows Vin with a distance equal to 0.7V, voltage drop across the diode. Then when the voltage decreases below Vhold, the capacitor has stored that value and it goes in the hold condition and it stays like there until the signal becomes high and chases it again.

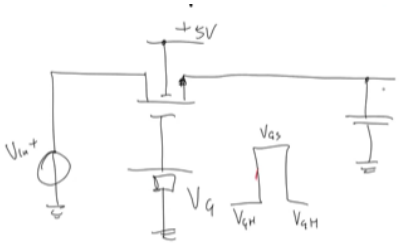


The green one is the behaviour we want to avoid, so we cannot use a 3 wires transistor, we have to use a 4 wire transistor, with the bulk contact connected to the most negative PS (if nMOS, positive is pMOS). If the bulk is sufficiently negative, we don't suffer no more of the problem because we sufficiently reverse bias the bulk-source and bulk-drain diodes.

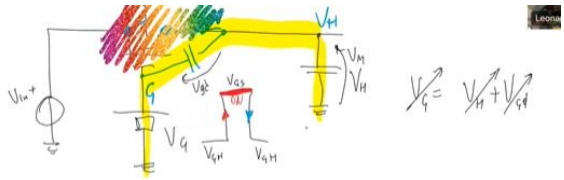
So we cannot use a transistor with the source connected to the bulk.

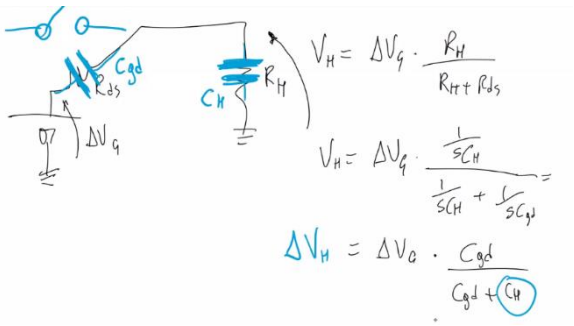
Every time we open and close the switch, any parasitic capacitance between a node that goes up and down (gate) and the other one that is connected to a floating component (the capacitor), that parasitic capacitance (Cgd in the image) causes charge injection. Since the sampling capacitor is connected to anything, we have this problem.

So we have our Vg that moves from Vg_hold to Vg_sampling and then again Vg_hold.



During the rising edge transition of the Vg, the transistor is on and closed, so there is current and the capacitor charges up to Vin. When we have the falling edge transition, soon or later the transistor will be open and the switch will be open, so there should be no variation at Vhold, but there is a coupling due to the parasitism of the capacitance. If we change the volage at the Vg, we change the voltage at the series of the two capacitors. If Vg changes, both Vh and Vgd changes. How much? We use the partition of voltages.



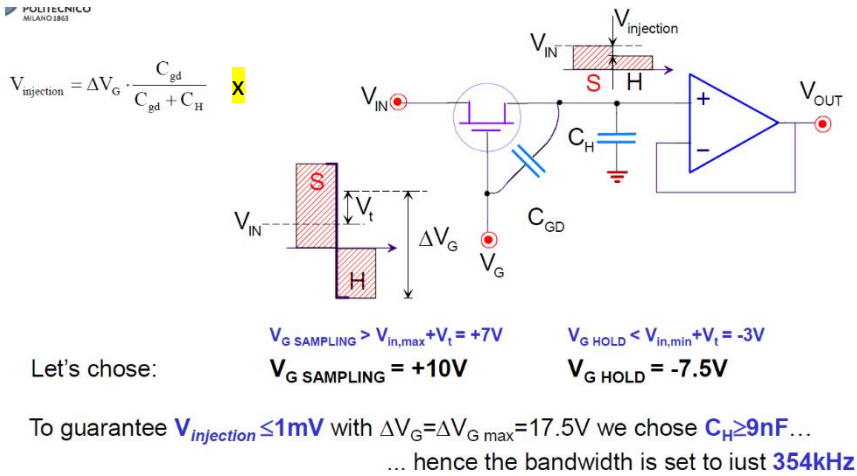


So the higher Ch is, the lower the change in the Vh stored in it. So to reduce this error in the falling edge transition we need to increase the Ch value.

For instance, if delta_Vg total = -17.5 to move from sampling phase to the hold one, we should store the value of Vin, but due to charge injection we see a drop in the voltage stored on the capacitor, delta_Vh, due to charge injection.

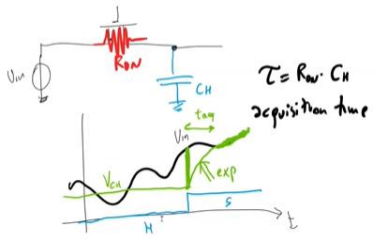
It is called charge injection because if we change the voltage delta_Vg, voltage across the capacitor is given by C = delta_Q/delta_V, so given a capacitor Cgd, if we change the charge across it also the voltage across it will change. The same for Ch. Since Cgd and Ch are in series, the two delta_Q are equal, so if we change delta_Vg we cause the injection of a charge in both, so both Ch and Cdg must discharge.

So charge injection happens every time we open a transistor and from the node of the gate to the high impedance node where there is a capacitor we have a stray capacitance.
We know the value of Cdg and the full jump from Vg up to Vg down, -17.5V, so we know the value of Ch that we can choose.



The highest capacitance possible is not a good idea because then when we close the switch there will be a given r_on that will go in series with the Ch that cause a tau that gives us the acquisition time of the system.
If we have Vin and the S&H signal, the stored voltage is on the capacitor, and when we close the switch we should quickly reach the new value of Vin but the charge on the capacitor will increase with an

exponential curve and it depends on the tau. So either we use a lower r_on or we choose a low Ch, as low as possible to speed up the acquisition but not too low to have relevant charge injection.



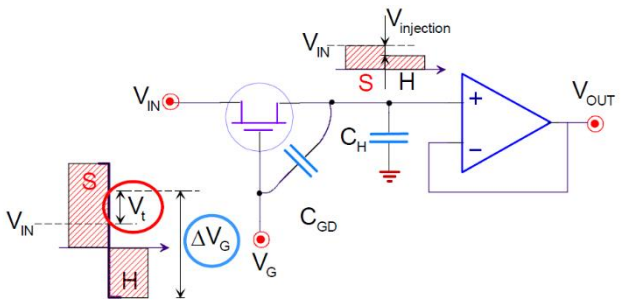
To be more precise, in the equation x we considered delta_Vg as the full transition, from +10V to -7.5V, but when the switch is closed, the voltage on the capacitor is set by Vin. The capacitor is then alone when Vg is sufficiently low and given Vin, if we apply a Vg > Vin + Vt the transistor is on. So the delta_Vg to be used is not the full Vin swing, but just the portion where the Vg determines the openness or closure of the transistor. The problem is that delta_Vg, in this way, depends on Vin. This also means that Vinjection changes with Vin.

Hence the Charge Injection is not a constant error.

APERTURE-INDUCED NON LINEARITY

POLITECNICO MILANO 1863

$$V_{injection} = \Delta V_G \cdot \frac{C_{gd}}{C_{gd} + C_H}$$



$\Delta V_G = V_{IN} + V_t - V_{G\,HOLD} = V_{IN} + 7.5V$ It is not constant, but depends on V_{IN} !

Charge injection error is not constant:

$\Delta V_{G\,min} = V_{IN\,min} + V_t - V_{G\,HOLD} = -5+2+7.5 = +4.5V$ $V_{injection} = 250\mu V$

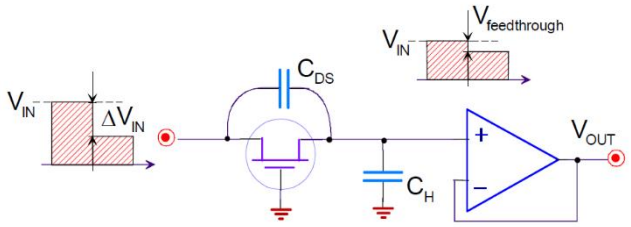
$\Delta V_{G\,max} = V_{IN\,max} + V_t - V_{G\,HOLD} = +5+2+7.5 = 14.5V$ $V_{injection} = 806\mu V$

This is not a constant error. If when I take a Vin and close and open the switch and the Vinj is constant, there is no problem, because I know how to compensate it eventually, because I know the gap between Vin and the value of voltage I'm storing on the capacitor. But since Vinj changes depending on Vin, this is bad.
In the samples I'm acquiring from the microcontroller, we have different errors if the sample has a high or low value. The error gets reducing the more the value sampled becomes low.

So a constant error can be compensated, such as the offset of the opamp, but if it is not constant it is bad.

SIGNAL FEEDTHROUGH

$$V_{feedthrough} = \Delta V_{IN} \cdot \frac{C_{ds}}{C_{ds} + C_H}$$



With $\Delta V_{IN\ max}$ = 10V we get $V_{injection}$ = 111µV

… negligible (in this case) compared to the requirements of <1mV

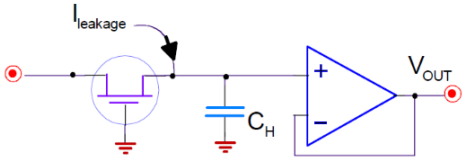
When the switch is open there is a parasitism between source and drain that causes the values across the capacitor Ch to change, even if the switch is open.
If we apply +5V and the switch is close, then I open it and I should have 5V on Ch, but if the input drops to -5V I get a Vinj on the capacitor of 111uV, and it's ok because I'm tolerating errors up to 1mV, so we can neglect it.

Of course we cannot neglect it if the Cds is high, the swing in input is high or the Ch is low.

DROOP

When the switch is open, the capacitor stays constantly charge in theory, but if there is a leakage, so a current in the transistor or the bias current of the opamp, this leakage current causes a voltage drop.

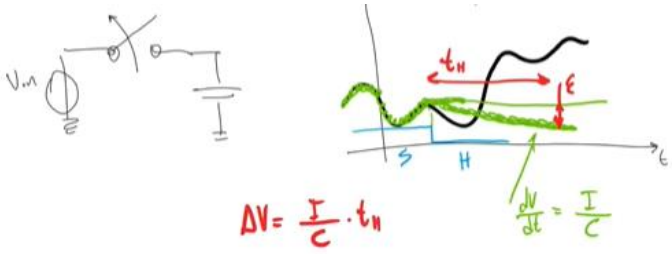
$$\Delta V_C = \frac{I_{leakage}}{C_H} \cdot t$$



With $I_{leakage}$=100pA and C_H=9nF, the stored voltage will droop by 11mV/s

Hence with f_S=100kHz we get ΔV_C=0.11µV every 10µs Hold duration

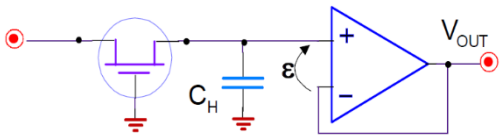
This means that when the transistor gets open, Vin can move and the Vh should remain constant, but it is not like that, we have a droop that has a slope dV/dt = I/C, so the error after a given aperture time is the red one.



If in our case we considering the operation with a sampling frequency fs = 100 kHz, it means that in the worst case scenario the sampling is very short and the hold is very long (not the same duration for both).

If so, the error due to the droop is 0.11 uV, which is still negligible because Ch is sufficiently large and Ileakage is low. But if we change the opamp, then droop can become an issue.

BUFFER-INDUCED NON LINEARITY

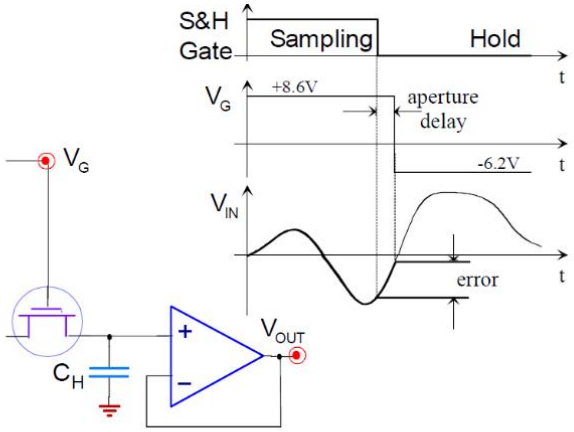


To guarantee ε<1mV we must have $A_0 > V_{out,max}/\varepsilon = 5{,}000 = 74dB$

We know the buffer has a limited A0, so ideally the epsilon is 0, but since A0 is finite, to have a given voltage at the output the epsilon will be Vout/A0. Since I want epsilon < 1 mV, then Vout/epsilon = 5000, meaning that we cannot buy whatever opamp we want, but an opamp whose A0 is higher than this value.

APERTURE DELAY TIME

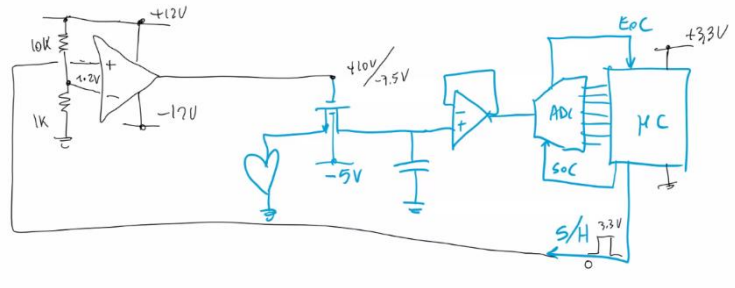
$$\Delta V_{aperture} = \left.\frac{dV_{in}}{dt}\right|_{\max} \cdot T_{aperture} =$$
$$= 2\pi \cdot f_{\max} \cdot V_{in,\max} \cdot T_{aperture}$$



With $T_{aperture}$=1ns we get $\Delta V_{aperture} = 2\pi \cdot 20kHz \cdot 5V \cdot 1ns = 0.63$ mV

It's a dynamic error, so far we have considered static ones. Data is sampled in the falling edge transition because during sampling the switch is close and when I move from S to H the switch gets open. So the sampling transition is the falling one, because I remain constant in the H phase when I open the switch. But maybe the Vg voltage requires some time to commute from high to low when the microcontroller gives the command.

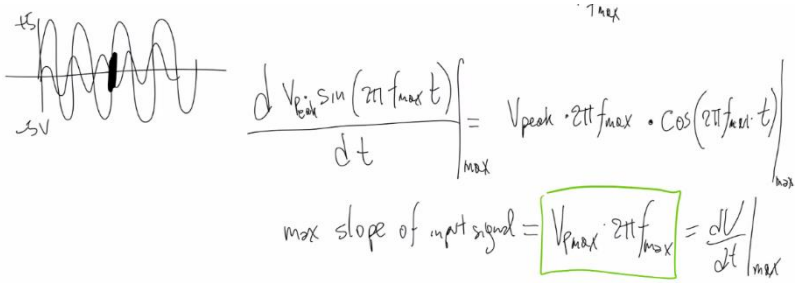
The uC commands the ADC ad the S&H phases. The typical PS of uC is 3.3V, so the signal to the S&H circuit is e.g. from 0 to 3.3V. But to the gate we need 10V/-7.5V. so we cannot directly connect the uC, we need a voltage amplifier stage between the uC and the S&H circuit.



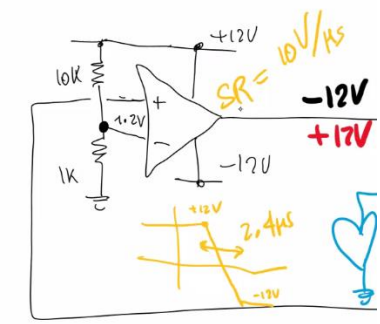
In the hold phase we apply -12V hence, and in the S phase 12V. This is good because even higher and lower the one requires. But from the time when the apply the command and the swing is done by the transistor some time instants appear. So the aperture delay time causes a sampling error, because we are not sampling the value at the time instant we want.

Let's compute this error. It depends on how fast Vin is, so if I don't know Vin I cannot quantify it. The highest error occurs when Vin changes with the fastest speed. The maximum speed of variation is where we have the maximum frequency fmax. So the maximal speed is like having a sinusoid at fmax of the signal.

So let's consider a sinusoid at fmax: Vin = Vp*sen(2*pi*fmax*t)
Let's now compute the maximum value of the slope of the signal, so we compute the derivative and we take the maximal value. So the maximal value of a sinusoid is:



If we know this and we multiply this for the maximum aperture delay we get the maximal possible error. Aperture time Taperture depends on the electronics, such as due to the SR of the opamp that drives the gate signal.

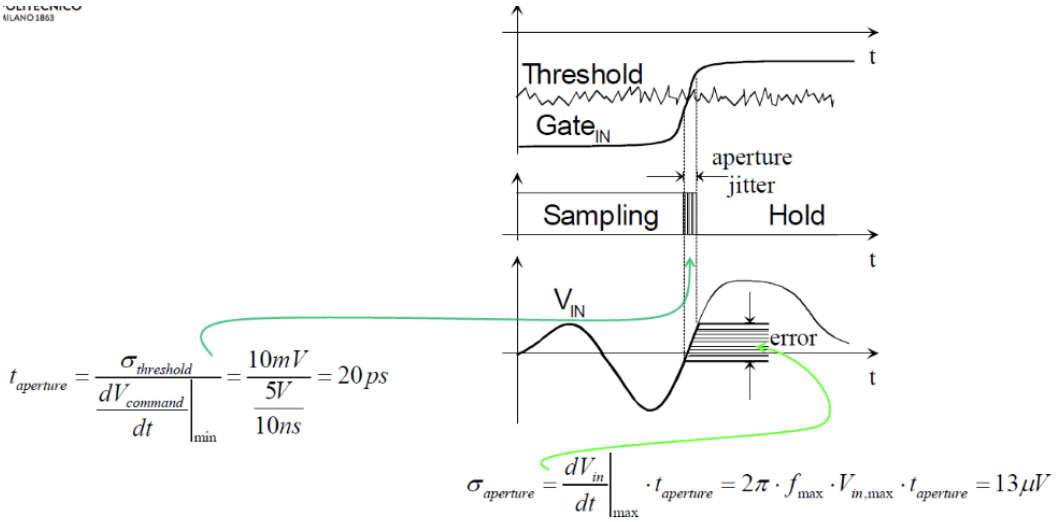


This error is pretty high with a Taperture = 1ns even. If the SR is so large, so us, this error can be very huge.

Is this delay really an issue?
It depends on the application. In some ones, like in the audio applications, we just need a sufficient number of samples with a proper spacing coherent with the Shannon theorem. But the Shannon theorem never tells where to start, so the comb with which we sample the signal is different, if the delay is constant along all the samples, then there is no problem at all. So if the aperture delay is constant for all the samples it is not a problem.

However, there are also other applications where it is important to properly sample the signal and measure its intensity to know the position of, for instance, of the peaks starting from a specific time instant. In this case, even if the delay is constant we run the risk of sampling a sub-maximal value, not the maximal one.

APERTURE TIME JITTER



In case of a CMOS drive: $\Delta V = 6 \cdot \sigma_{apertura} = 75\mu V_{pp}$ (**18bit**) trascurabile

The problem is that the aperture delay could be not constant, so it is really a big issue in this case.

Let's consider the following example. Previously, when the uC applies the commutation, we crossed the threshold of the Schmitt trigger, output went low and we had the opening of the switch. And this happens at a precise voltage.

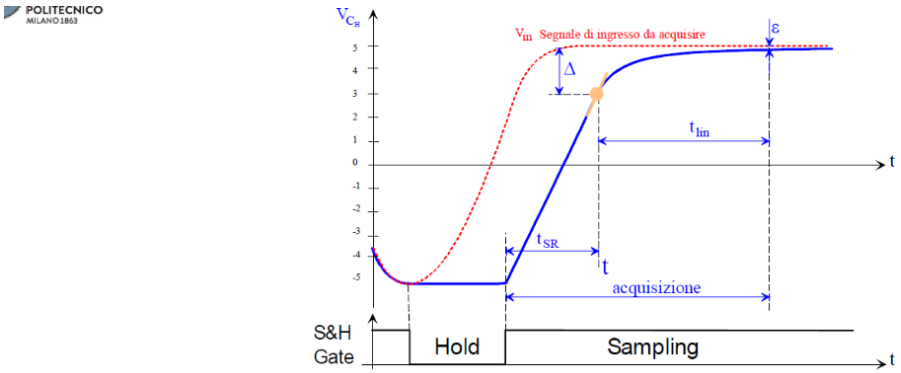
However, there may be some fluctuations on the power supply and so on the threshold. The opamp is noisy as well as the resistors, so the threshold is not constant but noisy.
If the threshold is not constant, apart from the delay to the time the threshold is really crossed, there may be a jitter.

To quantify it we need to know the vertical fluctuation of the threshold, then the slope of the commutation so that we can compute the jitter in amplitude. Given the vertical sigma, the horizontal one depends on the slope of the commutation. Once we have the horizontal jitter, the vertical jitter of the stored value depends on the slope of the Vin signal, and we can use the same equation used for the aperture delay time.

The 13uV value we get is either the sigma or the peak-to-peak value depending on the way in which we quantify the fluctuations in the threshold value.
To understand how big this error is we need to look at the following ADC. Depending on the number of bits of the ADC, the ADC can quantize the analog input signal with a LAB resolution of mV, uV or less.

So depending on the number of bits this error is negligible if bits are few, but if the ADC has a lot of bits, then the error is significant because the ADC resolution is very high.

ACQUISITION TIME



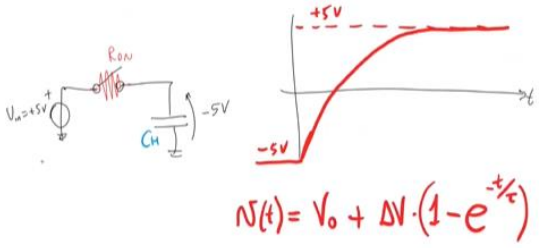
Acquisition (Sampling): linear ramp $t_{SR} = \frac{10-\Delta}{SR} = \frac{(10-1.25)V}{2.8V/\mu s} = 3.1\mu s$

and exponential charge $t_{lin} = \tau \cdot \ln\frac{\Delta}{\varepsilon}$

For example: with $I_{Out.max}$ = 25mA we get $\Delta = 50\Omega \cdot 25mA = 1.25V$

It's still a dynamic error. We are in the sampling phase and we track the signal. Then we move to the hold phase. Vh is fixed and Vin moves. Then we move again to sampling; in the worst case scenario we have stored the minimum input signal on Ch and then we need to acquire the maximum input signal. So we need to move e.g. from -5V to +5V. The time needed to reach the final value is called acquisition time.

If the circuit was the following one, when I close the MOSFET it acts as a Ron, so I have an exponential increase to 5V from -5V.



The tangent of the curve in the origin touches the asymptotic value after tau. Hence the following.

$$\left.\frac{dV}{dt}\right|_{max} = \frac{\Delta V}{\tau}$$

From this equation we can compute the time that it takes to reach the final value. Theoretically, the error epsilon = 0 will be reached after an infinite time. Once defined the maximum error we can accept we can define the acquisition time (equations to be remembered).

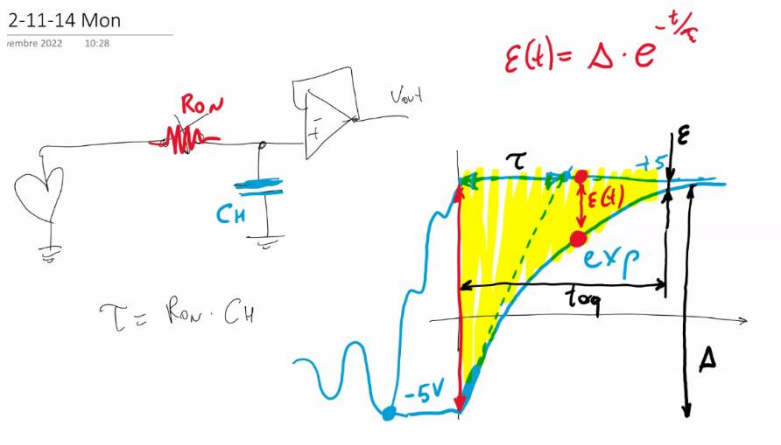
$$\varepsilon_{max} = \Delta V \cdot e^{-t_{acq}/\tau} = \varepsilon$$

$$t_{acq} = \tau \cdot \ln\frac{\Delta V}{\varepsilon}$$

In our case we have to perform the following calculation.

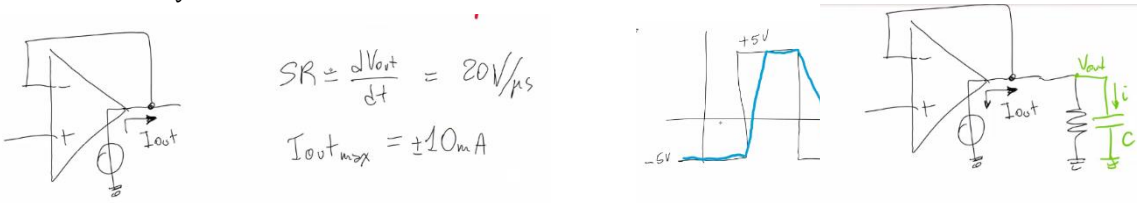
$$t_{acq} = R_{ON\,max} \cdot C_H \cdot \ln\frac{10V}{1mV} = \tau \ln 10000 = \tau \cdot 10$$

Time that the sample and hold takes to move from the previous stored value during the hold phase up the then new value during the sampling phase. It is a transition modelled with an exponential charge.
When the switch is closed, it behaves like a Ron, the capacitor has its capacitance and so we have a tau = Ron*Ch. Hence if this is the circuit, Vout will move from the previous value (-5V in the worst scenario) to the new value (+5V). We need the time for the capacitor to charge to the new value. We will have an exponential increase whose tau is the one previously calculated.
The time needed to acquire the proper value depends on the epsilon value we are willing to reach at the end of the transition.

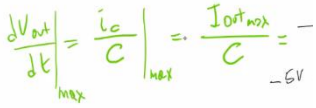


From the equation for epsilon we get the time of acquisition.
Unfortunately, in a real circuit we can have also other limitations. The circuit could be driven from a previous stage, e.g. another opamp, that could introduce a further limitation.

The opamp has two issues: the slew rate and the maximum current the opamp can provide to the external circuit.
Hence if the load is a resistive one and the input varies too fast, due to SR the output of the opamp will have a transition limited by the SR.



Moreover, if the output of the circuit is not just a resistor but a capacitive load also, the voltage output depends on the current that flows in the capacitor divided by the capacitance itself. If the output is current limited, the maximum variation we can reach by charging the capacitor depends on the maximal current of the opamp.

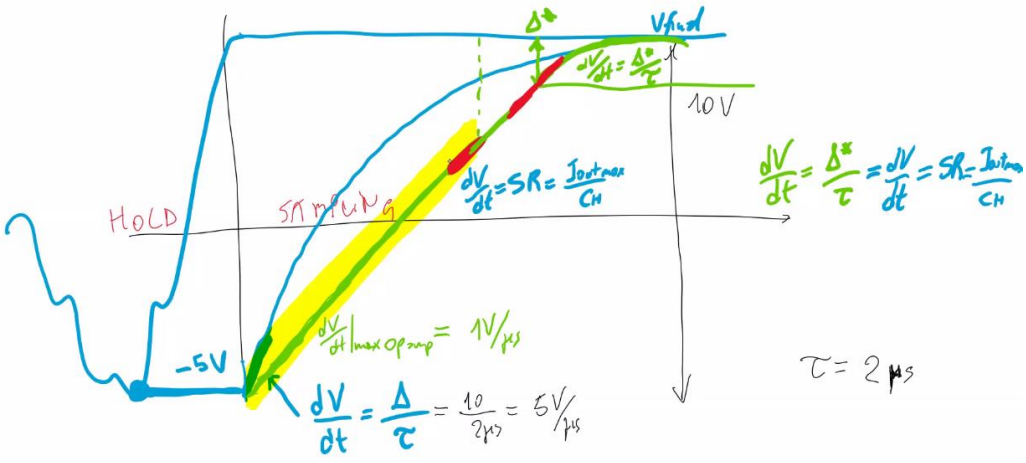


Hence the output with a capacitor has a limited slope that can achieve depending on the capacitance value. Hence we have a limitation in the swing from one value to the other that is slower than the one predicted by the external SR of the capacitor, but that is due to the Iout,max capability of the opamp. So the voltage swing can be determined by the SR or the Iout,max depending on which is the minimum.

If before a S&H stage we put e.g. a buffer, when we close the switch we have the limitation set by the tau Ron*Ch, but also another limitation set by the Iout,max of the opamp or the SR of the opamp. We need to understand which between these last two is the lowest (i.e. the more limiting) and we consider it.

If this is the case, in the transition from on value to the other, it is not an exponential curve till the end, because it is given just by Ron*Ch (and so delta/tau initial slope), but the slope at the beginning of the transition may be smaller if the input opamp cannot provide the one we need.

So what happens is that we store a value and once we move from the hold phase to the sampling phase the switch closes and we need to reach the new final value. But the transition is not an exponential curve with an initial slope of delta/tau if the opamp has a limitation that is smaller. We will have the dV/dt max of the opamp. This limitation of the opamp causes the transition to bel slower, and it will stop acting on the circuit when the residual delta we need to perform will have a new value delta* so that it is left just an exponential transition, and this happens when the two slopes are equal, the one of the starting exponential to the one limited by SR.

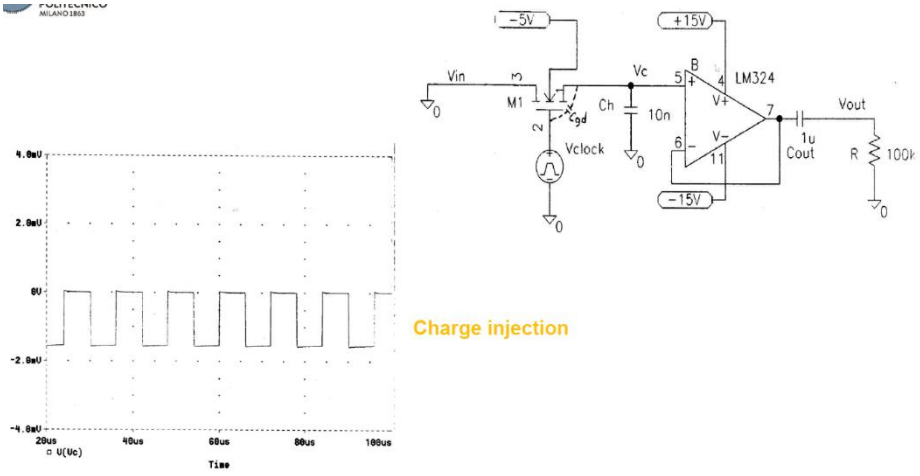


We need to find the value of delta*, that is the portion attended with an exponential curve.

$$\Delta^* = \frac{I_{out\,max}}{C_H} \cdot R_{ON} \cdot C_H$$

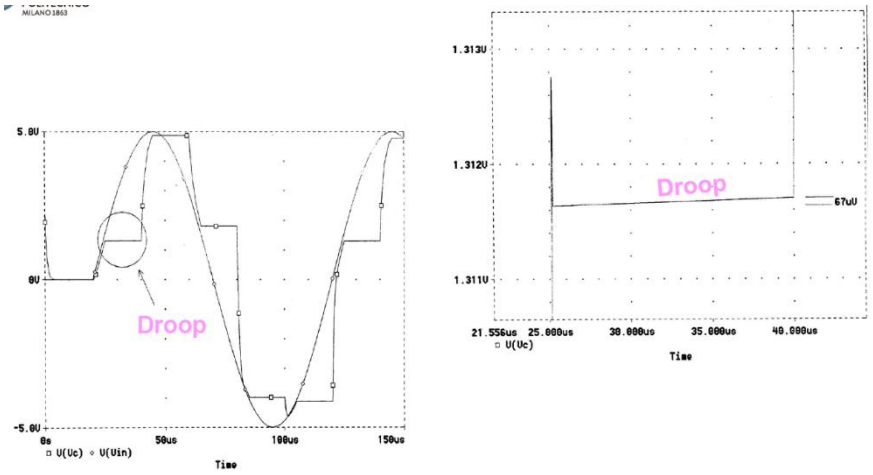
So on the top of the transition t_acq*, where the transition is exponential, we need to add a t_SR limited transition, where the transition is linear. The total t_acq is the sum of these two times (at the written test we can forget about the SR limitations and consider the transition exponential).

Example of SPICE simulations



The switch is made by a n-channel transistor. In the + voltages we are in the sampling phase. Then when the gate goes negative there is charge injection that stores charges on the capacitor, and voltage goes down to -2.2mV.

If we consider a sinusoid and we keep closing and opening the transistor, if the transistor is close we are in the sampling phase. Then we close it and we open it again, but Vout doesn't reach Vin immediately, but later due to t_acqu. If we open the switch before reaching the input signal we are sampling a wrong value.

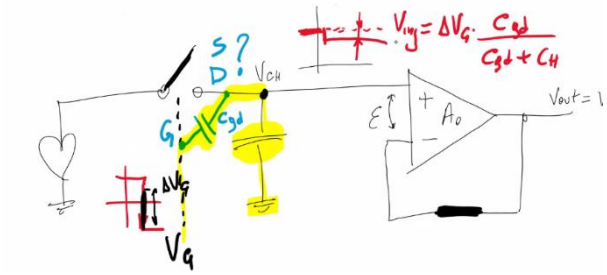


If we zoom the region where we have the hold phase, we have a peaking when we open the switch (due to charge injection) and then during the hold phase we are not flat, but increasing or decreasing due to droop.

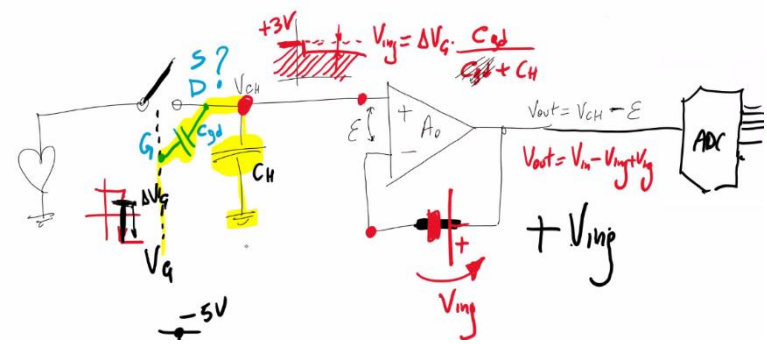
CHARGE INJECTION COMPENSATION

In order to compensate charge injection we introduce a dummy pair. In fact, in a standard S&H, in theory Vch = Vout when the switch is close, but when we open the switch, due to the parasitism and the capacitance Cgd, we have the voltage on Ch to change because there is a path through Cgd. So value stored on the capacitor is not the ideal value, but we have an injection error.

Depending on Vin we have a certain value of delta_Vg. So this error varies with Vin so it cannot be subtracted after the ADC conversion.



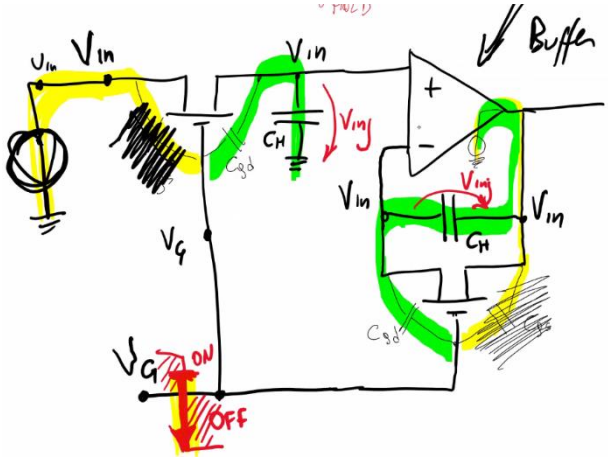
What we can do is trying to compensate the error in another way. If we know the value, we can sum that error in the circuit. Since the value stored on the capacitor decreases by Vinj we can place a battery equal to Vinj so that the output voltage has this quantity readded to the output.



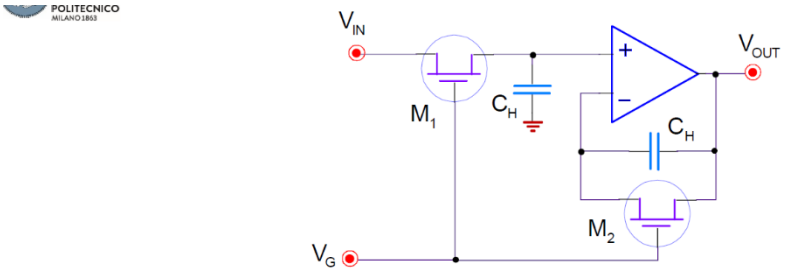
To put a battery in series to the feedback loop of the opamp we can create a dummy cell. To compensate an error due to a transistor we introduce another transistor. The buffer has in feedback a transistor that closes when the sampling transistor closes (so we have a buffer) but that it opens also when the other transistor is open. So we add also a capacitor in feedback so that when we open the transistor the capacitor acts as a battery and the feedback is connected, so the opamp behaves like a buffer. Of course we shouldn't wait too much or the capacitor decreases.

The advantage of this configuration is that both the two transistors have parasitisms Cgd, when Vg is high we charge the capacitor in feedback to 0V. Then, when we apply a transition to Vg, there is a portion of the transition that causes charge injection. But the value of Vg that cause the transistor to move from on to off happens in the same way for both transistors. So the two capacitors will experience the same charge injection. This charge injection through Cgs in the input mosfet is of no importance because we are injecting in a voltage generator, so Cgs is not impacting on charge injection. The effect is on Cgd.

In the dummy cell, again the Cgs is connected to the voltage source of the opamp, so it has no effect, while Cgd touches a node that is floating, because the transistor is open (high impedance) and then we have the – terminal (high impedance). So charge injection affects in the same way the two capacitors.



This applies if the two cells are equal. If the two have some mismatches, the final error at Vout will be the Vinj,1 – Vinj,2 (Vinj,2 being the one in the feedback of the opamp). What remains is x.



M2 mimics M1 and eventually inject the same charge, but in oposite direction

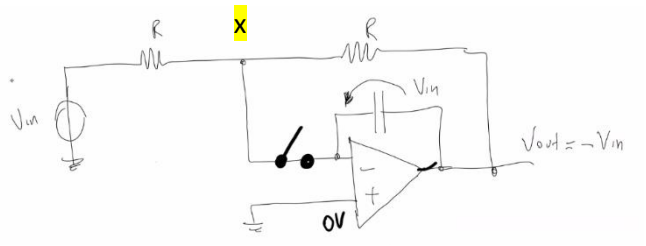
Only the residual contriution due to the mismatches remains $V_{injection} = \Delta V_G \cdot \frac{C_{gd}}{C_H} \cdot \left(\frac{\Delta C_{gd}}{C_{gd}} + \frac{\Delta C_H}{C_H} \right)$ x

In this way it is possible to reduce C_H from 9nF down to just 450pF, with the same charge injection.
So the acquisition time improves (300ns instead of 6.3µs); the signal-feedthrough worsens

In any case, Vinj is now lower, so we can use also a Ch that is smaller, which allows to have a shorter acquisition time. The problem is that, if we reduce Ch, there is also the Csd parasitism, and signal-feedthrough increases. If Ch is huge, this latter error is negligible, but if we reduce Ch it becomes relevant.

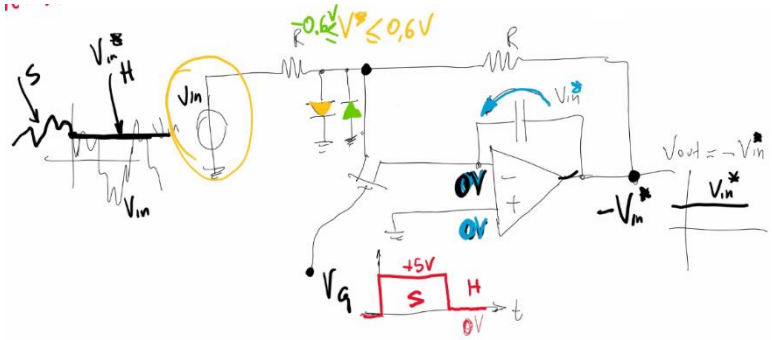
SIMPLIFY DRIVING

Let's suppose to have an inverting amplifier with the resistances of the same value, so a gain of -1. We also put a capacitor in feedback to charge it during the sampling phase. Now we want to introduce a switch, and we have to place it like in the image.



Thus we still have a feedback through the capacitor and, even if the input changes when we open the switch, in output we have Vin. So the switch must be placed between nodes where the voltage is 0 so that the transistor will be easy to turn on or off. In fact, node x will be 0V in the sampling phase when the transistor is on. Moreover, in this brand new S&H the voltage we need to apply to the mosfet is simply +5V or 0V to keep it open or close.

However, it the switch is open, node x varies because it is between a varying source (Vin) and the output. But this is not a problem because the output voltage is set by the capacitor, so it is not changing.
The problem is then to bring node x back to ground, and if it was to a high value, the transistor may have problems. So I don't want node x to be too far away from the VG value.



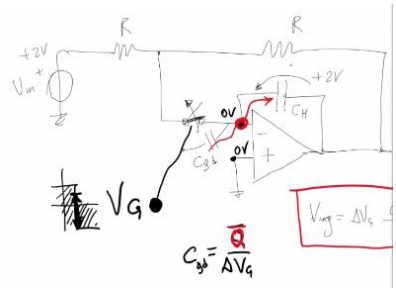
So what I can do is to introduce two diodes in counter opposition so that the voltage at node x is less than 0.6V or below -0.6V. Now when we close the switch the t_acq is faster.

As a final concept, we have to consider charge injection. Every time we open the switch we change the value stored on the capacitor due to charge injection. Due to the parasitism Cgd the voltage Vch changes.

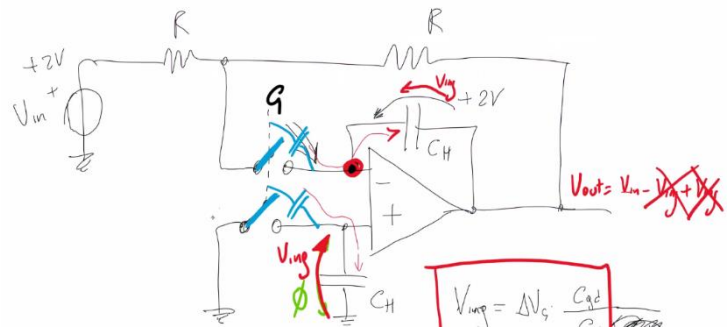
$$V_{inj} = \Delta V_G \cdot \frac{C_{gd}}{C_H}$$

There is no Cgd at the denominator because of the virtual ground.

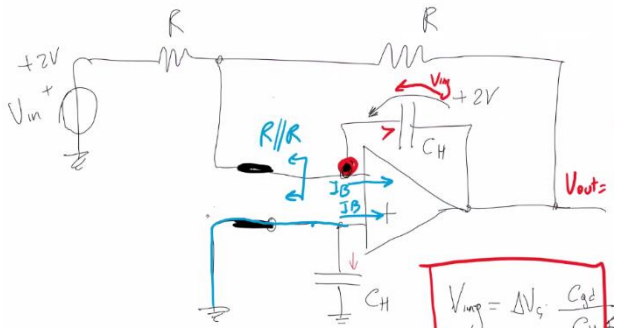
So if we apply a transition at the gate delta_Vg, we apply it between Vg and ground, so charge variation across Cgd is C = Q/delta_Vg. But this charge must flow into Ch, and Ch touches ground, so its output changes because the Q is the same but the capacitor is different.



Hence every time we open the switch we cause Vinj in Ch.
To compensate for Vinj we cannot use a dummy cell in feedback because Ch is already in feedback, so we have to play with the other pin. In order to compensate charge injection we have to include a dummy cell by moving the + terminal by the same amount of the – one. The two transistors are driven with the same gate voltage. When we close both switches, the added Ch has 0V stored because of ground. When the switches are open, the added capacitor charges up to Vinj and the same the feedback one, and again they cancel out.

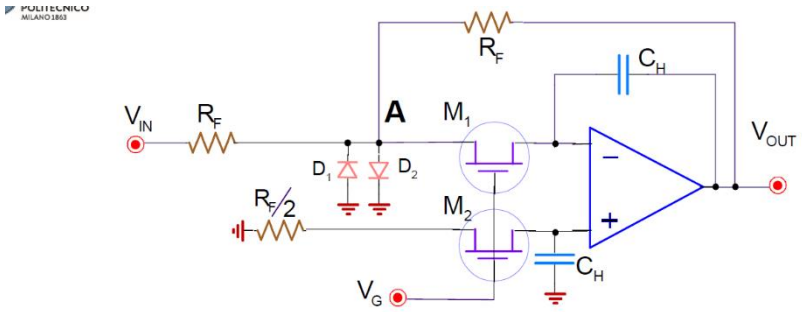


A final issue; opamp requires Ibias that flows in the two branches, and they see R||R on the – termina, and GND on the other pin. This will cause an error in the output, so we need to compensate the bias current. To do so, we introduce a resistor equal to the impedance we see on the other path.



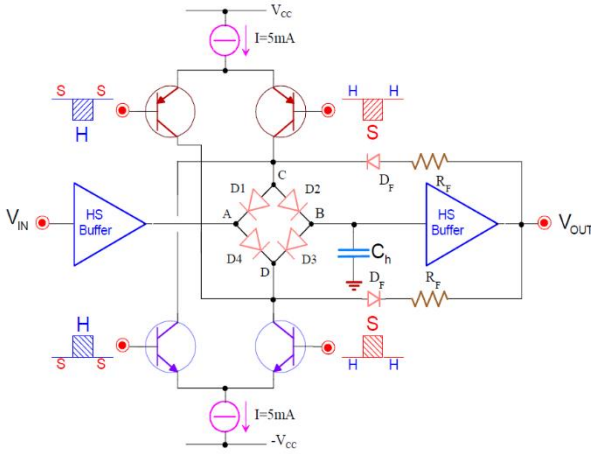
The final implementation is the one below.

What happens in terms of droop?
We have two capacitors, and the two Ibias will cause droop on the two capacitors. But does this circuit suffer from a double intensity droop error with respect to the classical S&H configuration? Does droop compensate?



- V_G can have simple low-voltage CMOS levels (0-3.3V), independent of V_{IN}
- the *Aperture-Induced non-linearity* can be drastically reduced

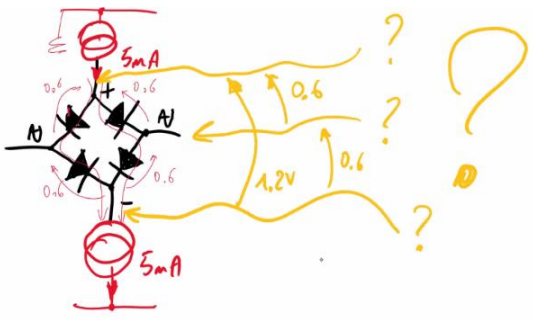
SPEED UP THE SWITCH



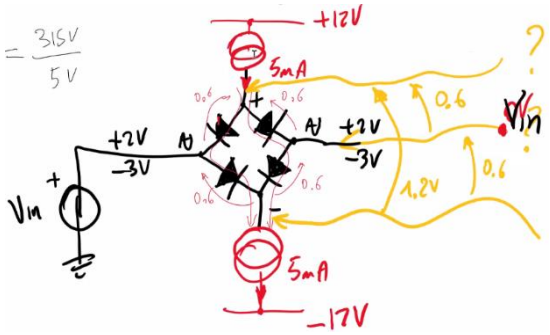
- fast diodes
- criss-cross

- poor accuracy
- timing issues to solve

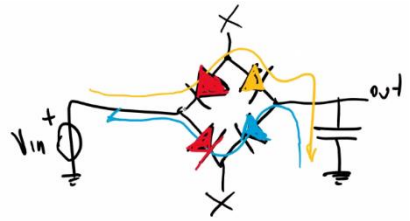
This is a real schematic. We are designing a new S&H circuit. I want to enter with a signal Vin, put a switch and then place the capacitor and the buffer; I want to improve the switch, not using a slow problematic MOS transistor. We could use a BJT transistor, but it is not a good switch because its Ron is not symmetric for positive and negative voltages.
A diode bridge can be used as a switch. This configuration was used after a transformer to convert an AC signal into a DC one. The diode bridge converts the negative peaks into positive one, and then we use a capacitor to have a DC value.



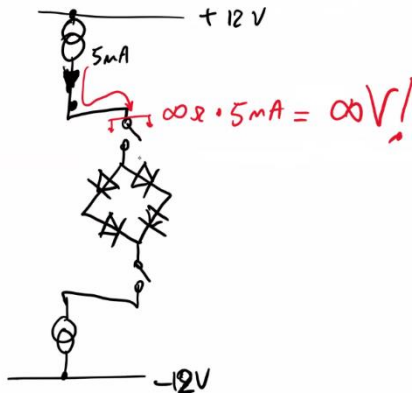
Moreover, if the connect the bridge as in the image, and we pump current in the bridge, it will split and if we have a current source also on the other side and the two currents are matched, the one below drinks the current. So all the diodes are on and hence we have the following voltages. The voltages there (?) will be set by the r0 of the current generator, but I don't care.
The idea is that if we apply Vin to the other node of the diode bridge, the output will be Vin. If now we remove the current generators, the diodes cannot stay on.

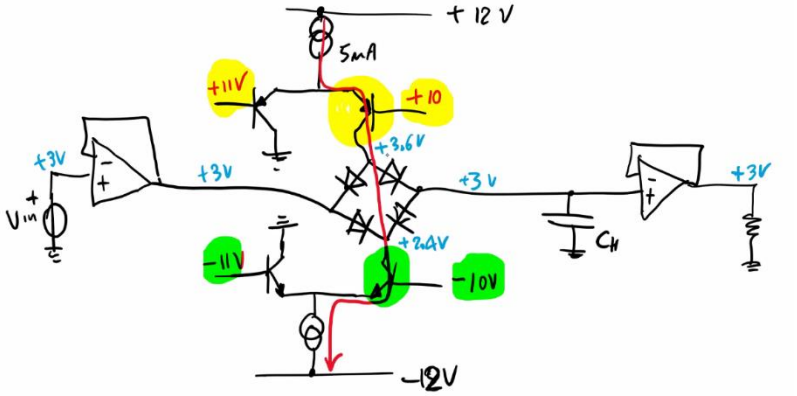


If the diodes are open we have no longer control on the output with Vin voltage, because diodes are counteracting, so neither I charge the capacitor nor it discharges (red diodes in the image are stopping the current flow).



To stop the current generator to provide a current is to put a switch in parallel to the current generator so that when it is closed the current recirculates, but we have that in this case the PS voltage is put to the upper node of the bridge and we break everything. To stop the current flow we hence may add a transistor. But this is not feasible because if the switch is open, the current it pumped into a node with infinite impedance, which gives infinite voltage at that node, and the switch breaks.



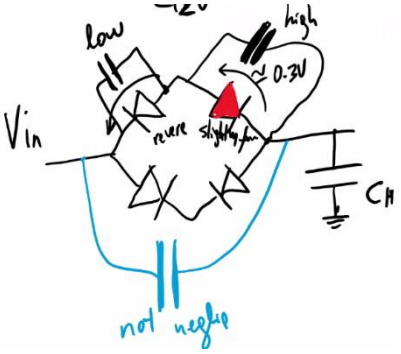


So what we should do is to use a differential pair (pnp and npn transistors). If the bias is the one on the right image, the right transistors are on. So we can use a buffer to apply a Vin voltage, and the output goes to a capacitor and eventually an output buffer to provide a current to the load without discharging Ch.

If we invert the bias voltages on the bases of the transistors, the current is in the other branch and the bridge is off.

Of course, the driving of these transistors is not so easy with these high values of voltage.
The basic advantage of this configuration is that the diode bridge can be turned on and off very quickly because made just by pn junction, so we don't have to charge a polysilicon gate as in the MOSFET to create the charge in the channel.

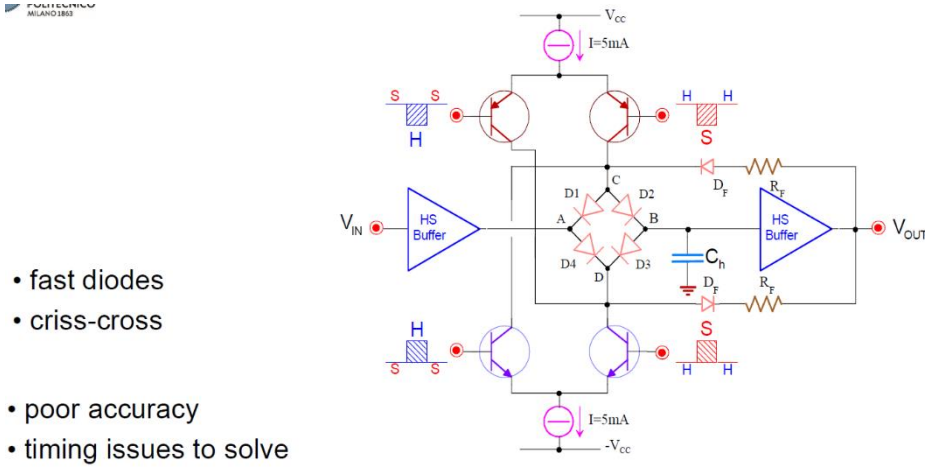
This is hence a very fast configuration where we can use high values of current. So we can provide a switch where the r_on, impedance from input to the output when the bridge is conducting, is (1/gm + 1/gm)||(1/gm + 1/gm).
A minor disadvantage is the signal feedthrough. In the off configuration, transistors are off but, if Vin tries to go positive, e.g. 5V, and the voltage on the capacitor is 3V, one diode is on slightly and the other off, both on the top and bottom paths. If one diode is slightly on, then the voltage drop across is almost 0.3V (not forward bias 0.6V because it is not conducting current). So we have the input/output capacitance of the diodes that is not nihil.



To reduce signal feedthrough across the bridge we need to decrease the parasitic capacitance. To do so, we need to increase the reverse bias. But since Vin can be at most 5V in this config, and the voltage across the capacitor -5V, we have at the most 10V of reverse bias on the diode.

In fact, given a diode, if we increase the reverse bias, the capacitance between the anode and the cathode decreases.

In the current configuration, current in the OFF stage is drunk from ground or place to ground. So we change the configuration as below.



- fast diodes
- criss-cross

- poor accuracy
- timing issues to solve

In the off configuration we don't ground the currents, but I pump them in D in off config, and pump in C. Thus the current cannot go upward in D point or in the BJT at the bottom, and the same for the C point. So I introduce a path, resistor and diode, where the current can come from. The diode is introduced to have only one direction of flow for the current. And the same current that is drunk from the top node is pumped in the bottom node. Hence the two currents compensate and the output buffer has no current to provide.

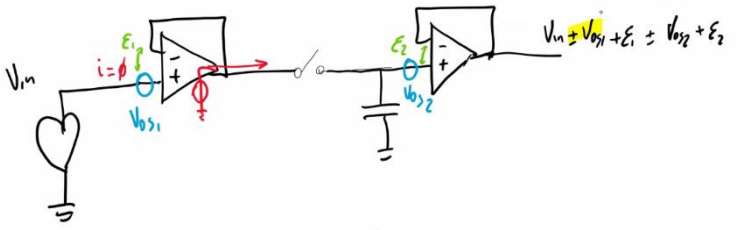
These currents will generate a voltage across the resistors Rf (5V) and we have node C that goes below (5.6V below 3V) and D goes above (5.6V above 3V).
Hence the reverse bias is so high that the parasitic capacitance is low and the signal feedthrough is drastically decreased.

So far the, in the ON configuration, 5mA are flowing in the red and purple BJT, so there is no current in the buffer and no current in the capacitor Ch. What happens if there is a mismatch between the two current generators? A possible answer is that the capacitor keeps charging up until it saturates. Another possibility is that the current flows in the left buffer. Or, the right answer, is that there is the loop of diodes and thanks to the translinear principle we know the current in the diode bridge and for sure the difference will flow in the left buffer. But if so, voltages of the diodes will mismatch, so the voltage on Ch won't be Vin, but with an error due to the mismatch. So the mismatch in the current values causes a offset.

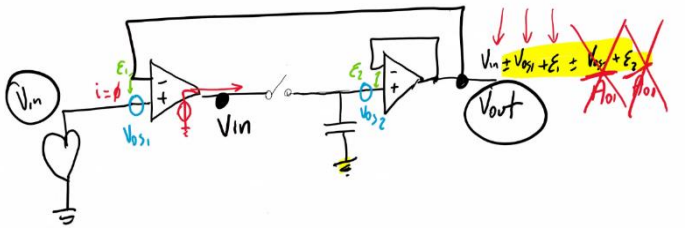
The other possibility is that there is a mismatch in the timing of the transistor, for instance when we turn on the red BJT and off the purple one.

FEEDBACK S&H CONFIGURATION

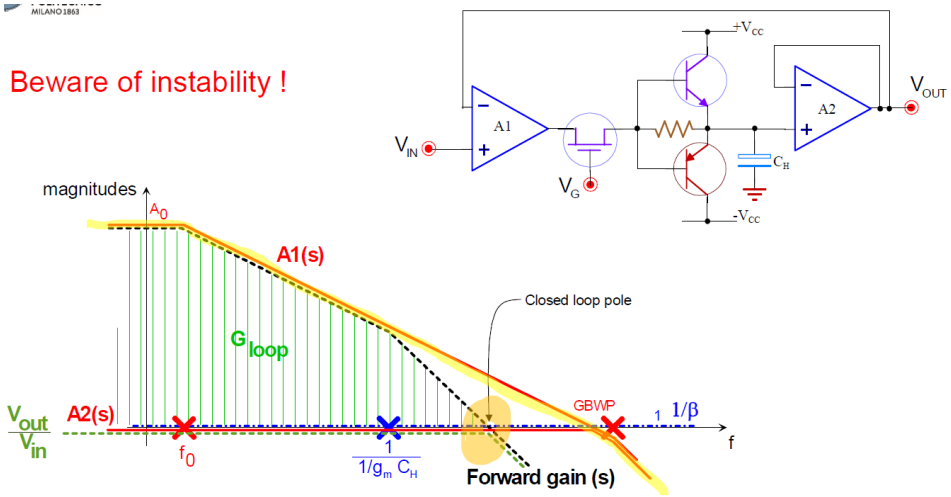
So far we have seen configuration without a feedback. But since we have introduced buffers, we have the Vos error and the epsilon error. Hence we have an error between the input and the output.



We want Vout to be Vin, so we have a global feedback. Thanks to this connection, Vos1 and epsilon1 are reduced by A0, so negrligible.



Moreover, if we want r_on to be low, we can put a buffer in series with the switch to have a reduced impedance, as in the configuration below.

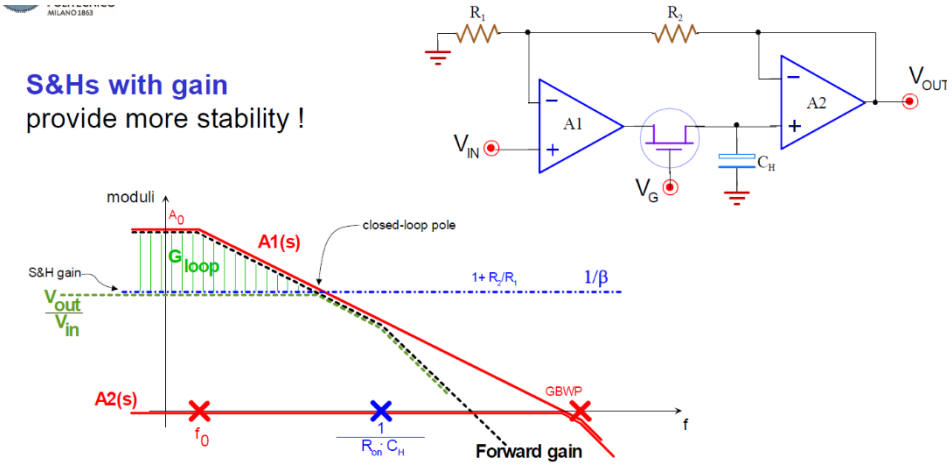


The buffer is just a follower, no need to introduce an opamp. Of course, the follower has a voltage drop of 0.6V across it, but thanks to the feedback, the output is still Vin. Pnp transistor and npn transistor are needed to provide both the directions for the current, otherwise the capacitor will just charge or discharge.

The resistor is needed because if the switch is close, without it, the loop is ok. If the switch is open, now the base of the transistors is floating, so we run the risk that one of the two transistor is on and we charge or discharge the capacitor, corrupting the output voltage.

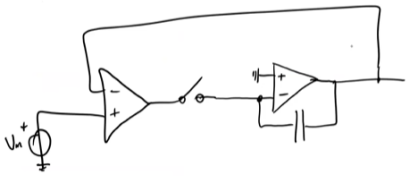
Now the Ron I see is not the Ron of the MOSFET, but it is the 1/gm of the BJT that is on plus Ron of the mosfet divided by beta of the BJT. So I also reduce the impedance.
Unfortunately, however, this circuit can be unstable. The first opamp is compensated; then we study the beta network. The beta has a pole due to Ch and 1/gm. As A(s) we are considering the forward path, so opamp, follower and opamp, that is the forward gain (s). We notice that the closure angle between 1/beta and forward gain (s) is 40-40, so the circuit is unstable.

In fact, every time we close the switch, if the circuit is unstable, we have some ringing when approaching the Vin signal during the sampling phase, so we need a higher settling time.
To regain stability we cannot add a capacitor in the global feedback, because we don't have a resistor, so what we can do is to introduce resistors as below.

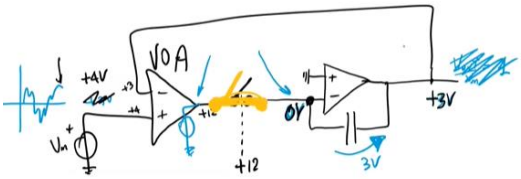


In this case we have the same stage as before, but now the beta is lower and so 1/beta is higher. This S&H is stable and it introduces also amplification. However, the issue here is that if we want to store the full voltage on Ch (e.g. 5V), Vg should be very high, even 10V. Instead, the simplify driving configuration seen previously was lovely because we can drive the transistors with the logic voltages, because of VG.

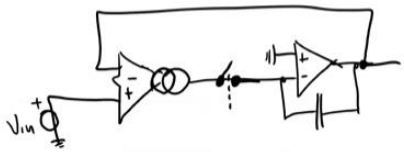
So better to use another configuration. Better not to use a buffer as a final stage but something like below. It's not an integrator.

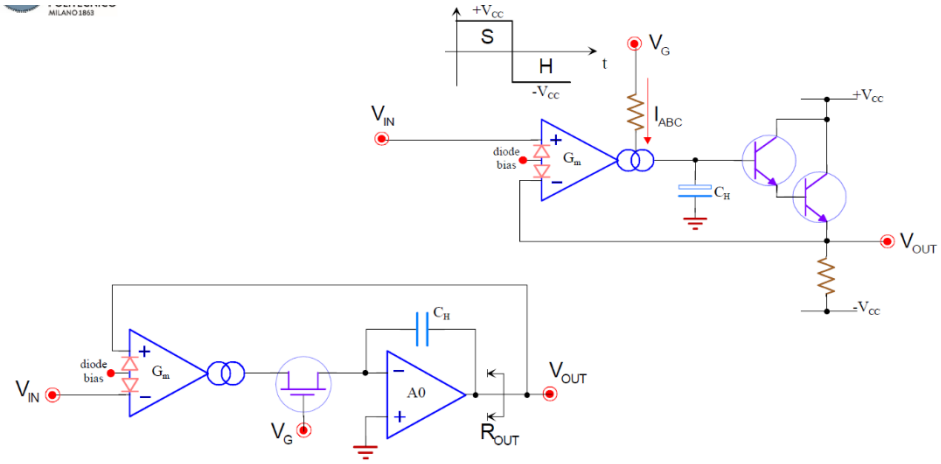


When the switch is closed, the gain is 1 and capacitor charges to 3V. If then the switch gets open, the capacitor provides 3V to the output even if Vin changes.
The problem is that if the first amplifier is a VOA and we have e.g. Vin = 4V, the output is at 3V, so the epsilon in input is 1V and the output of the VOA saturates to the PS, that is e.g. 12V. Now the switch is between 12V and 0V (VG), so even if we apply 12V on the gate we need a huge amount of time to close the switch.



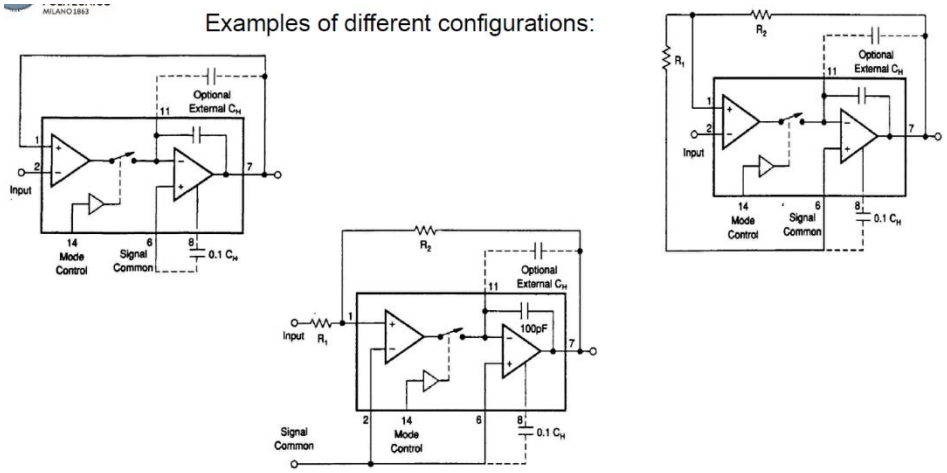
Hence is better not to use a VOA but an OTA followed by an opamp connected as an integrator.





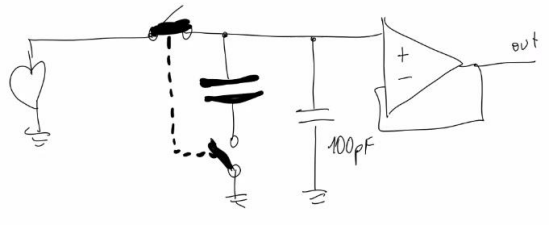
Many opamp manufacture developed an IC with the OTA, the switch, the second opamp, the capacitor.

Example of IC connections



Ch is very important in S&H and it must be very large to hold the voltage for a long period, but if it is big, the t_acqu is high. So we want a Ch that is high value when we need it and low value when we don't need it.
During the sampling phase I want Ch to be low value, so that we are fast in tracking the signal. Vice versa, in the hold phase I want Ch to be very large so that the voltage stored on the capacitor doesn't discharge due to droop.

To do so, for sampling when the switch is close I want a small Ch, so I could include a very big capacitor that I leave open during the sample phase and I use another one that is smaller. The two switches are connected one with the other.



Then when I open the switch, the bigger capacitor is attached and the two capacitors are in parallel. But this configuration is not working, because when the switch is closed we store the voltage on the tiny capacitor, but the other one is discharged. Then when we connect it, we place two capacitors in parallel, and the big capacitance was charged to 0V. Hence we need to equalize the charge between the two, so the voltage across the bigger capacitor will be like below (assuming 3V stored on the small capacitance).

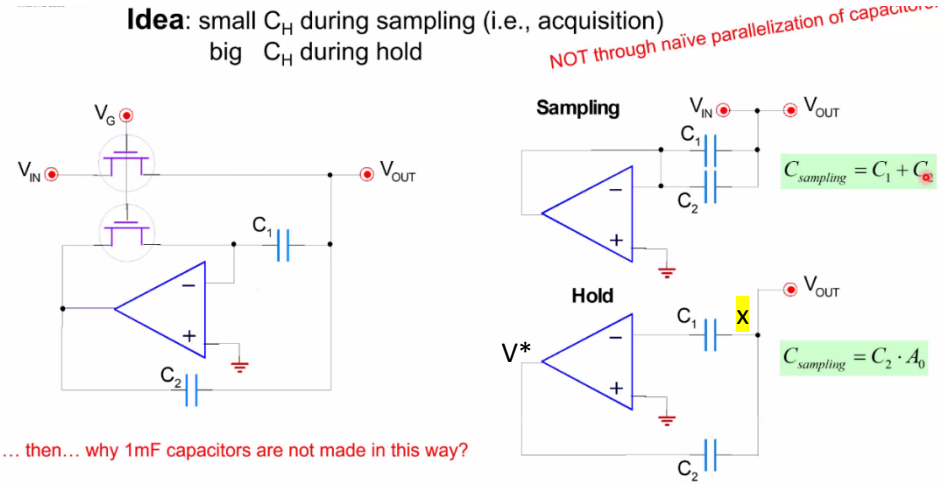
$$C = \frac{Q}{V}$$

$$\begin{cases} Q = 0{,}1n \cdot 3V \\ Q = 10n \cdot V^* \end{cases}$$

$$V^* = 3V \cdot \frac{0.1n}{10n} = 30mV$$

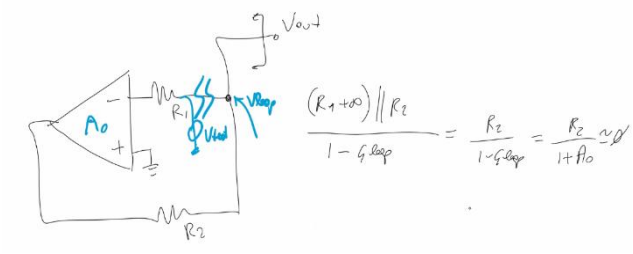
We could introduce an amplification in the following opamp, but in doing so also the offset, error and noise get amplified, so it is not a good solution.

The proper configuration to be exploited is the following.
C1 goes to VG, and when the switches are closed, the capacitances are in parallel during sampling, and during hold phase the opamp is an open loop opamp, not a buffer. The overall capacitance we see is not C1, because in series with infinite (terminal), so we see C2 multiplied by the internal gain of the amplifier.

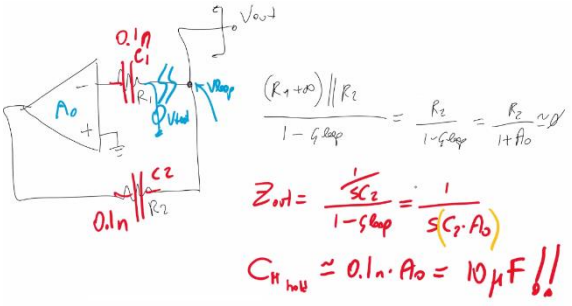


If C2 discharges because e.g. an ADC drinks the current, then node x decreases, but if so, - terminal decreases and output increases, so the loop keeps the voltage fixed. C1 is not acting because no current flows in it, it was charged and it remains charged as a battery.

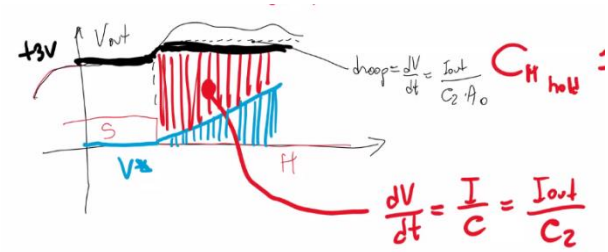
The impedance we see in one node of the loop is then the stupid impedance divided by 1 – Gloop, and at node x we see, in the case of resistors, is:



In case of capacitors we have the following.



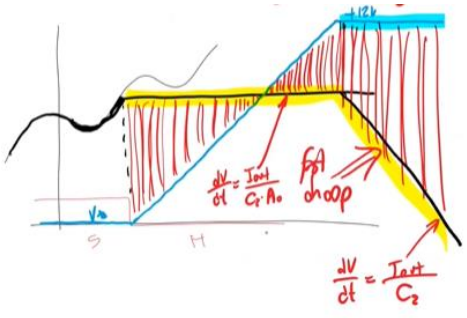
When the switch is closed, V* is at VG, so 0V, but as soon as the switch is open the voltage across the capacitor discharges, but V* increases so that the Vout is almost flat during the hold phase. Of course there will be a droop, but it is reduced by the feedback.



This configuration works because during sampling we charge both capacitors to 3V. So when we open the switch we don't have a capacitor charged and the other one not so we have to wait for the other capacitor to charge due to redistribution of charge, because now both are charged.
Since in many applications we need large capacitors, why don't we place C2 big, e.g. 1uF and A0 = 1000 to have an equivalent huge capacitor?
The discharge of Vout is very small, but there is something in this circuit that at a certain point makes the circuit not working.

If we look at node x, it is not discharging, and if the V* reaches the power supply value, the configuration stops working.

Let's see what happens. We have Vin and the S&H signal; when V* reaches 12V (PS) and saturates, the capacitor discharges (red area) and then it charges in the opposite way and in the end the capacitor keeps discharging. So the Vout after some time collapses, so we have a very fast droop.

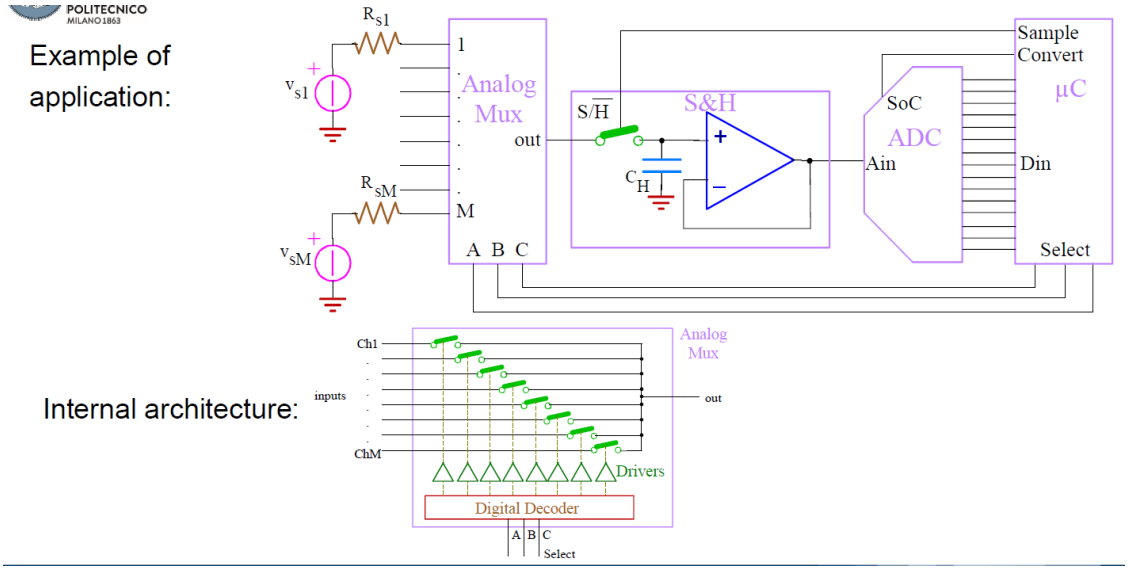


Hence at a certain time we need to close the switch and reapply sampling before we reach saturation of V*. Otherwise, C2 discharges very quickly.

MUX AND DIGPOT

ANALOG MULTIPLEXER

An analog mux is a multiswitch. We have different input signals and we want to choose which signal to send to the S&H. We use a digital decoder to select which is the channel to sample.

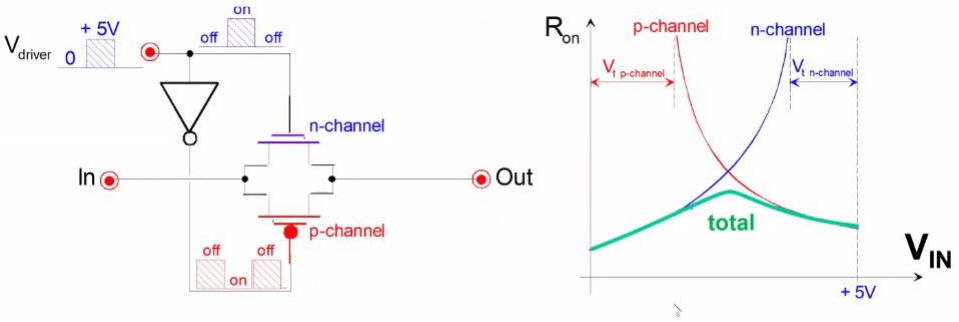


Pass transistor

Switches must be analog switches, so transistors, but one transistor is not enough, because the input is analog and it can be positive or negative, so we should use very high voltages to drive the switch if we are controlling the MUX with digital levels, so we need charge pump and voltage translator. Hence the best solution is to use a pair of n-channel and p-channel transistors driven with opposite phase. A 0 signal keeps off the n-channel and on the p-channel if Vin is sufficiently high.
In the n-channel transistor, if Vin is high Ron increases. If Vin increases and we keep Vgate fixed, the n-channel transistor is less closed and Ron increases.

If Vin reaches 5V (gate voltage) – Vt, the Ron becomes infinite. Vice versa, if we apply 0V to the p-mos, while Vin is low the transistor is off, but if it increases the Vin the Ron decreases.

Analog switch implementation:



Also useful to compensate charge injection!

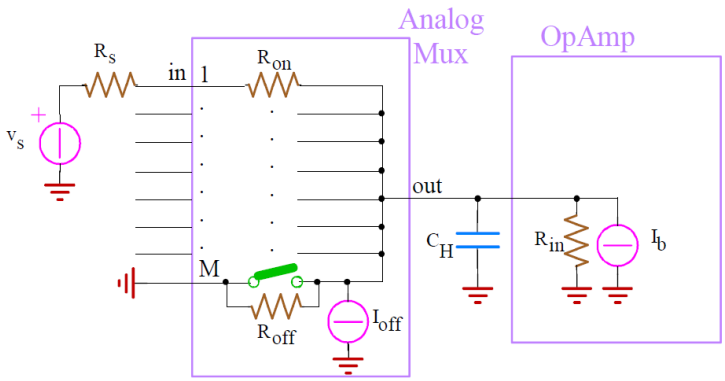
If we put them in parallel, we have to compute the parallel between the two Ron, and we get a value that is almost constant. Moreover, this configuration allows to compensate the charge injection, since we have the Cgd of the two transistors. But when we apply a swing to turn the nMOS from off to on, we have charge injection. But we apply a falling edge to have the nMOS of and a rising edge to have the pMOS

off, so the falling edge at the gate of the n channel and the rising edge at the gate of the p channel cause opposite charge injection to the output node, so they compensate each other.

Analog multiplexer's errors

Sources of errors:

R_{on}	10Ω÷10kΩ
R_{off}	>10MΩ
I_{off}	<100nA
T_{on} T_{off}	<100ns



$$\Delta V = \pm v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}} \pm \left(I_{off} \cdot (M+1) + I_B\right) \cdot \left\{(R_s + R_{on}) \| R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}$$

When the transistor is closed, it will show its Ron, that usually its hundreds of ohms. If the PS of the chip is very low or the temperature is too high or too low, maybe Ron can increase. Similarly for Roff, which should be infinite, but it is in the range of MOhms, so eventually also Roff can decrease.
Hence **each pass transistor has a leakage current from drain to the bulk**, so from the drain towards the ground, for each switch, and this in both sides of the transistor. But one side is attached to the source, and the right side we have the S&H, so we care about the leakages in the S&H part (they act on Ch), because the leakages on the left causes something in the source, that is not relevant when the switch is open.

Moreover, after the S&H circuit we have and opamp or the ADC.
Let's compute the error and to consider one switch close and all the other opens, at the output of the MUX. I see Rs in series with Ron and then a parallel of Roff to ground. In fact, if is a pin is not connected, it is connected to ground, or to another source, so eventually I would see Rs + Roff towards ground, that is still Roff because it is much bigger than Rs. I expect the voltage drop across the MUX to be 0, but it won't be zero.
In the equation I should put the highest possible Rs and Ron, and at the denominator the lowest Roff, Rin, Ron and Rs.

The second contribution is the one related to Ioff, because when we open the switches Ioff drinks current from the resistances. There are M Ioff and also the Ib bias current. Of course I don't care about the Ioff on the left of the switches because they die toward ground if the switches are open, but if it is closed I have to consider it, hence the M+1. Then the impedance I see from the node of the capacitor is a parallel

Example: ADG508A mux, Analog Devices

M=8 channels	R_{on}=400Ω R_{off}=10MΩ I_{off}=100nA (@ 125°C)
v_s=±15V	Rs=1kΩ
OpAmp	R_{in}=500kΩ I_b=0.5μA

we get a maximum error of ΔV=±56mV±0.98mV±0.7mV

the first term is equal to ½LSB for a 10bit ADC
the sum of the other two contributions is equal to ½LSB for a 15bit ADC

between all the resistances. In the worst case all the resistances must be the highest possible value, as well as for Ioff.

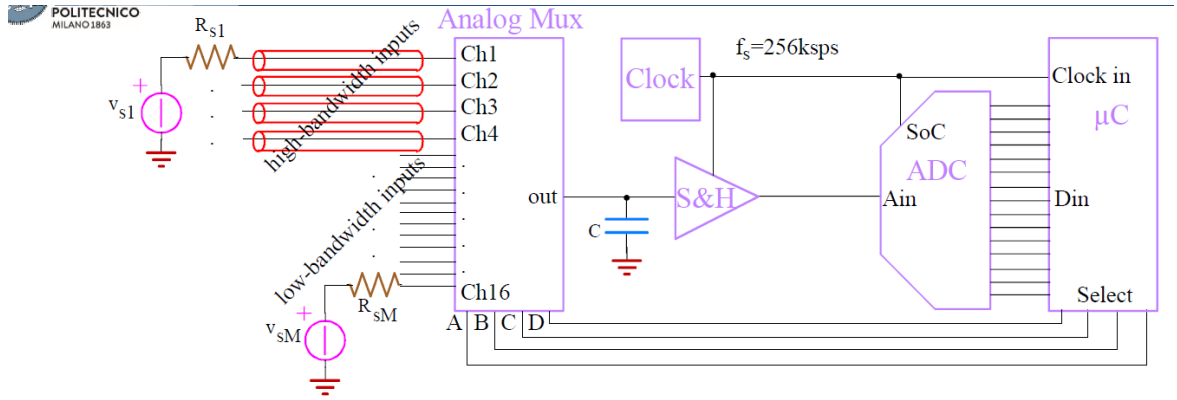
If the error is less than ½ LSB we are ok, but if not, we are converting the value in the next bit interval. Hence the error of the MUX is large or not depending on the following ADC.

Timings

Let's suppose to have 16 channels to acquire, and some inputs are high BW, other low BW, and we have just one ADC, uC and S&H.

We select a channel to sample, e.g. the first one. We acquire the first input, we store the value on the capacitor, we close the switch, we wait to track the signal and then we open it in the hold phase, we convert it, we wait for the ADC result. After this the uC reads the digital bus and changes the channel.
In order to acquire 16 channels we need to dedicate one slot per channel. The uC decides when to change the channel, apply a pulse to the S&H and to the ADC. Or we can have a master clock that commands the S&H and the ADC, that then tells to the uC if the data is ready or not.

Of course, the master clock must be correctly chosen according also to the conversion frequency of the ADC (sps = samples per seconds). A 256 ksps corresponds to 3.9 us, so the uC will read the data and provide a new selection every 4 us.



Example: ADC with f_S=256ksps
every channel scanned at 16ksps ($1/f_S$=3.9µs and not 1/16ksps!)
therefore the maximum admitted bandwidth of each channel is less than 8kHz

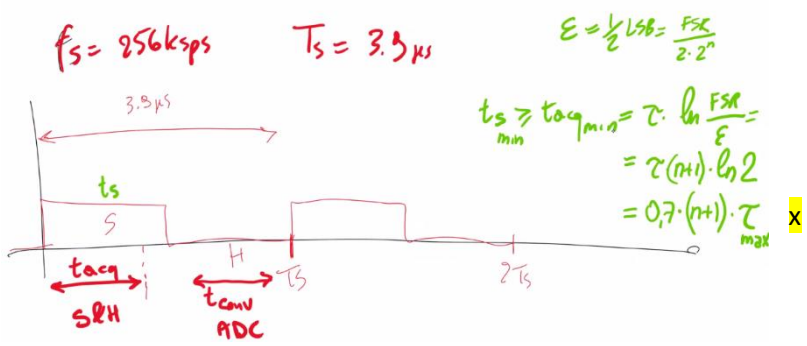
What about if M=16, but only the first four channels with a 20kHz bandwidth? Here is the solution:
Ch1, Ch2, Ch3, Ch4, **Ch5**, Ch1, Ch2, Ch3, Ch4, **Ch6**, Ch1, Ch2, Ch3, Ch4, **Ch7**…
for the first 4 channels, the sampling becomes $f_S/5$=51.2ksps, while the other 12 have $f_S/60$=4.2ksps

If we want to acquire 16 channels it requires 16*3.9 us, that is 1/16 * 256 ksps. So every channel is sampled at 16 ksps. But then it means that fmax for each channel should be lower than half that value, not 256 ksps.
But this approach is not good because it is too slow for the channel attached to high BW signals and too fast for channels attached to low BW signals.

So we should choose another timing. What we could do is to read channels 1, 2, 3, 4, 5 → 1, 2, 3, 4, 5, 6 → 1, 2, 3, 4, 5, 6, 7 so that we scan the first four channels more often, and the other less often.

It takes 5 readings for channels from 1 to 4 and then 12 conversions to read channel 5 again. It takes 12 times 5 times 4 us, so once every 60 conversions for the last 4 channels.

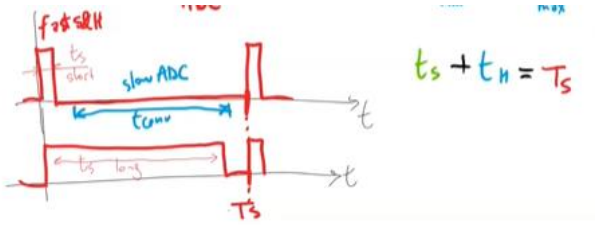
Of course, as for the S&H I relaxed the BW of the signal, but since the S&H is running at 256 ksps, every 4 us the S&H must close and open. So since fs = 256 ksps, it means that in my signal I need to open and close the switch every Ts = 3.9 us. When the switch is close in the sampling phase, the duration to be provided has to be at least the t_acqu, sufficiently long, and in the hold phase I should have at least the t_conv of the ADC (conversion time of the ADC).



So t_sampl,min > t_acq,min. t_acq,min is tau multiplied by the logarithm above. The error epsilon is half the LSB, so FSR/(2*2^n). tau is Ron*C. The Ron is not the Ron of just the MUX, but the total Ron, so Rs + Ron,mux.
We have to compute t_acq,min in the worst case scenario, when tau is maximum. In fact, by minimum value of t_acq we refer to the minimum reliable one, so the longer worst acquisition time.

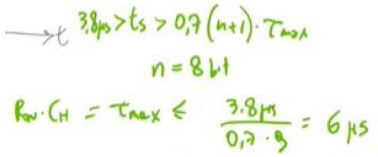
Another consideration. The minimum duration of t_hold should be longer than t_conv of the ADC, and the worst case scenario of t_conv is when it is maximum, at it can be 10ns in the case of flash ADC, 10us in case of SAR ADC, 10ms for dual-slope ADC and so on.
Moreover, t_sampl + t_hold = Ts.

So depending on the S&H circuit and on the ADC, we could use a very fast S&H so that t_sampl is very short, because maybe the ADC is slow and t_conv is very long, or we can have the opposite.

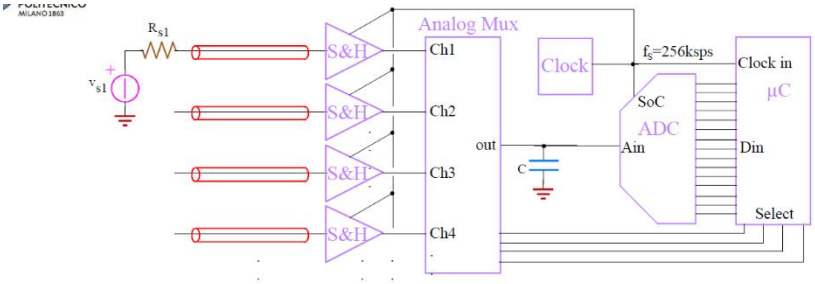


Coming back to the circuit of the slide, the S&H must be able to close and open within 4us. So at at least the ADC should convert within 4us, but if the ADC converts in 3us, we have only 0.9us left for S&H. so better to use a flash ADC that converts in 10ns so that we have 4 us for S&H.

The tau_max we have to achieve can be computed from equation x.



There are some applications where we cannot use this technique, because in this technique we have many different signals and we perform multiplexing before the S&H. Instead, there may be other applications where the signals must be acquired on the same time instant. In this case, we need to have a S&H for each signal, and then we multiplex them to the ADC.



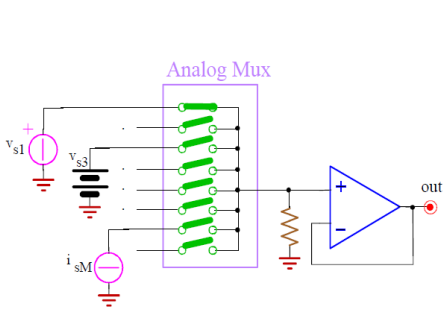
Example: ADC with f_S=256ksps
4 "fast" and 12 "slow" input channels

With just one S&H after the mux, each "fast" channel should have available a time of $5/f_S$=19.5µs

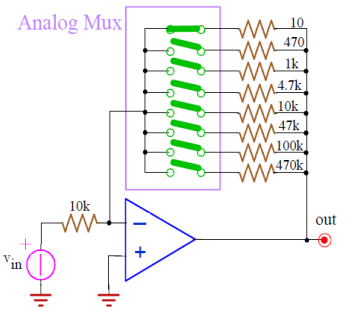
With one S&H at each channel input, each channel should have $16/f_S$=62.5µs

The master clock should open all the switches of the S&H and open them synchronously. Then it opens and closes the switches of the MUX alternatively.

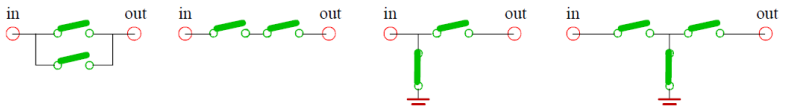
Improvements



ALD4201 (**Break before Make**)



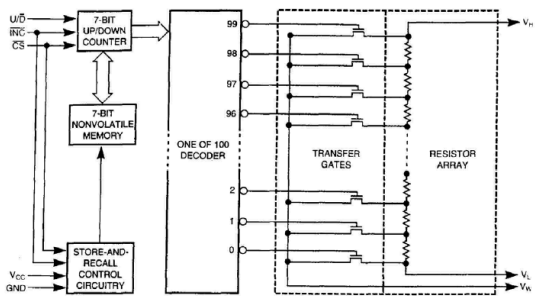
ALD4202 (**Make before Break**)



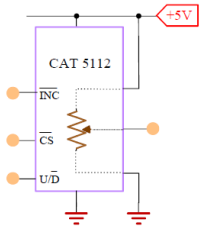
Make before break is needed when the MUX is in feedback, otherwise the output may saturate.

DIGIPOT

Block diagram:



Example:

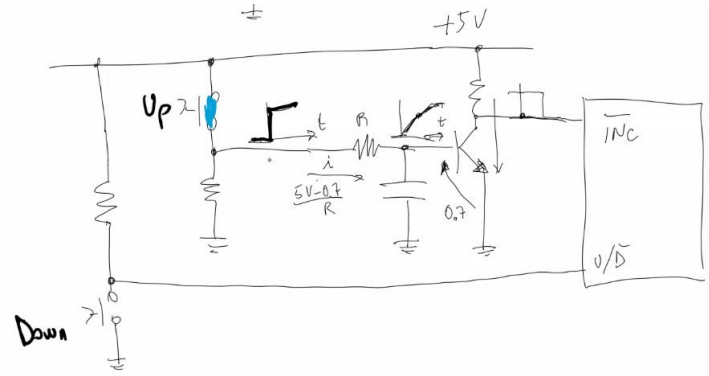


Model	Steps	Resistance (kΩ)	Configuration	Non-volatile	Packages	Interface	Operating power	Features	Price (1000)
AD8402	256	10, 50, 100	Dual	No	DIP-14, SO-14	Three-wire	2.7 to 5.5V, 5 µA	Full ac specs, nA shutdown current	$1.66
AD8403	256	10, 50, 100	Quad	No	SOL-24	Three-wire	2.7 to 5.5V, 5 µA	Full ac specs, nA shutdown current	$2.51
DS1267	256	10, 50, 100	Dual	No	DIP-14 SO-16, TSSOP-20	Three-wire	5 or ±5V, 650 µA	Stackable wipers for 512-step resolution	$2.45
DS1867	256	10, 50, 100	Dual	Yes	DIP-14, SO-16, TSSOP-20	Three-wire	5 or ±5V, 650 µA	Nonvolatile version of DS1267	$3.14
DS1802	64	50	Dual	No	DIP-20, SO-20, TSSOP-20	Three-wire and pushbutton	3 or 5V, 2 mA	Log taper, mute, audio specs	$2.56

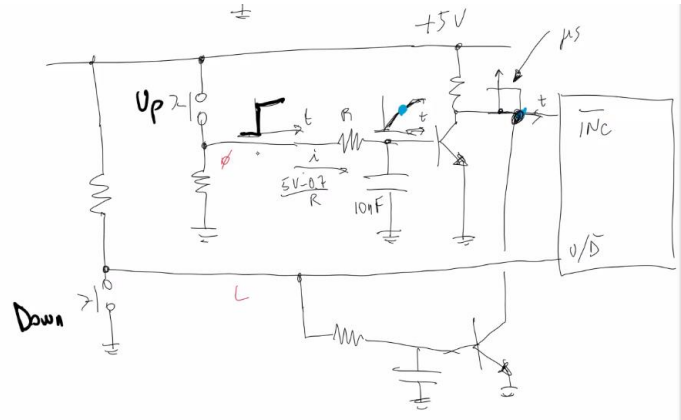
The value of the potentiometer is changed by providing pulses to the component. It is a set of resistors one in series with the other, then we have an analog mux that can be modelled as a set of switches. To change the switch we provide the digital address of the switch or, to simplify the connections, we can use 2 or 3 pins, and inside the digipot there is a counter and we decide if to count up or down.

Some digipot can be non-volatile, which means that if we switch off the PS the component remembers the position of the cursor.

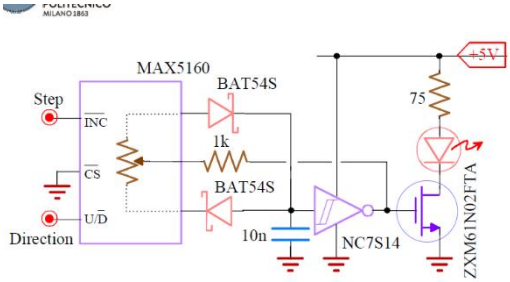
If we provide a pulse, the component increments, after having chosen the direction of the counter. One push button can be used to select the direction of the counting, the other to give the pulses. We would like one push button to go up and one to go down.
To do so, when I push the down counting button I also want to change the INC pin. So I can introduce a certain delay, but we cannot simply attach the output of a RC circuit, because we could cause a shortcircuit with the up counting button. So we can use a transistor.

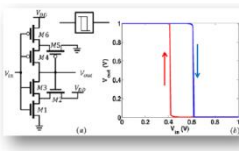


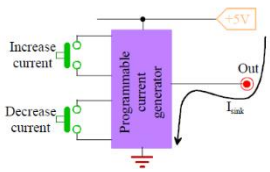
We need then to perform an OR connection to the INC pin to have the up and down push buttons.



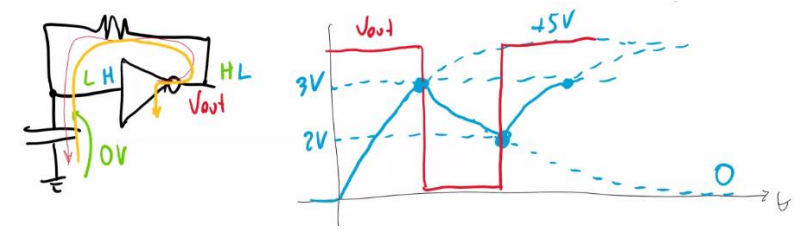
Let's consider the following circuit.



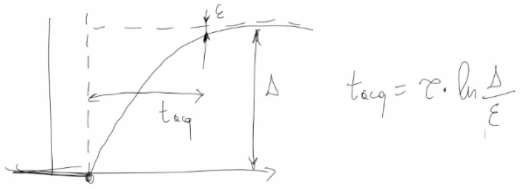




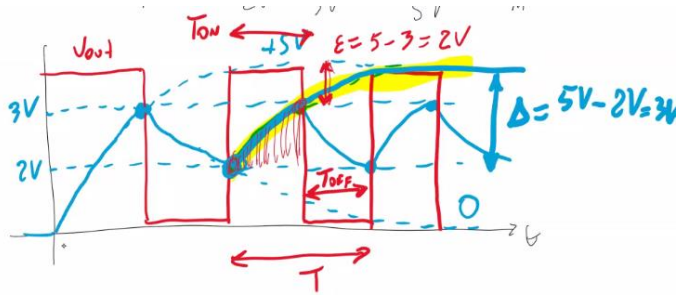
We have a NOT port that is a Schmitt trigger gate, with an hysterical behaviour (upper right image). The inverter has two logic levels, 2V and 3V if it is biased at 5V. Above 3V we have the commutation to the high level. So we have an inverter whose output feeds the input with a capacitor to ground. It is hence a digital oscillator.



We can use a similar equation used during the S&H studying, the one on t_acq.



We can use it because we have an exponential curve also in the oscillator, and the delta we need to reach in the increasing exponential is delta = 5V – 2V, and epsilon is 5 – 3 = 2V. So we can properly compute Ton. The same for Toff, but it changes delta and epsilon.



$T_{ON} = \tau \cdot \ln \frac{V_{dd} - V_L}{V_{dd} - V_{TH}} = \tau \cdot \ln \frac{3}{2} = \tau \cdot 40\%$

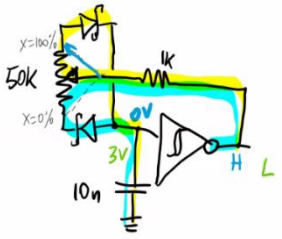
$T_{OFF} = \tau \cdot \ln \frac{V_{TH} - 0V}{V_{TL}} = \tau \cdot \ln \frac{3}{2} = \tau \cdot 40\%$

$T_{osc} = 2 \cdot 40\% \tau = 0{,}8\tau$

$f_{osc} \simeq \frac{1}{\tau} \cdot 1.25$

If now we look at the circuit, we have the NOT gate whose output feeds a resistor that then feeds a digpot, which feeds then two diodes and the input of the NOT gate through a capacitor. To reduce as much as possible the voltage drop across the diodes, we are using Schottky diodes, not standard ones. Its voltage drop is 0.2V

If at the beginning the C is full discharged, the output of the NOT will be high, and the path that charges the capacitor is the yellow one. Once the capacitor has been charged, the output will move to the low level, and the discharge will be through the blue path.



$R_{charge} = 1k + 50k(1-X)$

$R_{discharge} = 1k + 50k \cdot X$

$T_{ON} = \tau_{charge} \cdot \ln \left(\frac{V_{dd} - 0.2 - 0}{V_{dd} - 3} \right)$

$= C \cdot R_{charg} \cdot \left(\ln \frac{2.8}{2} \right) 34\%$

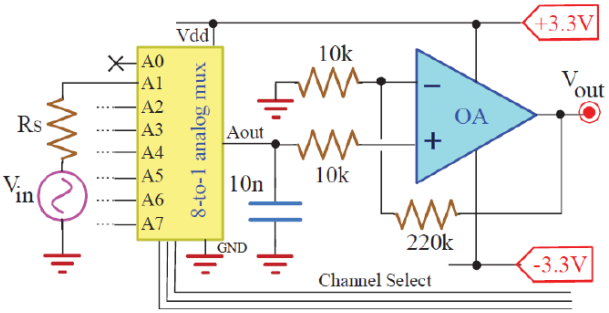
$T_{OFF} = C \cdot R_{discharge} \cdot \left(\ln \frac{2.8}{2} \right)$

The total Tosc = Ton + Toff is a constant value. Hence oscillation frequency won't depend on the position of the cursor but on the total value of the potentiometer.

$$T_{osc} = T_{on} + T_{off} = 34\% \cdot C \cdot [R_{charge} + R_{discharge}]$$
$$= 34\% \, C \, [1k + 50k]$$

By changing the cursor position, Tosc remain constant but I'm changing the duty cycle.

Exercise 1



Seven inputs with R_s=100Ω-1kΩ. OpAmp with sourcing I_B=1nA and V_{OS}=0.2mV. Mux with R_{on}=5-50Ω, R_{off}=2MΩ-20MΩ and I_{leak}=5nA.

a) Compute **sampling** and **hold** times for 12 bit resolution and $V_{in,max}$=±50mV and specify if they are max or min values.

b) Compute all static errors and properly add them to compute the total output error in LSBs.

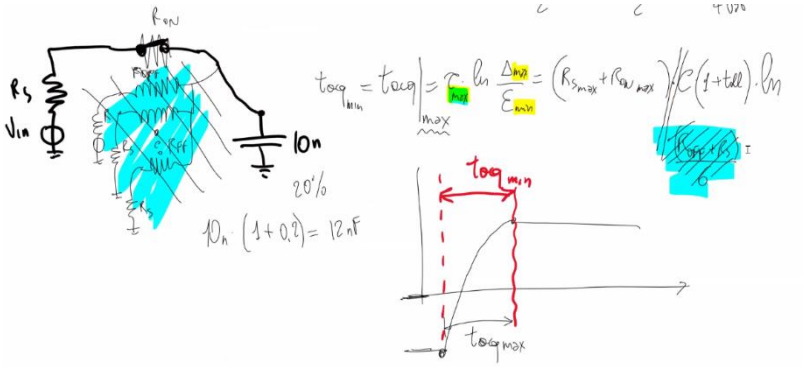
The capacitor acts as a S&H circuit, so we need a configuration where the analog MUX can be connected to one input but eventually to have also all the switches open. Because of this, pin A0 is left floating because we can never have all the pins of the MUX open, one must always be closed (even 0000 is the bits code for a channel to be closed).

Point a)

Once we have acquired the signal, the signal gets stored on C and then goes in the OA that is connected as a standard non-inverting configuration, so the G = 1+R2/R1 = +23. Vout_max = Vin_max * G = 50mV*23 = 1.15V.

The input signal can be either positive or negative, the FSR of the opamp is 6.6V and Vout can be both positive and negative, so Vout = +-1.15V.
As for the LSB of the ADC: LSB = 6.6V/2^12 = 1.6mV.

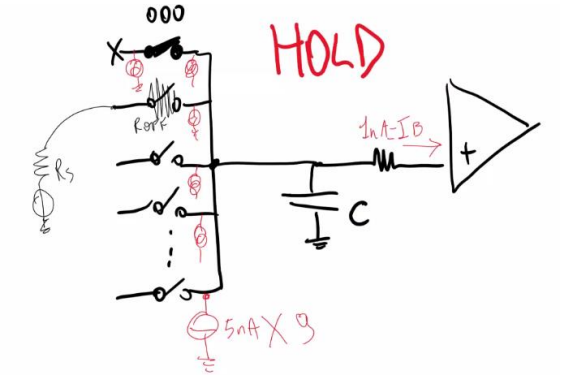
When we want to acquire a signal we need to wait the sampling time, when the switch is close and it will show its Ron. The switch is connected to a source and the signal of the source goes to the 10nF capacitor. To quote t_acq we need to think at the worst scenario, so when it is maximal.
t_acq,max is the minimum acquisition time I need to wait. Of course, also the capacitor must be considered with its own tolerance to have the maximal value. Moreover, in the equation all the other Roff are not considered, and these Roff are in series with the Rs of different sources, so the Rs,max + Ron,max should go in parallel with the Rs + Roff, but since this parallel would decrease the overall

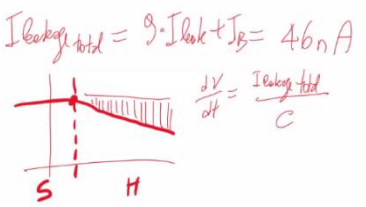


resistance, let's neglect the Roff. The delta_max to be considered is the maximum swing the output can experience, so the FSR (if not indicated otherwise). Instead, the epsilon is ½ LSB of the following ADC.

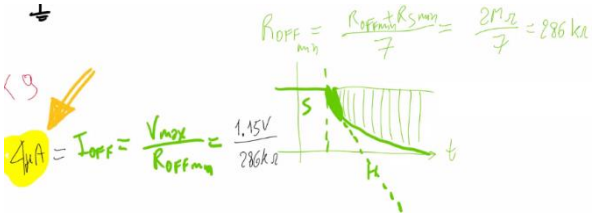
If we want to be less conservative, the maximum possible output swing will be due to the maximal Vin the user will use, so it can be reduced the delta_max term. In this case delta_max = 2.3V
In the end, t_acq,min = 100 us.

Now we have to compute the hold time. I've already acquired the signal so I can leave the switch open, switch A0 is closed but all the other switches are open. We will have the Ibias of the opamp and all the leakages currents of the MUX (5 nA * 9, because we have 8 switches + 1 switch closed). Then each switch will have its own Roff and Rs.





Due to the leakage current the total voltage store on the capacitor will decrease with a constant slope, with a droop that is Ileak/C.
The second source of error is due to the resistors. Even if there were no leakages, due to the Roff, the capacitor will discharge due to the RC network. The minimum Roff is due to the parallel of all the resistances. The Roff,min causes a discharge in the time domain that will be exponential from the sampling phase to the hold phase, while previously it was a linear ramp.

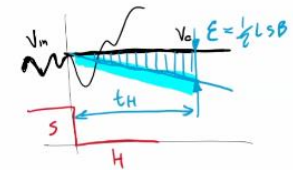


To simplify the computation, let's consider only the first part of the discharge and suppose it linear. The worst case scenario is when the discharge current is due to Vmax stored divided by Roff,min. Vmax = 1.15V.
This discharge must be compared with the previous one to see which one prevails. This last one is the dominant, because bigger.

We can hence compute the droop rate.

$$droop = \left.\frac{dV}{dt}\right|_{max} = \frac{I_{OFF}}{C_{min}} = \frac{4\mu A}{10nF \cdot (1-tol)} = \frac{4\mu A}{8nF} = 500\frac{V}{s}$$

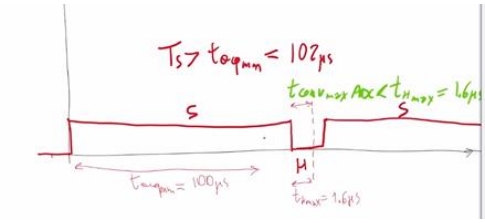
As for the hold time, during sampling we track Vin and then when we move from sampling phase to hold phase Vin can vary but Vh on the capacitor should remain constant. Due to the droop, the actual voltage on Ch decreases, and now we want the maximum t_hold that causes the voltage to decrease until we reach the maximum error we can stand. The maximum error we can stand is ½ LSB.



If the ADC has a conversion time of a given amount, we need to be sure that the conversion time is lower than t_hold,max, otherwise the voltage across the capacitor may change too much and we convert the wrong value.
Delta_V is the maximum one we can stand, that is the error epsilon. This is the time we need to wait at the most when we open the switch.

$$t_{H_{max}} = \frac{\Delta V}{I_{off}/C} = \frac{0.8mV}{500\frac{V}{s}} = 1.6\mu s$$

We must keep the switch close for a very long duration and open for a very short time duration. This is the working operation of this S&H circuit, because the circuit has a t_acq,min = 100us, so Ts > t_acq,min and instead t_hold should be smaller than t_hold,max. t_conv,max of the ADC should be shorter than t_conv,max.



So this S&H is pretty bad, because we need a lot of time to acquire the signal (100us) and we also need an ADC able to convert in less than 1.6us.
The fs of this sample and hold is fs = 1/(Ts + Th) = 1/(102us + 1us) = 9.7 ksps.
This means that the maximum frequency of the input signal, according to Shannon, has to be smaller than 4.7 kHz

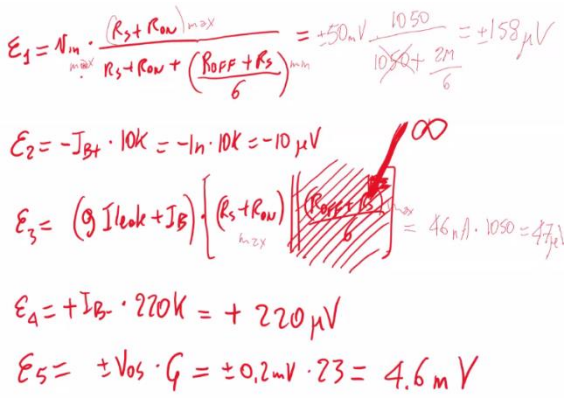
Point b)
When we acquire a signal we close a switch, which has its own Ron. Then we have the Rs of the source, so the overall voltage drop is Vin * (Rs + Ron)/(Rs + Ron + (Roff + Rs)/6). This is the voltage drop between Vin and the voltage stored on the capacitor. This is the first error in the circuit; its highest value can be computed considering the highest possible Vin, Rs + Ron maximal and the Roff + Rs minimum.

Then, there is another contribution. If we have a given amount of voltage stored on the capacitor, due to the bias current in the opamp, we have a voltage drop on the 10k resistor, that introduces the error epsilon_2. This error is always negative or positive depending on the direction of the bias current. This error has to be multiplied by the gain of the stage to refer it to the output (missing in the bottom image).

Even when the channel is closed, so we have Ron, the Ibias causes another issue. We have in fact 8 + 1 leakages currents in the MUX, and the 9 Ileak plus the Ibias flow in the Ron, causing an error in the voltage drop stored on the capacitor. This is error epsilon_3. In the worst case scenario, if all the inputs are unconnected, the Roff are in series with infinite, so it is better to consider only the Rs +Ron in epsilon_3. Also this error must be multiplied by the gain (missing in the bottom image).

Then there is the final error due to the Ibias- of the opamp. When we consider it all the inputs are at 0, so V+ of the opamp is 0 because voltage on the capacitor is 0, so also V- is 0, hence Ibias- flows just in the 220k resistor. Since Ibias- is inward going, we have a positive contribution. This error epsilon_4 is already at the output, so no need to multiply by the gain

The last final error is due to the offset of the opamp, that is 0.2mV. When we turn off all the other sources, Vos is applied straight to the opamp which has a gain of G.

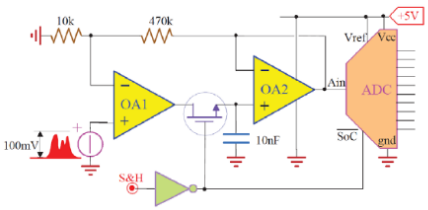


The total error at the output at the output is the sum of all these errors. Usually Ileak has a not known sign.

The non-negligible contributions are the epsilon_1, and cane be reduced by reducing Rs, then the one due to the leakages, that can be minimized by reducing Rs or the leakage currents and the epsilon_5 due to the offset. Eventually, the total error is +- 9.3 mV. To see how much the error is big, we need to compare this value with the LSB of the ADC by doing 9.3/LSB to get the digits that deviates our conversion.

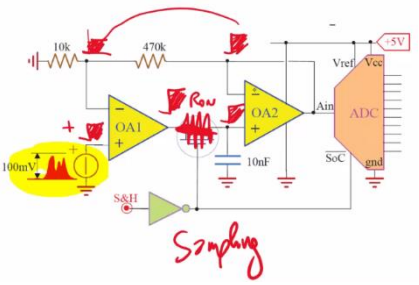
In the computations there is an error. In fact, when we study the errors, sometimes we find the error on the capacitor and some other times we compute it at the output. Of course we cannot sum the errors as they are, but we need to move them in the proper position, so if I find it at the input node I need to multiply it by the gain of the stage to have it comparable with the one in the output. For instance, epsilon_1 should be multiplied by the gain to find the error at the output of the stage. The same for epsilon_2 and epsilon_3.

Exercise 2



OpAmps: A_0=100dB, GBWP=100MHz.
MOS: R_{on}<100Ω, V_T=0.8V. ADC: 12bit.
a) Compute the acquisition time.
b) Compute static error on Ain in LSB, due to I_B=5nA (sink) and V_{OS}=5mV.

Let's start by computing the acquisition time. We know where the Vin is, the circuit has a loop and during sampling the switch is in close configuration. Let's see if the stage has a negative feedback. OA2 is a buffer, so the output increases, so we have a negative feedback, because we are forcing the error signal epsilon to decrease. If we have negative feedback, we can apply the VG concept, so V- of OA1 is at Vin.



Vout = (Vin/10k)*(10k + 470k) = 1 + R2/R1 = 48.

This is the gain of the circuit. Let's now compute the acquisition time. The switch is open, so the voltage stored on the capacitor is buffered at the output of OA2. On the V- of OA1 we have an attenuation of

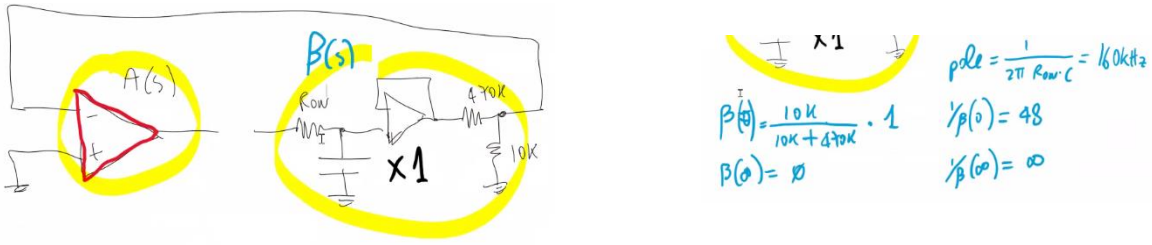
these 2V, and since the switch is open, OA1 output keeps saturating to PS high or PS low. When we close the switch the circuit has a feedback and we can compute the acquisition time.

We know the equation for t_acqu, but tau is not Ron*Ch in this case, because the circuit has feedback.

$$t_{acq} = \tau \cdot \ln\frac{\Delta}{\varepsilon} \qquad \tau \neq R_{ON} \cdot C_H$$

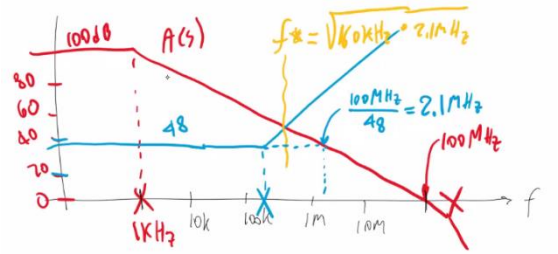
We need to compute the BW of the loop. Once we compute the BW, the tau to be used in the equation for the acquisition time is the inverse of the BW. So we need to plot the Bode diagram of the Gloop, hence extracting the pole. Then the BW will be 1/f_pole.

We can consider the first opamp to be the A(s) and the remaining circuit is the beta, considering the buffer ideal eventually, so with a gain of 1.



Let's now plot the Bode diagram.

To compute f*, we prolong the 1/beta trend, we compute the frequency and then we get f* with the geometric average.



The circuit is not stable, the closure angle is bad, but f* is not far away from the 160 kHz pole. Let's compute the phase margin.

$$PM = 180° - 90° - \arctan\frac{f_x}{160kHz} - 0 = 90° - \arctan\frac{580k}{160k} = 15°$$

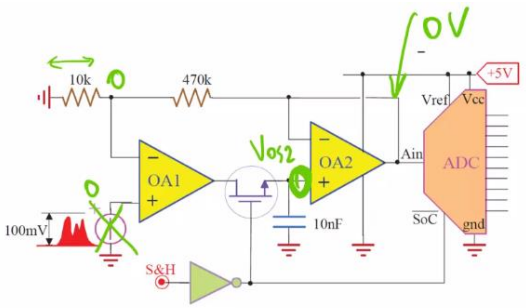
So the circuit is not stable, but we don't want to compensate it, let's compute the pole of the circuit. The pole is at f*, and since the closure angle is poor, we have two c.c. poles at f*.

$$\tau \simeq \frac{1}{2\pi f^*} = \frac{1}{2\pi\, 580k} = 275ns$$

$$t_{acq\,min} \simeq \tau \cdot \ln\frac{\Delta}{\varepsilon} = 275ns \cdot \ln\frac{5V}{0.6mV} = 275ns \cdot 9 = 2.5\mu s$$

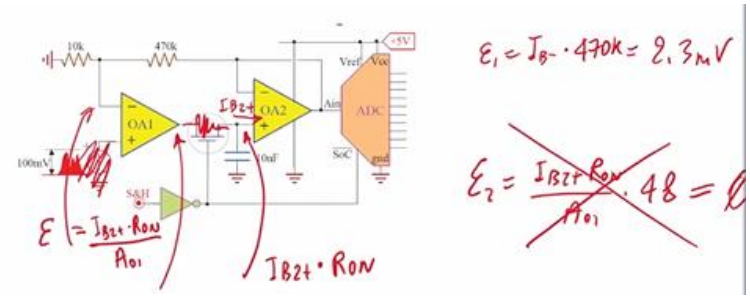
Once the tau is computed, we can compute also t_acq. Delta and epsilon must be considered where the capacitor is. Epsilon is FSR_adc/2^12, so 5/2^12 = 1.2 mV. This is the LSB at the entrance of the ADC, but I need it at the capacitor side. Since we have a buffer, the two are coincident. Then, epsilon = ½ LSB. Now we have to consider delta. We have two possibilities, the worst case amplitude is when at the output we reach 5V out of OA2. If so, delta at the capacitor is also 5V. However, maybe the output voltage at the input of the ADC will never reach 5V if we have a certain maximum input signal.
If the Vin,max is 100mV, the output maximum voltage is 100mV * 48 = 4.8V. So 5V was ok.

Let's now compute the static errors.
We have the Ibias of the opamp and the Vos of the opamp. Let's consider Vos1, at the + terminal of OA1; the gain of the stage is 48, so for sure the error in output is Vos1*48 = 240mV.
LSB = 1.2mV, so this error correspond to 240/1.2 = +-200 digits error.
As for the offset due to the second opamp, we have negative feedback. If the circuit was ideal, no current in 470k and the output is 0V whatever is Vos2.



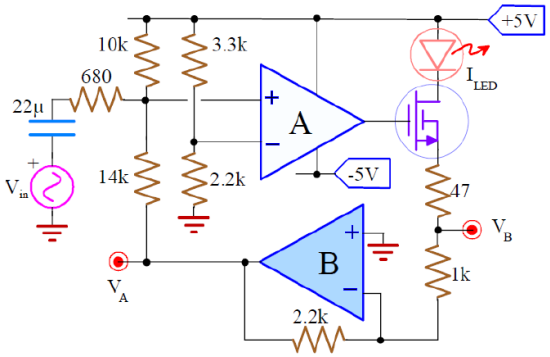
This because the ideal gain doesn't depend on A(s) and so on its errors along the path. In the real case, if we have Vos2, it means that OA1 provided Vos2, and this can happen is epsilon at the input of OA1 is Vos2/A0, since Vin is off. Once we have this voltage, we can propagate this voltage back to the output. So the error at the output will be Vos2/A0 * G that is almost 0.

Eventually, in this circuit we have two more errors, due to Ibias. Ib+ of OA1 has no error because the input is shorted, while Ib- gives an error flowing through the 470k resistor.
Ib- of OA2 gives a negligible effect because it flows and comes from a voltage generator (the output one of OA2), while Ib+ of OA2 is inside the loop. Ideally, this current doesn't have an effect, because the output is 0 no matter what. But to be precise, Ib+ of OA2 flows through Ron of the MOSFET, so the voltage drop that develops on it is Ib+ * Ron. OA1 should provide this voltage, and this is done if at its input we have an epsilon = Ib+ * Ron / A0. Then this epsilon experiences the gain G = 48, so also this contribution is negligible.



EXERCISES

Ex. 1



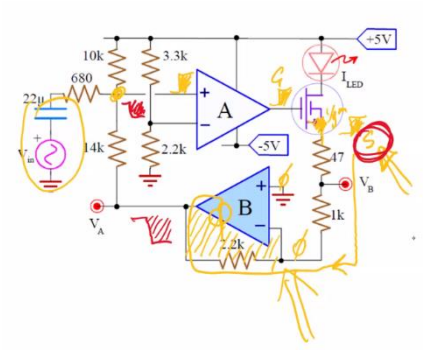
OpAmps with A_0=50V/mV, V_{OS}=5mV and I_B=200pA. MOSFET with k=10mA/V² and V_t=0.8V.

a) Compute the relationships of V_A, V_B and I_{LED} vs. V_{in}.

b) Compute the input pole and plot $V_A(t)$ and $I_{LED}(t)$ waveforms when the input is a high frequency 200mV_{pp} sinusoid.

c) Compute the maximum V_A static error due to V_{OS} and I_B of the OpAmps.

Resolution



We must investigate the operational regime of the schematic. The signal enters, reaches the node x and then we have OAA. The output will increase if we have the positive input, the transistor is in source follower configuration, so the source will follow and increase. If so, the second opamp has a local negative feedback and so the – terminal will behave as a virtual ground. So the current will increase and flow through the feedback resistance of OAB and so the output voltage decreases. If so also the input node will try to decrease due to the feedback → overall negative feedback.

If we have a negative feedback, it means that we have the concept of VG that applies. Amplifier B is in inverting configuration with G = -2.2k/(1k + 47). So the infinite gain will be provided by opamp A. So the virtual ground will be the epsilon = 0 at the input of OAA, so the + and – terminal are at the same voltage.

Due to the power supply we have 2V at – terminal of OAA, and also at the + terminal due to the negative feedback.
If there is no input signal, then the capacitor is open and the current that flows in the 10k resistor should the be same of the 14k resistor.

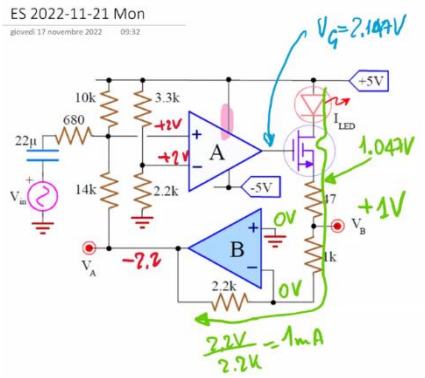
$$i = \frac{5-2}{10K} = 0{,}3m = 300\mu A$$

This current flows in the 14k resistor so, in DC:

$$V_A = 5 - i(10k + 14k) = -2.2V$$

Then, in order to provide -2.2V at Va and since B is in inverting configuration and its terminals are at 0V due to VG, there must be a current in feedback of 1mA, and this current will flow in the mosfet and in the led. Hence Iled = 1mA in DC, and the led is on.

We can also compute the voltage at the gate of the mosfet if we know the equation of a mosfet transistor.
Vov = Vgs – Vt = sqrt(Id/k) → Vgs = sqrt(Id/k) + Vt, with Vt = 0.8V and k = 10mA/V^2
In the end, Vgs = 1.1V.

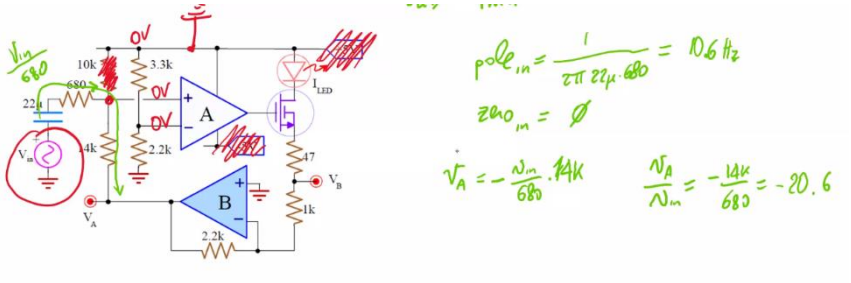


This Vg value is coherent with the PS values.

Point a)

Let's now move to the signal analysis. For the signal we can consider the PS to be off and we let just the signal operate. If so, for the signal the input of OAA is 0V on the + and – terminals. So the current that enters is in 0 in DC, but now if the capacitor is shorted, is Vin/680. We can also compute the input pole of the network. Of course there also a zero at 0 frequency.

Since the PS is off, there is no current in the 10k resistance, because it is between 0V, so the current will flow in the 14k resistance.



I know that the signal propagates clockwise, but I'm computing everything moving counterclockwise, till the position I want to reach.

To compute Vb, we know the voltage Va and I know that Vb*(-2.2k/1k) = Va.

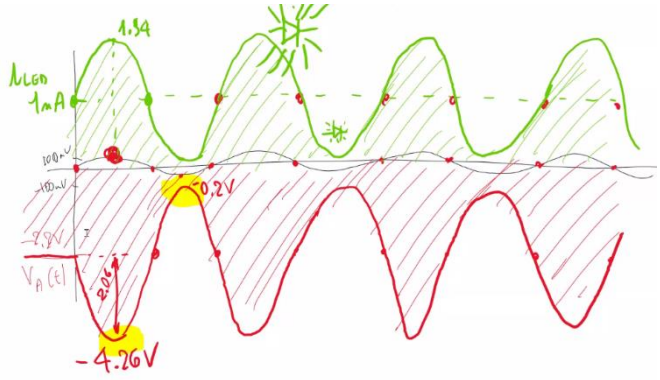
Now we can compute the Iled current, because we know the value of Vb. If Va is positive, the current has a negative direction to be coherent. Hence we can compute the final gain Iled/v_in = 9.4 mA/V.

Point b)

Since we are asked at high frequency, we are above the pole of the capacitor. So the input signal is a sinusoid at high frequency, 200mV pp, so between 100 and -100mV. Moreover, in DC Iled = 1mA constant and Va = -2.2V constant.

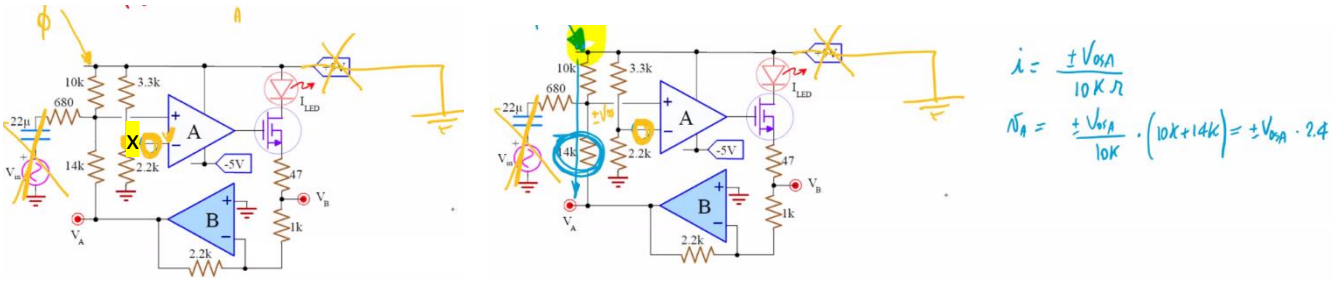
Since the gain between Vin and Va is 20.6, at 0V I'm at -2.2V, at -100mV I've reached a value -4.26V. the red one is the ideal Va waveform. In fact, in principle the opamp could reach the value according to the PS. In this case -4.26V is ok because theoretically OAB has the PS that is +-5V. But I should also consider the value of Vb and if also the ranges of OAA are compatible with this value of Va.

The gain for Iled is 9.4mA/V, so from 1mA I will reach 1.94mA. Also for the Iled it's valid the reasoning on PS done for Va. Moreover, I notice that the current Iled is never becoming negative, and it is ok because the current cannot flow in the opposite direction in the MOSFET.



Let's see if Va can be -4.26V. Of course we can because Vb = Va/(-2.2) = 1.9V, that is also equal to the source voltage because the 47Ohm resistance is very small. Let's now compute the Vgs of the mosfet, that is 1.6, giving 3.6V in output of OAA. So the very negative voltage can be reached. We should also do the same for the -0.2V peak, but it is soon verified and reasonable.

Point c)
We are considering the Vos on the – terminal of OAA. If we have the offset there, we must turn off Vin and the PS. If so, x is at 0V, so + terminal of OAA is at Vos.



Since the offset is a DC value, the current cannot flow in the capacitor because it is open. Since Vos = +-5mV, Va varies between +-12mV.

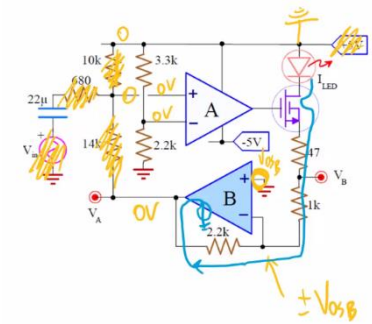
I can now compute also Iled.

$$i_{LED} = \frac{V_A}{2.2K} = \pm \frac{V_{OSA} \cdot 2.4}{2.2K} \simeq \pm 5\mu A$$

This value is so small with respect to mA that is negligible.

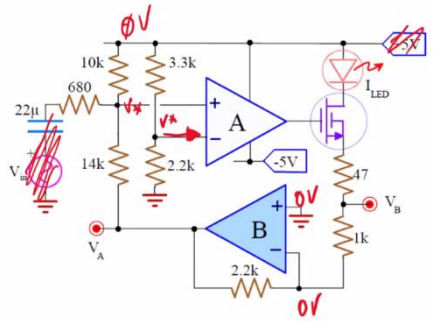
Let's now move to Vos of OAB, that we place on the + terminal. The – terminal of OAA is at 0V, so also the + terminal of OAA it is, so no current in 10k resistance and hence no current in the 14k resistance, so Va = 0V. Hence we will have a current in the 2.2k resistance given by +-Vos/2.2k. This is the Iled current.

We see that the offset of OAB on Iled is smaller with respect to the one of OAA.



$i_{led} = \frac{\pm V_{OSB}}{2.2k} = 2.3\mu A$

Now we have to consider the bias currents contributions of the two opamps.
For Ib- of OAA, once V* is known, also the + terminal will be at V+, so we will have a current in 10k and we can compute Va and Iled.

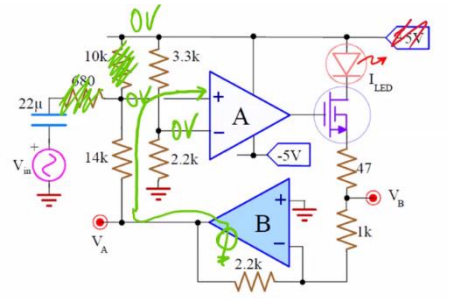


$V^* = I_{B-} \cdot (7.2k \| 3.3k) = 264nV$

$V_A = \frac{V^*}{10K} \cdot 24k$

$i_{led} = \frac{V_A}{2.2k} = \frac{264nV}{10K} \cdot \frac{24k}{2.2k} = 230pA$

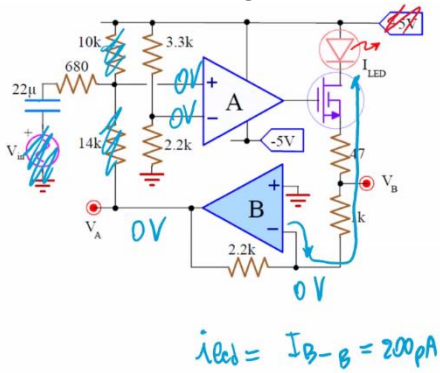
As for Ib+ of OAA, it comes directly from OAB.



$V_A = I_{B+} \cdot 14K$

$i_{led} = \frac{V_A}{7.2K} = \dots$

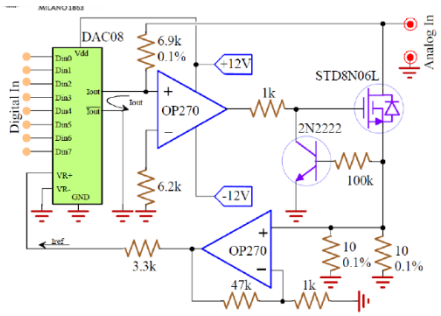
For OAB, Ib+ has no effect, while Ib- will flow through the led.



$i_{led} = I_{B-B} = 200pA$

These value of currents don't have a direction, they simply add up on the top of the bias current.

Ex. 2



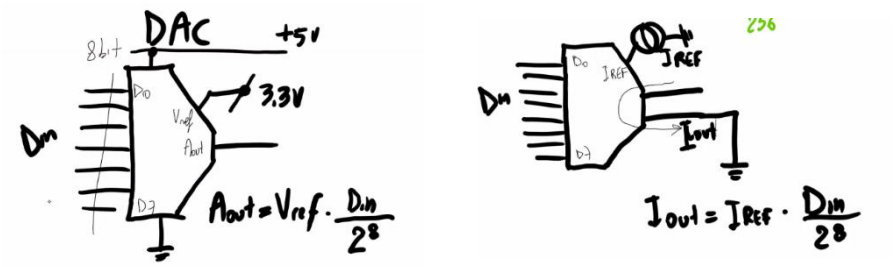
The DAC provides $I_{out}=I_{ref}\cdot D_{in}/256$ (D_{in} is the Digital In value) and **virtual ground** at its V_{R+} input. OpAmps have A0=1500V/mV and GBWP=5MHz. A voltage generator is applied at the Analog Input.

a) Obtain the analytical relationship V_{in}/I_{in}, as a function of N.

b) Reckon if the stage is stable or not when D_{in} =255. Moreover, tell if stability improves by reducing D_{in}. (hint: assume $1/g_{mMOS}$=495Ω and ignore the role of 2N2222 BJT).

Resolution

We have a DAC converter, which takes a digital input and gives an analog output.



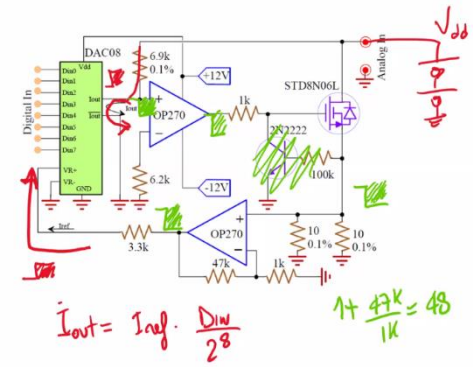
In the exercise we are not using a voltage DAC (VDAC), but a current DAC. We can provide a reference current to the circuit, Iref and the current Iout will be given by the same equation than in the case of the VDAC.

Point a)

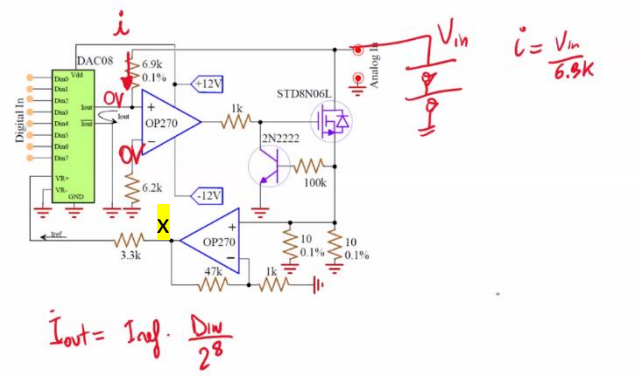
We apply something in the Analog In port and in a Digital In port.

$$I_{out} = I_{ref} \cdot \frac{D_{IN}}{2^8}$$

If Iref increases, then also Iout increases. But if Iout increases, then the voltage at the – terminal of OAA (upper opamp, OAB bottom one) decreases, and so also the output. If so, let's forget the BJT, we have a source follower and then a non-inverting configuration, so also the output decreases, so we are causing a Iref inside the opamp → the loop acts trying to reduce the Iref that we have at the beginning, so we have a negative feedback.

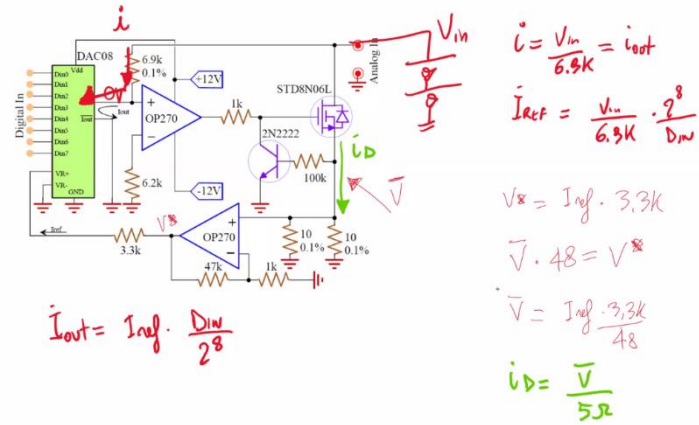


The component that has a high gain is not OAB, which has a fixed gain, not the mosfet, which is a follower, so it is OAA. So both + and – terminals of OAA will be at 0V.



The current i must flow somewhere, but we cannot force a current into the DAC, so when we apply Vin the current causes the + terminal of OAA to increase and due to the feedback and the loop, Iout starts to flow and the full i will start to enter in the DAC, but not because we are forcing the current in the DAC, but because we have a loop that reacts to have the DAC drinking it.

Since we know the relationship between Iout and Iref, I can compute Iref and then V* at node x.



Given V-hat we can compute the current i_d.

$$i_D = \frac{\bar{V}}{5\Omega} = \frac{3.3K}{5 \cdot 48} \cdot \frac{2^8}{6.9K} \cdot \frac{V_{in}}{D_{IN}} = V_{in} \cdot \frac{0.5}{D_{IN}} \; A/V$$

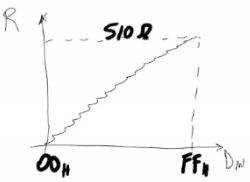
But I need the value of Iin in the Analog In terminal, that is:

$$I_{IN} = i + i_D = \frac{V_{in}}{6.9K} + \frac{V_{in}}{2\Omega \cdot D_{IN}}$$

The first contribution is negligible with respect to the second one. The final equation we get is the following.

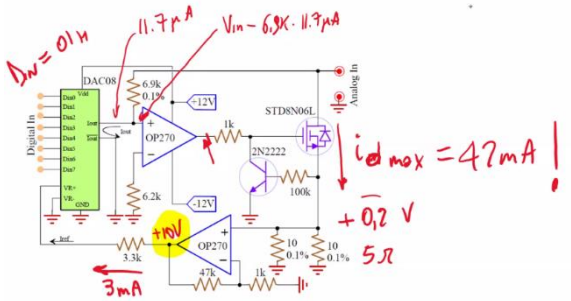
$$\frac{V_{in}}{I_{in}} = 2\Omega \cdot D_{in}$$

Hence this circuit is a programmable resistor. We apply our Vin, the current will enter and will flow in R, whose value will depend on the digital codification. Let's plot the relationship.



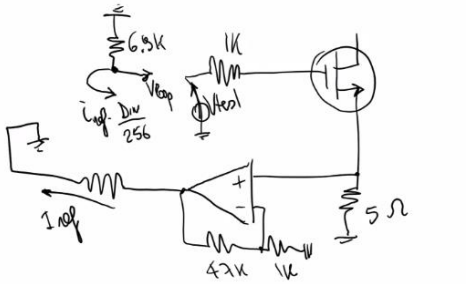
Of course, I cannot apply any Vin I want in input, because we could eventually get currents that must be compliant with the circuit. In fact, the value of current has to guarantee a valuable Vgs value on the MOSFET so that the opamp can provide it. So we should check that the circuit respects the biasing conditions and the power supply.

To compute the maximum ratings in the circuit we have to consider the PS values in output of the opamp (even if we should have some headroom from the PS). Let's suppose the PS is +-10V.

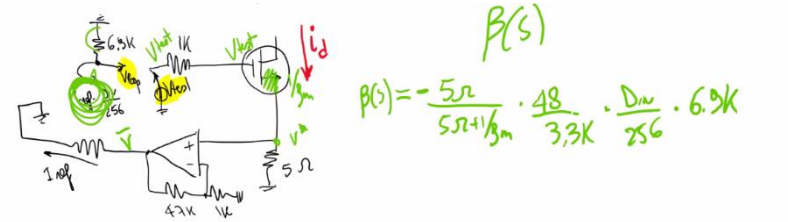


Since the voltage there is 0.2mV and the bottom resistor is 5Ohm, the i_d,max can be easily computed. 42mA is the input maximum current.
We cannot provide higher values of current because OAB saturates. To increase the 42mA value we can decrease the non inverting gain of OAB. If instead of 47k I use 1k, the i_d,max = 400mA.

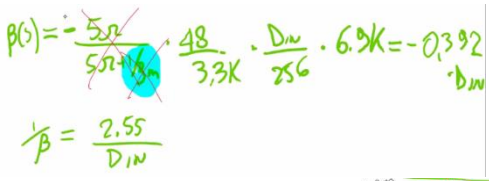
Point b)
Let's compute the stability of the circuit, so we need to identify A(s) and beta(s). To choose A(s) better to select OAA, because it is a OL opamp without a local feedback.
I apply Vtest and measure Vloop.



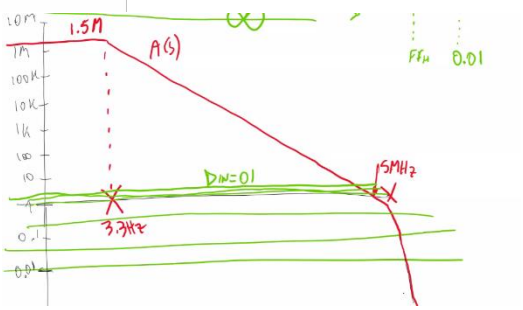
The transistor is a follower with a 1/gm, but in a first order approximation we can consider it negligible with respect to 5Ohm resistance.



It's negative because I drink current from a resistor connected to ground, the 6.9k resistor.

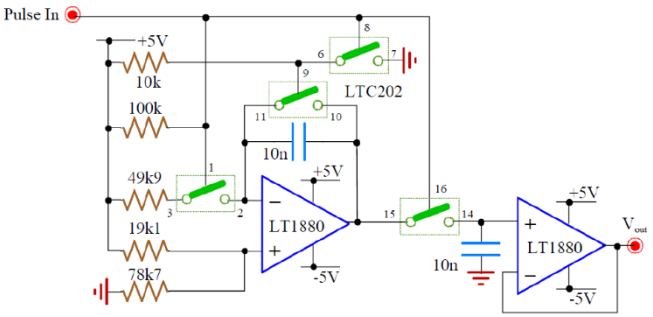


Let's now draw the Bode diagram. 1/beta depends on the value of Din. In the exercise we are asked when it is Din = 255.



In the worst case scenario, f* is a decade after the second pole and we are very unstable. If we don't have a second pole, the circuit is always stable.

Ex. 3



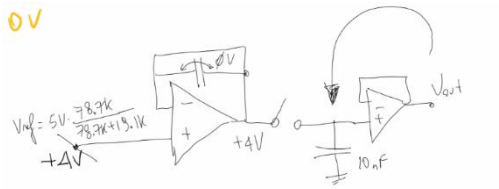
The LTC202 is a quad analog switch (closed when control pin is high). The input is a pulse, whose width T_{high} is in 1ms÷2ms range. The OpAmps have I_B<1.5nA in the -40°C÷+85°C range.

a) Compute V_{out} as a function of T_{high}.

b) Reckon the min T_{low} that guarantees a precision of 1µs.

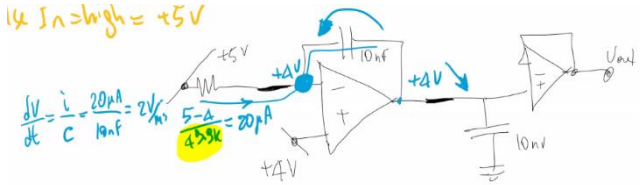
Resolution

The switch is closed when the pin is high. The input is a pulse. If the pulse is low, e.g. 0V, we have the switch 8 open, so 9 is closed and first opamp is a buffer. Switches 1 and 16 are instead closed.



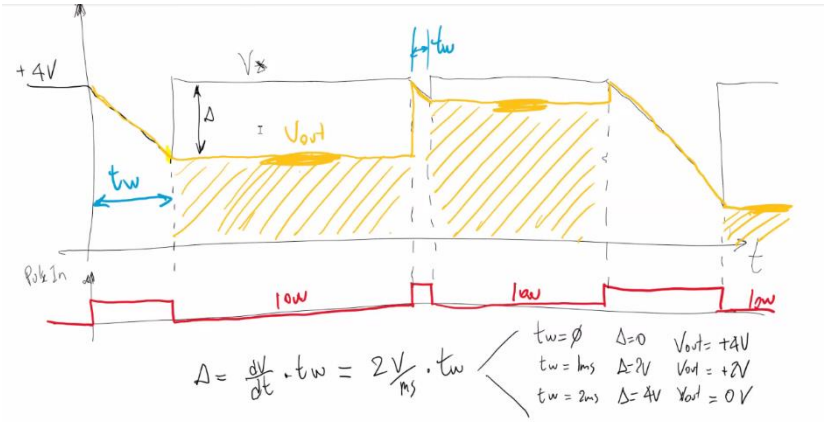
The capacitor 10u is charged to 0V and Vout is equal to the previous value stored on the 10nF capacitor.

If Pulse In is high, 9 is open and all the other switches are closed.



We have a constant current that is 20uA. So the capacitor in feedback will be charged and the 10nF capacitor will start to discharge.

Let's plot the time-dependent waveforms. When the pulse is high we start the integration and the output voltage decreases. Then, when the pulse resets to 0, the voltage in output to the first opamp resets to 4V.

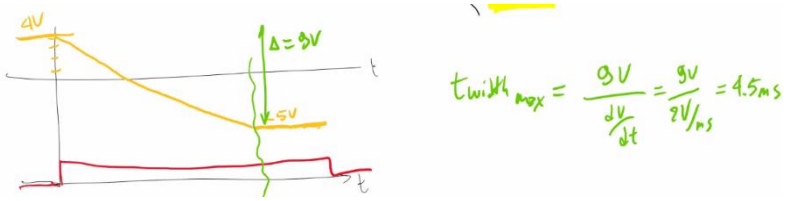


The output voltage is equal to V* when we have the switch closed, but then it remains to the previous voltage when the switch is open. Every time the pulse is low, the output is a constant voltage.

We can also compute the droop.

This circuit converts the duration of a pulse into a voltage.

If the width is longer than 1ms, the current won't go in the second capacitor, but it is drunk by the first opamp, and the capacitor is forced to obey to V* (output of first opamp). If the pulse lasts too long, V* goes negative, and this can happen because we have also negative power supply. However, it cannot go lower than -5V.

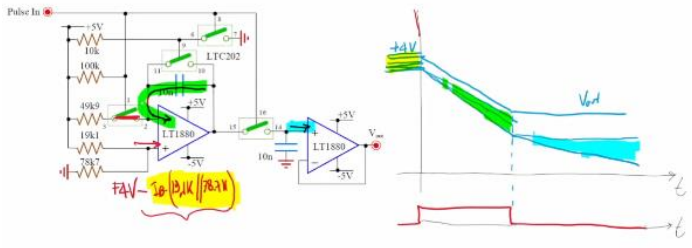


So better not to go longer than 4.5ms.

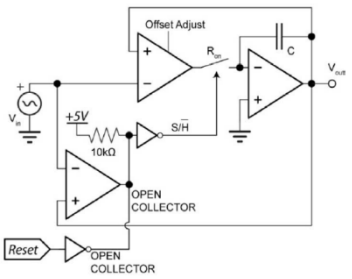
Point b)

I want to be capable of detecting a width of 1us. This question is related to the tolerances due to Ibias. Due to the Ibias current, the second opamp Ibias causes a droop. The Ib+ of OA1 causes an error, so V+ won't be 4V but 4V – Ibias*(19.1k || 78.7k), while the Ib- of OA1 flows in the feedback capacitor.

Due to the Ibias of the opamps if the pulse is rectangular, we start from +4V and then it should go down and remain constant. But due to possible errors, we start not at 4V, but at a different value. Then due to the current in the feedback capacitor we cause a droop, so the decreasing slope can be higher or slower. Moreover, due to the Ibias of the second opamp, even if Vout should be constant it goes down due to the droop.



Ex. 4



The Offset Adjust pin in set in order to provide a +10mV output offset. R_{on}=10Ω, C_H=1nF, GBWP=10MHz. The comparator has open-collector output.

a) Draw the quoted waveforms at all nodes when a triangular input from 0V to 2V is applied at V_{in}, with 1ms period, and a reset pulse of 50% duty-cycle is applied every 1.5ms.

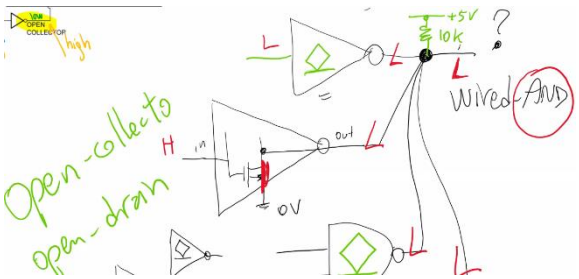
b) Explain the circuit behavior and the role of the offset adjust.

c) Compute the max input frequency that ensures an error lower than 1LSB for a 10bit ADC with FSR=5V.

Resolution

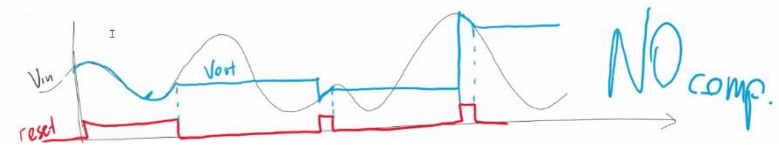
We don't know how the switch behaves, but we know that we have a S&H, and according to the indication on the circuit, when the pin is high, the switch is closed. Since we have an inverter, the input of the inverter must be low to have the output high, so either the output of the opamp is low or the reset pin is low.

The problem in this circuit is that we cannot connect a digital gain with an analog output, or we burn one of the two components because of short circuits. However, if we use gates that are open collector (also called open-drain) we can have the output connected. In them, we don't have the p-channel part. Then we use a pull up resistor that fixes the voltage. If then one gates wants to go low we are ok. So the output is high if no one is pushing it low, while it is low if one or more than the connected gates is 0.



So in reality the circuit is correctly wired.
The other opamp is a comparator that checks Vin and Vout. If Vout is higher than Vin, the output of it goes high. Hence I enter the sampling phase and the switch closes when either I apply a high voltage on the reset pin or when the comparator has the output low, that is when Vout < Vin.
If the switch is closed, Vout = Vin and I store Vin on the capacitor C. Then when the switch will open, the capacitor will remain charged to Vin.

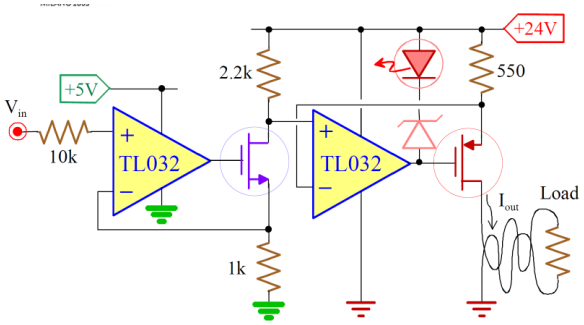
Without the comparator, but just the reset pin:



However, if we consider the comparator:

First we have a reset pulse, and then is the comparator that is acting. During the reset pulse, Vout = Vin, but then when the reset pulse ends, if Vin > Vout, the switch is closed, otherwise it is open.
This circuit is a peak stretcher.

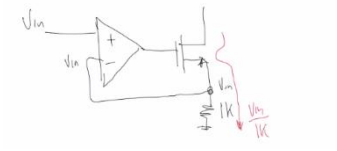
Ex. 5



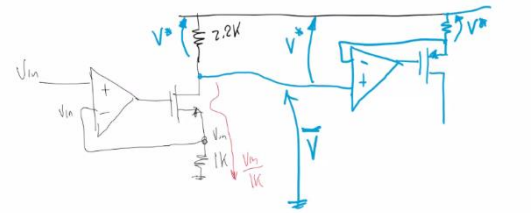
A 5V zener is in series to a 1.5V LED. MOSFETs have V_T=1V and k=½μC_{ox}W/L=2.5mA/V².
a) Find the relationship I_{out}/V_{in} and the Vin,max that ensures linear behavior for rail-to-rail OpAmps.
b) Change the first stage in order to employ the same +5V power supply, but providing a $V_{in,max}$=+5V.
c) Tell in which conditions the LED will light up and why the circuit is prone to burnings.

Resolution

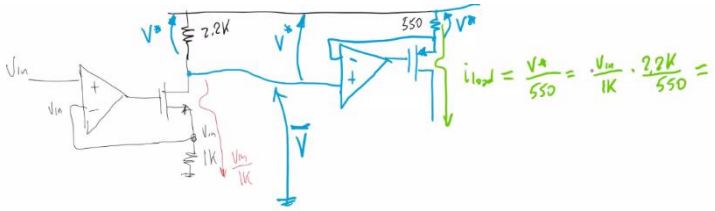
The first stage is a voltage to current converter, with a negative feedback.



Then we have a 2.2k resistor and another stage that reads the voltage at the + terminal and has a p-type transistor, still with a negative feedback.

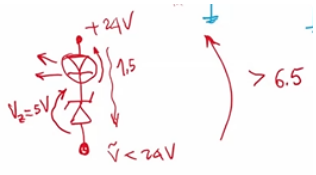


Given V*, I can compute I_load.



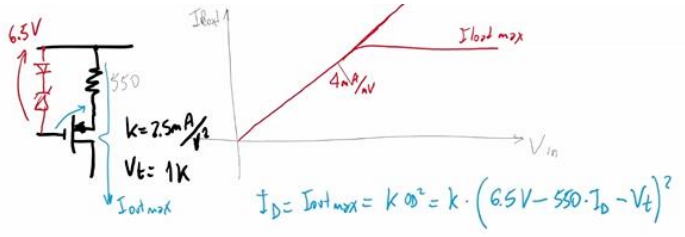
The output current is independent on the PS, it depends on the Vin only. So this overall circuit is a perfect current generator, because the output impedance is infinite and it provides a current.
In addition, we have a LED and a Zener diode.

If the upper voltage of the series (24V) is higher than the bottom voltage of the series, the current should try to move top bottom, and the LED allows this, because forward bias, but the Zener not, only if we exceed Vz, which is 5V.
Hence the two will turn on if we try to increase higher than 6.5V. The LED will be brighter and the Zener will force 5V across it.
If the voltage across the pair cannot be more than 6.5V it means that there is a maximum out current.

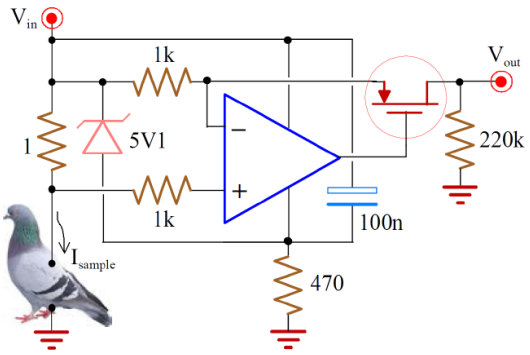


So when Vin is low the current is low and the voltage drop is low and maybe the LED is not on. if we increase Vin, the current increases, voltage drop increases and maybe we have more than 6.5V on the series but if so, even if we increase it even further, the voltage drop cannot increase more and the output current will be fixed to a certain value.

We can compute the value of Iload,max. We know the K value for the transistor, its Vt.



Ex. 6

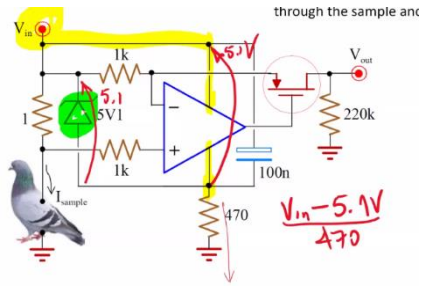


a) Compute the relationship between Vout and Isample.

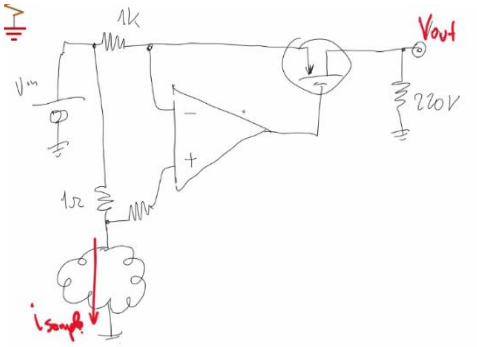
b) Design a new stage employing OpAmps that instead forces a constant current I_{sample}=10mA through the sample and measures the voltage developed across it with a gain of +20.

Point a)

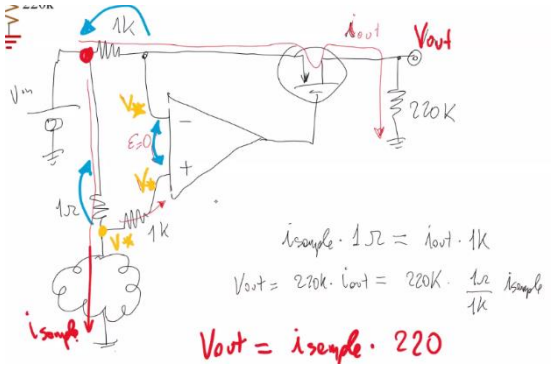
We apply a Vin voltage and there is no power supply, because it is already provided by Vin. Then we have a Zener diode, so if Vin is higher than Vzen, that is 5.1V in this case, then the diode is on and the voltage across the Zener is 5.1V, so also the voltage across the PS of the opamp is 5.1V. So the Zener is needed to provided a constant DC voltage as PS.



Then there is a capacitor connected between the two PS of the opamp, so it is just for noise reduction in the opamp. So if we remove the Zener, the capacitor and the 470 ohm resistance, the circuit reduces to the following one.



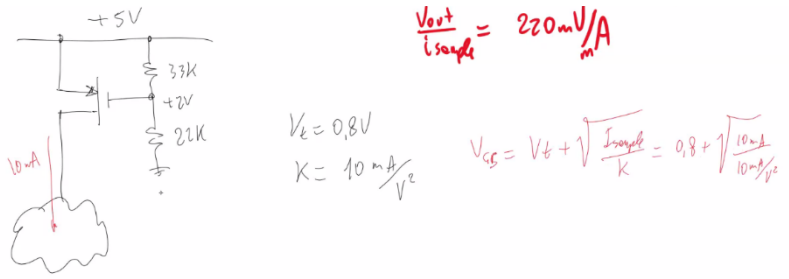
There is an overall negative feedback, so we can compute the ideal gain so that the epsilon error across the inputs of the opamp is 0. If so, no current flows in the 1k resistance on the + terminal, so the voltage V* is copied on the – terminal. So the voltage drop across the 1 ohm resistance is the same over the upper 1k resistance if epsilon is 0.



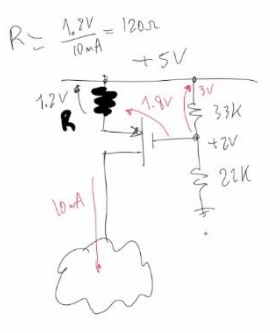
The gain is in mV/mA. So this circuit is a current to voltage converter, it measures a current flowing into a sample and then converts it into a voltage with a sensitivity of 220 mV/mA.

Point b)

We have to design a circuit that provides a constant current of 10mA through the sample. We can use a p channel transistor and fix the voltage at its gate. Then knowing the Vt of the transistor and the K we can compute the Vgs we need to provide a 10mA current.

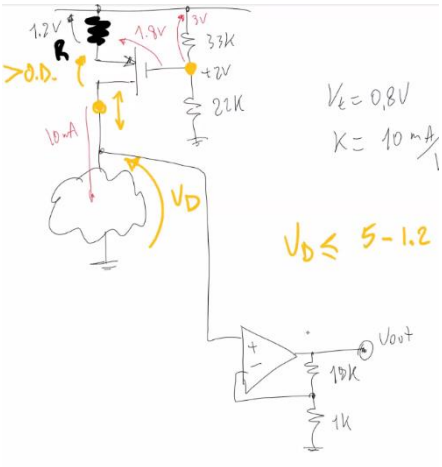


But 1.8V as Vgs is not true in the schematic drawn, so we need to add a resistor whose voltage drop will be equal to 1.2V for a current of 10mA, so R = 1.2V/10mA = 120 ohm.

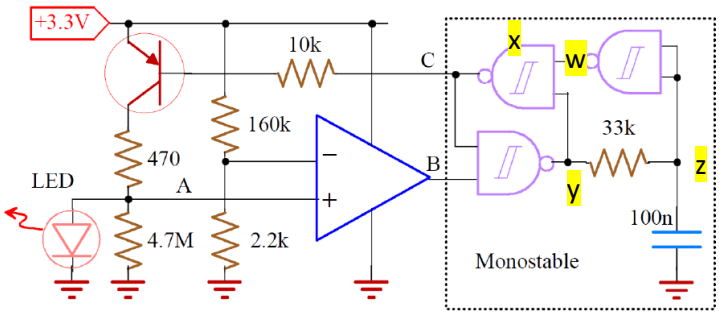


Vdrain can move up and down but Vds has to remain greater than 1 overdrive, so Vd < 5 – 1.2 – 1 = 2.8V.
So the circuit is a current generator if the voltage across the sample is smaller than 2.8V.

Now we want to measure the voltage across the sample, so I can use a simple non inverting configuration.



Ex. 8



The circuit switches on/off the LED if there is dark/light, without using any other photodiode, but exploiting the photovoltaic effect of the LED itself (with no voltage applied and in light condition it produces 60÷90mV).

a) Plot all voltage waveforms **with light**, explaining circuit behaviour.

b) Plot all voltage waveforms **with no light**, explaining circuit behaviour.

Point a)

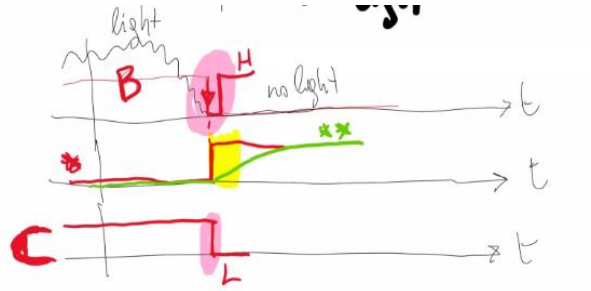
The LED turns on if we apply a voltage but, when the LED is off and there is light hitting it, then the LED behaves as a PD and produces a voltage.

Let's imagine that there is enough light on the LED, so the voltage across the LED is higher than 60mV, so even if the BJT is off even if the base is high, and the + terminal of the opamp is at 60mV. The other – terminal is at 45mV, so if there is light the comparator is operated as a trigger and the output is high level. Then we have NAND gates logic. The NANDS are Schmitt trigger because we have a RC network that provides an analog signal, so we have for sure to use it.

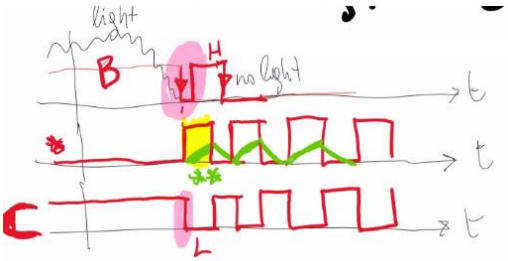
At regime, the C is fully charged and no current flows in the 33k resistance, and voltage across it is 0. At regime, the two inputs of NAND x are always opposite, so the output is always high and the transistor is off.

If there is light in the LED, output of the opamp is high and of x is high, so BJT is off. If we remove the light, then voltage at + terminal will be around 0, so the output of the opamp commutes to low level, that is GND. So node B experiences a high to low transition, and if so, and node C was high, we have low across the 33k resistance. But if B is low, now the first NAND has a H value in output. So if node y goes high, then node z starts to increase and the capacitor charges up until also node z is high. But before node w goes low, it is high, so the output of NAND x goes low. If so, the transistor turns on and it can be considered closed. So current will flow in the 470 ohm resistance. This current causes node A to increase until the LED turns on, e.g. at 1.5V (typical turn on voltage of a LED). But if so, then the voltage at the input + of the opamp is no more 0 but it becomes high, and the output returns high.

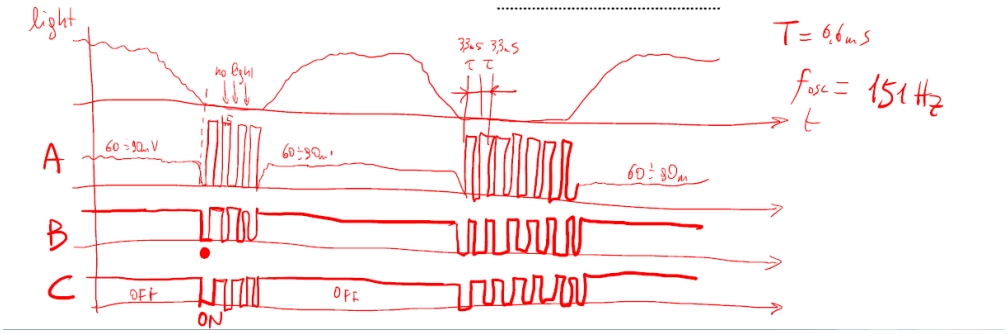
Hence node B, which experiences a falling edge when the light is off, thanks to the fact that node C goes low, immediately returns to a high value.



Hence voltage at node y stays high, so the voltage across the capacitor increases up until it becomes a high level. If so, node w becomes low level and node C, which was low, commutes to a high level (because node y is high and w is low).
If node C returns to a high value, node B decreases and, if there is still no light, we return to the previous situation. Hence the circuit continues to commute.

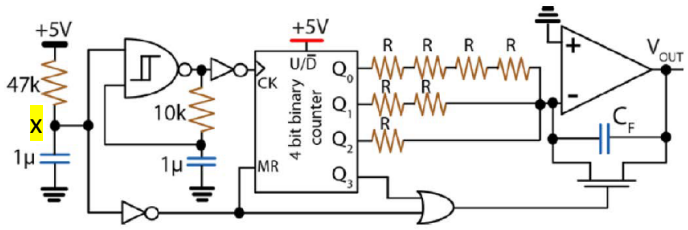


If there is light, node C is permanently high because node B is permanently light. If there is no light, node C goes low and the led is on, and also B goes low. The oscillating condition keeps running until light is on.



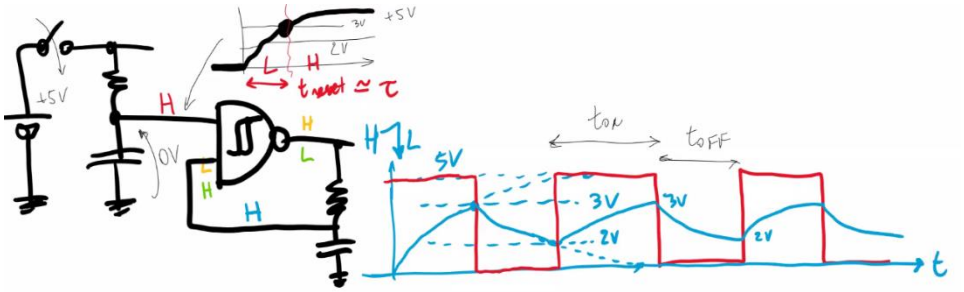
So the LED is used to illuminate the environment if there is light, but it is also capable to detect if there is light in the environment and switch off.

Ex. 11

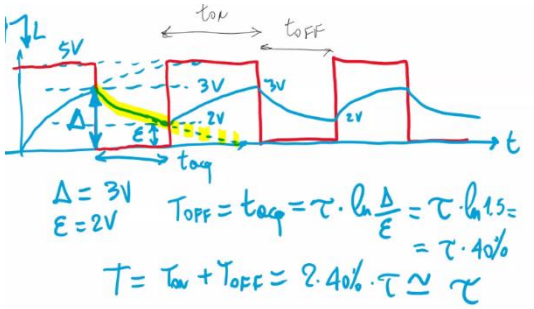


R=100kΩ, C_F=10µF. The digital counter provides +5V high levels and Q_0 is the least significant bit.
a) Plot the V_{OUT} quoted waveform during the first 3 clock periods (10ms period).
b) Plot CK and V_{OUT} waveforms during the first 200ms.

So we have a Schmitt trigger NAND gate (because we have an analog signal in input to it). When the output of the trigger is high, the capacitor charges and when the threshold is crossed the output of the trigger goes low → we have an oscillator and voltage across the capacitor keeps going up and down between 2V and 3V.

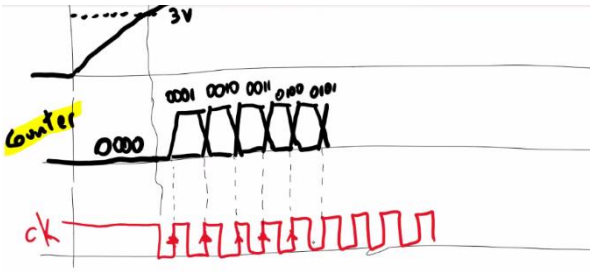


We can compute t_on and t_off. If we look at the commutation between 3V and 3V, the delta is 3V, but then we have to wait for a time t_acq until the error is 2V.
So we can use the equation Toff = t_acqu = tau*ln(delta/epsilon). Then Ton is almost Toff.



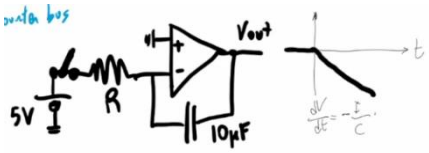
Hence we may say that f_osc = 1/tau = 100Hz.
The clock is then provided to a counter. The Master Reset pin is connected to the Power on reset network at the beginning, which provides a reset at the power on. When we turn on the PS, node x is low, so the counter is reset. When the reset phase is finished, the node x is high and the reset is no longer active.

The counter is a digital component with 4 output lines. At the beginning the output is 0000, then the circuit starts to oscillate, and every time we apply a rising edge, the clock increments its output.



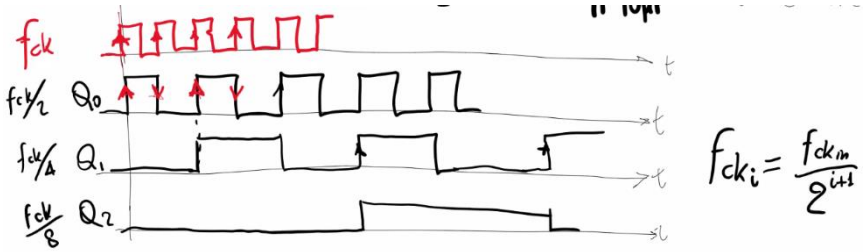
Then the output bus of the counter is connected to an opamp. When Q3 is low and also the other wire, the NOR gate is off and the transistor is off, so we have an integrator.

Let's draw the stage. We have an opamp with the + input grounded, the – input connected to a feedback capacitor and then we have a resistor R. If to R I apply a voltage, the circuit is an integrator and the output voltage will be 0 and then, when the switch is closed, the capacitor will integrate the current and the slope will be -I/C.



The problem is that here we don't have one resistor, but many of them, because we have different bits.

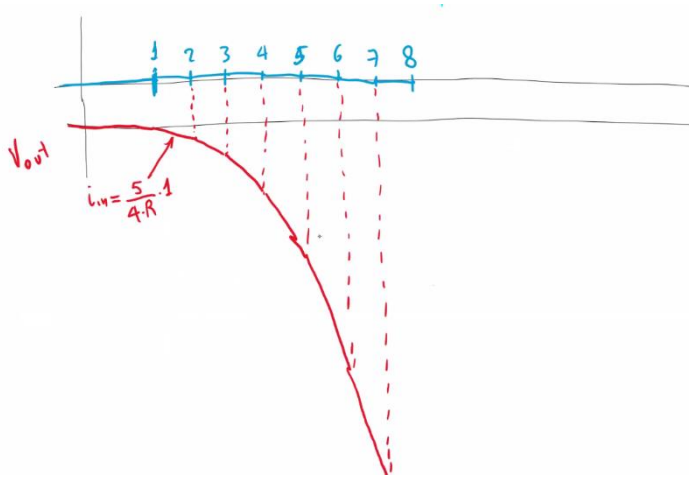
In a counter, we start from the LSB Q0 and at every clock pulse Q0 commutes, as for Q1, is commuting each rising edges of Q0.



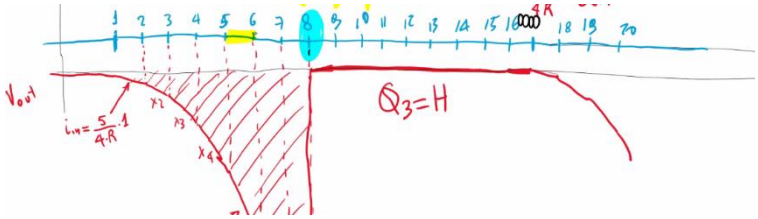
So given f_clock, Q0 oscillates at f_clock/2 and so on. so the f_clokc of the highest bit is f_clock/2^(i+1). Every time a bit is high, it contributes to the current in the integrator. Of course, the higher the resistor in the output path, the lower the current. Total current in the integrator will be the following.

$$i_{in} = Q_0 \cdot \frac{5}{4R} + Q_1 \cdot \frac{5}{2R} + Q_2 \frac{5}{R}$$

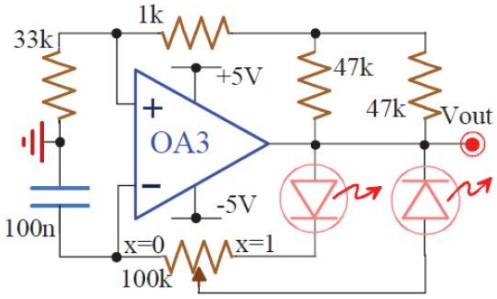
This is like writing i_in = 5/R * Dout, where Dout is the digital output of the counter. Once we reach the number 8, Q3 is high and the others are 0. If so, the NOR will have high in output and the n channel transistor behaves as a closed switch, so the capacitor is quickly discharged back to 0. Hence the output of the stage will be the following one. After the first clock, the circuit starts to integrate the current, and current that is integrated increments at each clock pulse.



Then at the 8th clock pulse the capacitor is discharged and Vout collapses back to 0. So the circuit provides an output voltage that is like a paraboloid curve.
Then for all the time Q3 is high the Vout is fixed to 0.



Ex. 13

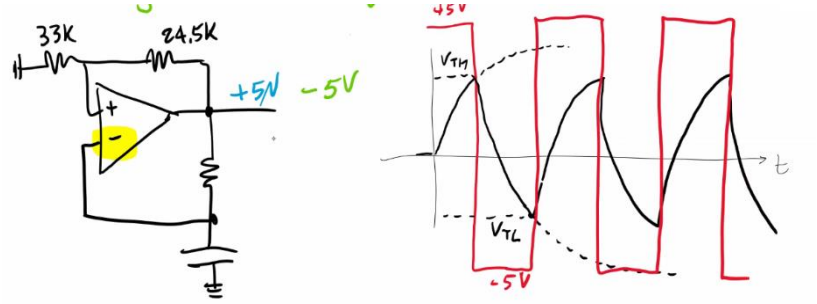


Employ 1.5V forward bias LEDs and a rail-to-rail OpAmp.

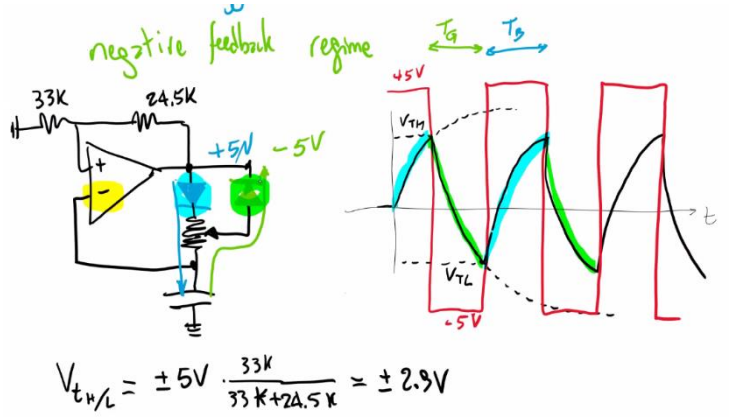
a) Study circuit operation when the pot is turned to 100% and draw the quoted output waveform.
b) Compute the analytical dependence of the output main parameters on x pot position.

The opamp has two possible feedback path, so we have both a positive and negative feedback. Which one prevails? Of course, during fast transitions of the input, the capacitor is a slow component, so the upper path is very fast, while the green one is slow because we have the capacitor. So in fast transients we have a Schmitt trigger and the output saturates to PS. But then, after a transient the capacitor charges up and, at regime, the capacitor will open and the negative feedback becomes 1. So at regime there is negative feedback that prevails. So trigger during the transient and negative feedback at regime.

This circuit is an oscillator. Once the output is at PS, it charges the capacitor and when the other threshold is reached, the output commutes to the other PS value.

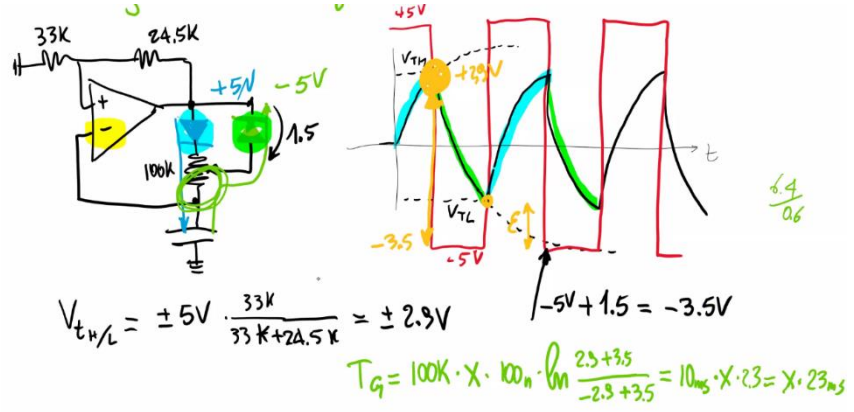


However, in our circuit we have some differences, a LED that goes to a potentiometer and then to a capacitor and then a cursor biased by another LED.
When the input is high, then the left LED is on, when low, the right LED is on. Again, we can compute the durations of the two phases and the thresholds.

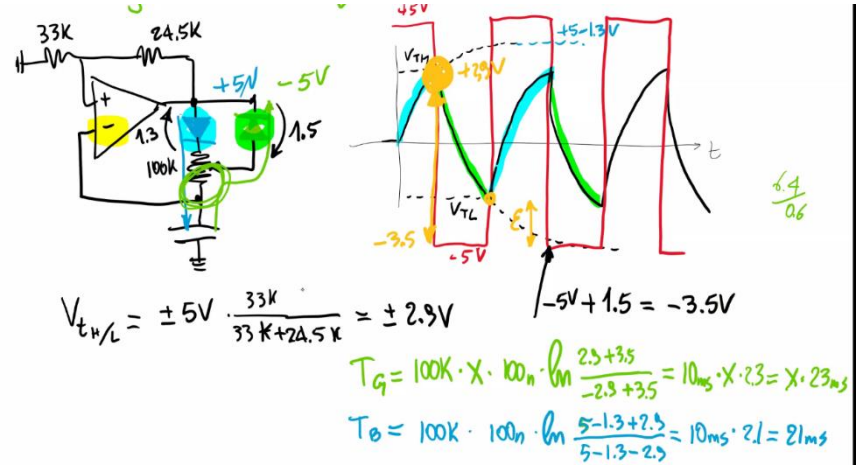


Once the capacitor has reached Vth, the asymptotic condition is not when we reach -5V, but -5V + 1.5V of the green LED, so it is -3.5V. Hence the delta is the distance between where we start (Vth) and what we reach (3.5V) (the plot is wrong, the swing is not +-5V).

In computing Tg we have to remember that the potentiometer changes its position, so we multiply its complete resistance 100k by x. Moreover, in this case the error is the difference between Vtl and the asymptotic value.



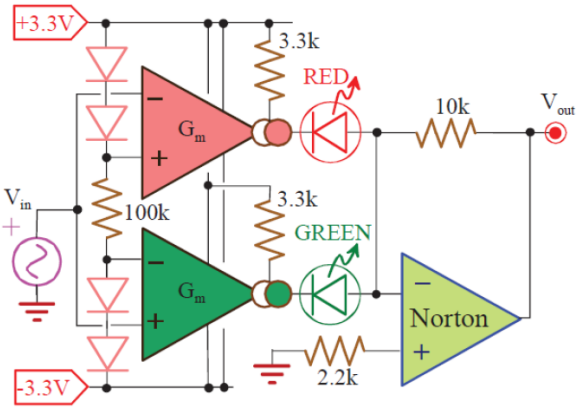
We can now compute Tb.



So this circuit is an oscillator with the following output waveform.

We have a duration that is constant and set by the blue led, while the low duration is variable, it can be long or decreased down to zero. So in this circuit the period is variable and also the oscillation frequency.

Ex. 15



OTAs with control pin at 0V. Diodes with on-voltage of about 0.7V and LEDs with on-voltage of about 1.8V. Norton amplifier with A_i=10.

a) Draw the light intensity of both LEDs vs. the V_{in} input voltage across the -3V ÷ +3V range.

b) Compute the V_{out} output voltage vs. V_{in}.

We have two OTAs and a Norton amplifier. We start from +3.3V and then we have the voltage drops of the two diodes, so at the + terminal of the upper OTA we have 1.9V, while at the – terminal of the bottom OTA we have -1.9V.

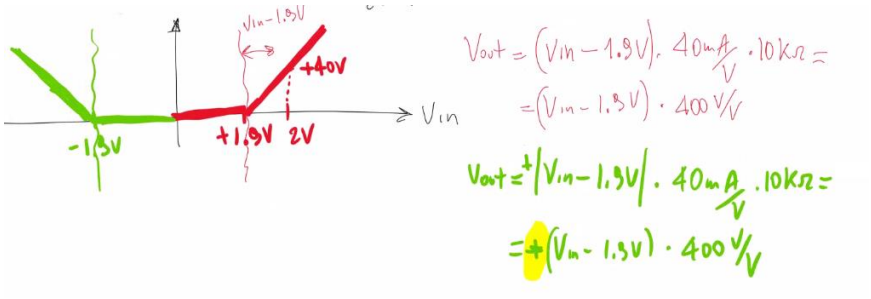
Gm|OTA = Icontrol/Vth = Icontrol/25mV = 3.3V/3.3k/25mV = 40 mA/V

Then the second stage is a Norton amplifier, but let's start by considering it a normal VOA. The red and green led allow current only from right to left and not vice versa. The red OTA allows such current if – is higher than + terminal, so Vin > +1.9V.
In this case, Vout = (Vin – 1.9V) * 40mA/V * 10k

So every time Vin goes beyond +1.9V, then the voltage difference gets amplified by the factor above that is 400 V/V. So if Vin > 1.9V the output is increasing with a slope of 400 V/V. For instance, if Vin = 2V, the output should be Vout = 40V, so probably the opamp saturates.

If Vin < 1.9V then the output is 0V.

As for the green OTA, current inward in the output is allowed if Vin < -1.9V. Still, in this case the Vout is positive. Again, if Vin = -2V, Vout = 40V.



So as soon as the OTA, either one or the other, activates, the LED is on and its light increases linearly.

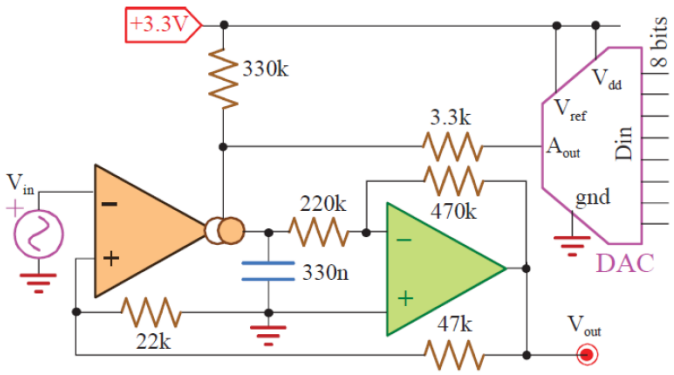
However, the opamp is not a VOA but a Norton one. If we have a Norton amplifier, we should specify the voltage of the Z node.

For a Norton amplifier:

$$i_{out} = A_i \cdot (x_+ - i_-) = A_i (\phi - i_{out} + i) =$$
$$i_{out} = A_i \cdot \lambda - A_i \cdot i_{out}$$
$$i_{out} = \frac{A_i}{1 + A_i} \cdot \lambda$$

i+ is 0 and i- is i_out – i, where i is the current either from the green or red LED. Since Vout = i_out * 10k, we end up with a very similar characteristic we found in the VOA case.

Ex. 16



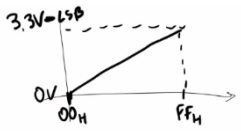
OTA with control pin at 0V and ±5V power supply.

a) Compute the OTA's transconductance as a function of D_{in}.

b) Compute the real $v_{out}(f)/v_{in}(f)$ gain, bandwidth, and stability vs. the input digital code D_{in}.

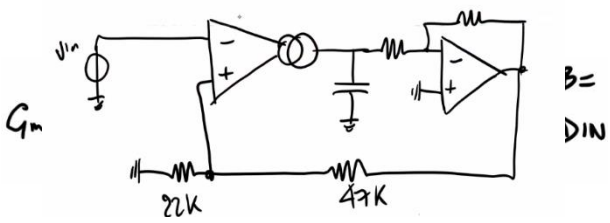
Given the Din of the DAC, the analog output Aout is the following, where Din ranges from 0 to 255, so I need to divide by 2^8.

$$A_{out} = V_{ref} \cdot \frac{D_{IN}}{2^8} = \frac{3.3V}{256} \cdot D_{in} =$$
$$= 12.9mV \cdot D_{in}$$

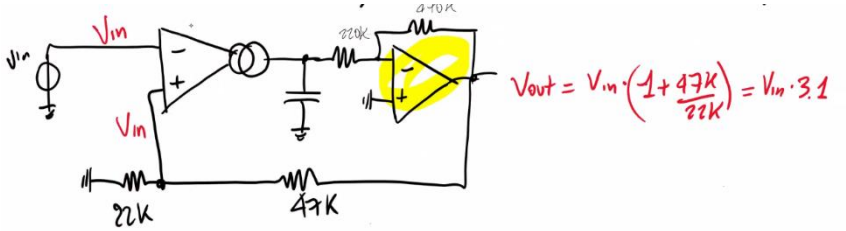


Now we can compute the Gm of the OTA, and Icontrol is the sum of the current from the PS and the current from the DAC. There is no current from the DAC that goes in the PS because the output of the OTA is at ground.
This spans with respect to Din being the highest or the lowest value.

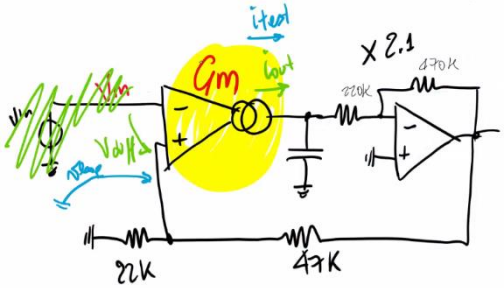
Let's now simplify the circuit and analyze it.



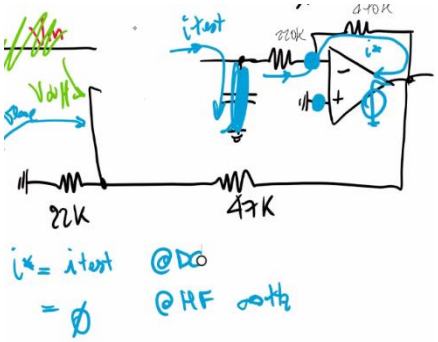
The second opamp is an inverting configuration with gain -2.1, and then we have another feedback in the OTA, but the feedback is negative.



Ideally, the gain is not depending on the Gm of the OTA. Now we want to compute the real gain and study the stability.
As A(s) let's consider the OTA, and all the rest is beta. In this case, A(s) is not the A of an opamp, but it is the Gm, i_out/v_diff. beta is wat remains, so v_diff/i_out = v_loop/i_test.



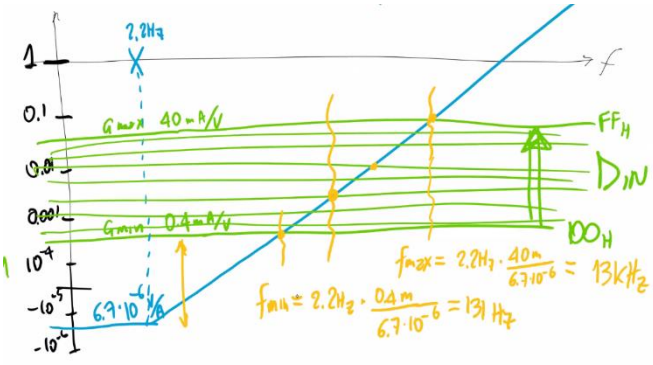
Let's study the beta. + and – terminals of the opamp are at GND and VG. We can say the following.



Then we have a pole set by the capacitor. Fp = 1/(2*pi*R*C) = 2.2Hz
So beta(inf) = 0 because of the pole. Instead, beta(0) is the i_test flowing in the 470k resistance and then in the input of the opamp.

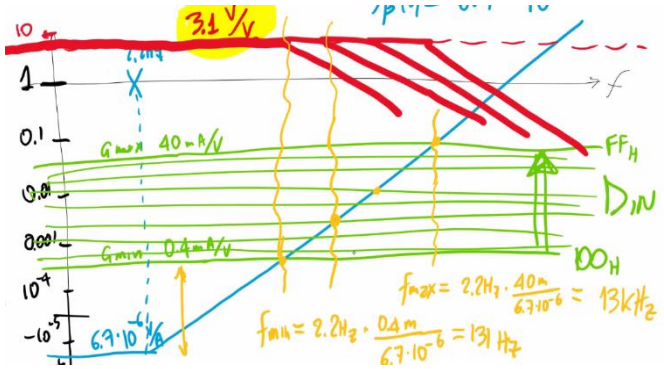
$$\beta(0) = 470K \cdot \frac{22K}{22K+47K} = 14990$$

Let's plot now. Green is the Gm, blue the 1/beta. Unfortunately, the Gm varies because of the DAC, it can be 0.4 mA/V or 40 mA/V, but it is constant. Instead, 1/beta has a zero at 2.2Hz.



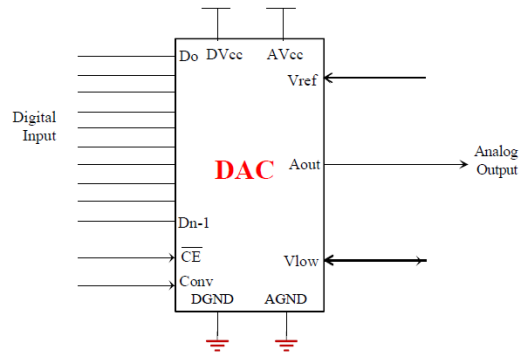
By varying the digital input code f* changes, so we have a minimum bandwidth and a maximum one.

The ideal gain is 3.1, so the real gain is the ideal one and it drops down at the corresponding f*.



We notice that by varying Din we can change Gloop.

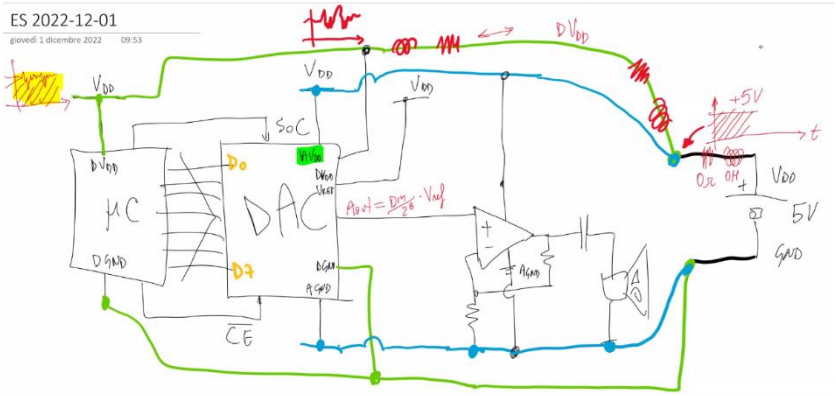
DAC



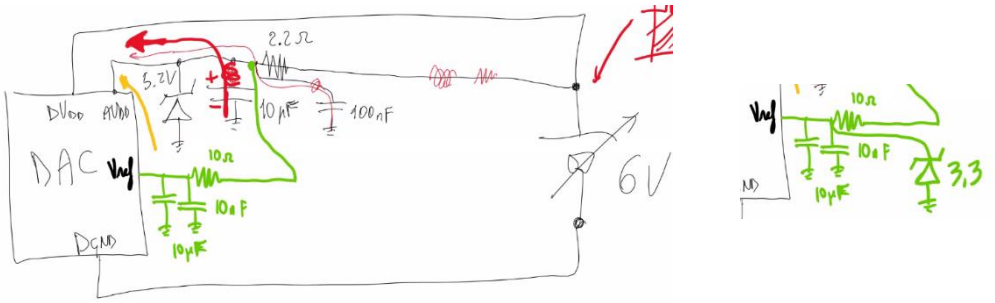
Number of bit:	n	8	16
Number of levels:	2^n	256	65536
$F_{ull}S_{cale}R_{ange}$:		5V	5V
Resolution:	$L_{east}S_{ignificant}B_{it} = FSR/2^n$	19.5mV	76µV
	$1/2^n$	$3.9^0/_{00}$	15ppm

Typically has both digital and analog components inside, therefore we have two different PS, so that one is not disturbing the other.
The need of two different PS is because typically the PS introduces disturbances due to the components attached to it. And if I drive with the same Vdd the analog and digital worlds, the disturbances are affecting both, and the quality of the output.

We have to keep the ground as much separated we can to avoid disturbances, so we connect all the digital parts of the circuit together and the analog ones together. Both Vdd and GND. To keep them separated, we try to reach the battery in just one position, as close as possible to it. Thus, any demand of current of the digital Vdd will experience voltage drops (parasitic resistances) but not at the battery level, so the analog world is preserved.

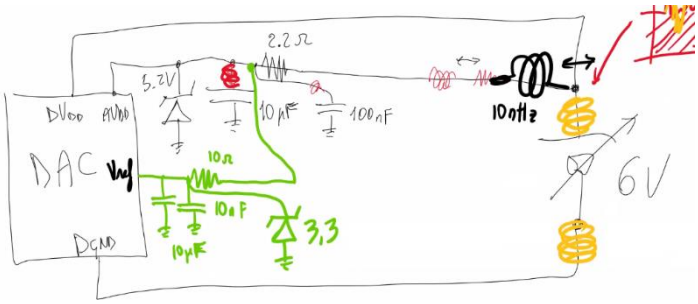


Also Vref should be theoretically filtered as much as possible and not connected directly to the PS. So what we do is the following. In fact, we add capacitor to filter the possible noise due to parasitic resistances and eventually we could also include a very small value of the resistance (2.2 Ohm) to LP filter the disturbances.
If the PS is variable, we can include another precaution, that is a Zener diode, so that if PS increases to much, the voltage at the analog Vdd will be fixed but the Zener. Moreover, electrolytic capacitors with high value have a very high parasitic inductance, so usually in parallel to a high value capacitor we place a low value capacitor to reduce as much as possible the inductor contribution.



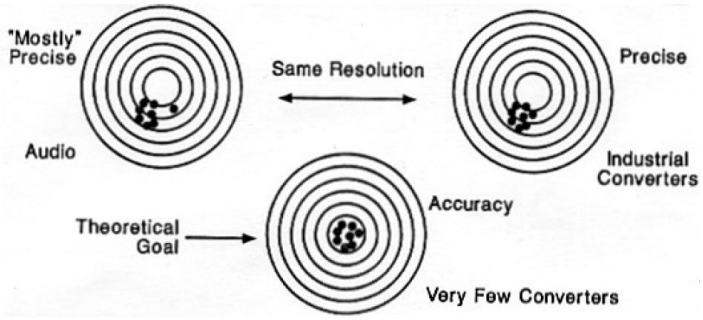
Moreover, we have also Vref that is better not to connect directly to analog Vdd, but to further LP filter it. If we want Vref to be very very stable we can also include a Zener diode.

To avoid any possible disturbances, e.g. HF fluctuations from the PS, we also introduce a specific inductance with a low value so that the HF disturbances won't pass to the other side (yellow are parasitic inductances).



Every time we have a DAC we have the CE pin, that makes the chip work or not. This is for power consumptions reasons, so that it works only when needed.
Moreover, the digital input could change and we have the Conv pin (or SoC), that is a pin that allows for the conversion of the input only when needed and asked by the microcontroller, not always.

RESOLUTION, PRECISION AND ACCURACY



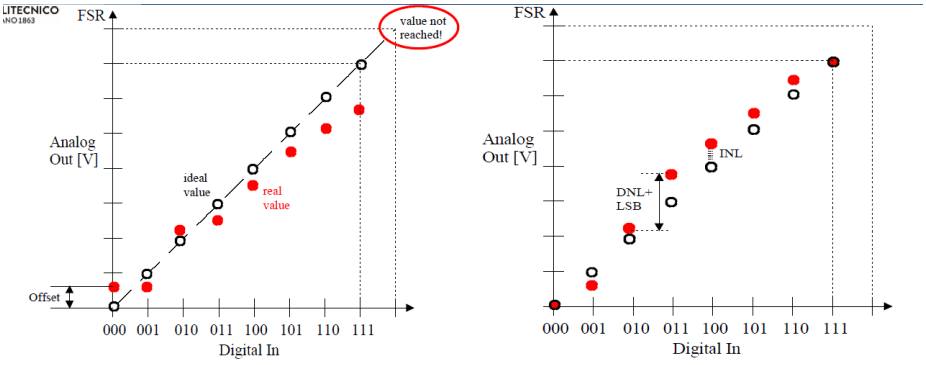
Resolution subdivision of output dynamics that the converter is able to resolve

Precision spread of output valuer, when the input is always the same

Accuracy maximum error between output analog value and theoretical expected one

In the case of DAC, the resolution is given by the number of bits. The precision is the maximum value minus the minimum one divided by 2. But there may be also another precision, not of different thermometers measuring the same quantity, but with the same thermometer on the same sample at different times.

ERRORS AND NON-LINEARITIES



Offset: output when input is 000

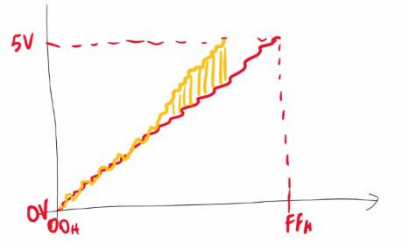
Gain: output when input is 111

INL: real-ideal distance

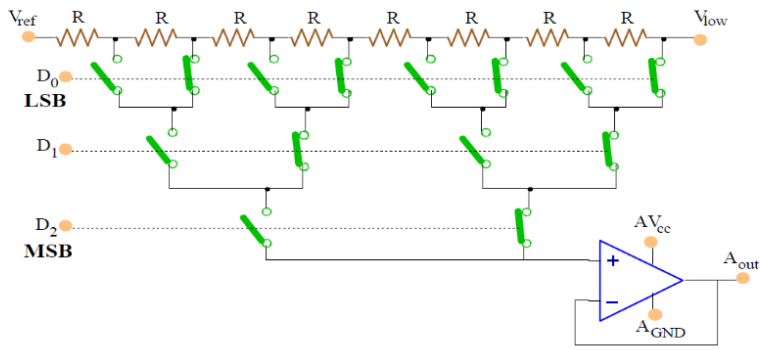
DNL: real-ideal step height

In the DAC we quote the input vs the output to see the errors. We have 8 bits and plot the value of the analog output, that is quantized. Since there are 8 steps and we want the first value to correspond to 0, the last point will be one LSB lower than FSR.
Due to errors in the DAC, if we apply 000 and the output is not 0, this error is the offset. Instead the gain error is measured when the input is 111 and the output is not FSR - LSB.
Even if we don't have these errors, we have some differences between the real and ideal behaviour. Given a specific digital input and the actual analog output, the distance between the ideal value and the real one is called INL. Instead, the distance from one point to the other should be 1 LSB ideally, but it is not always like this, so we have LSB +- DNL. This is obtained by measuring the distance between one value and the other, so the step.
If DNL goes below one LSB it means that there is a missing code, and the same if DNL > LSB.

Usually the manufacturer quotes the sigma of the DNL distribution, and the same applies for the INL.
The INL increases to much if the conversion curve, instead of being an ideal 45° one, diverges.



VOLTAGE SCALING DAC



Components: 2^n resistors and 2^{n+1} pass-transistors (i.e. 2^{n+2} MOSFETs)
1 OpAmp

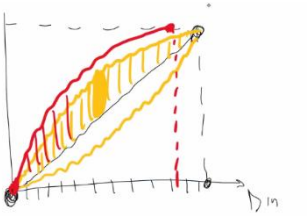
Advantages: easy scalability of resistors (all identical)
but OpAmp I_B current causes non-linearity issues

The idea is that we want to create all the possible combination of voltages, and in the case of n bits we need to introduce 2^n resistors. Then we have a decoder; the decoder could be an analog MUX where we provide in input the three digital lines, but it is very expensive. The smartest solution is to use switches, starting from 2^n, then (2^n)/2 and (2^n)/4 for each line.

The problem is that it requires 2^n resistors (big area) and also 2*2^n switches.

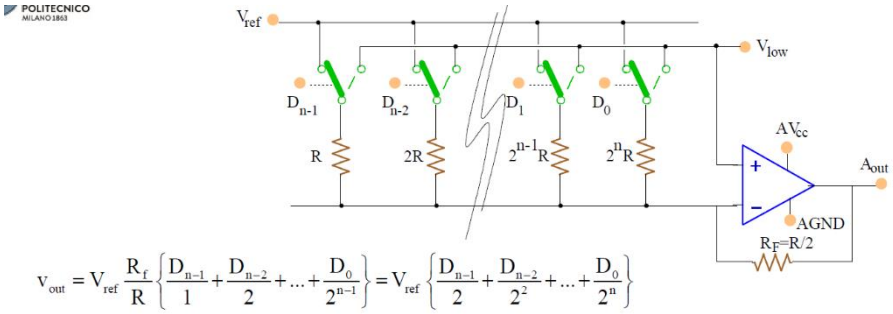
The transistors must be all pass-transistors, so one issue is the high number of transistors. Another issue is the bias current of the opamp. If we have 000 we drink the Ibias current from ground, but each switch has a Ron that causes an offset error when multiplied to the Ibias current. Moreover, if we apply 111 we see and error I*R, because R||7*R is R. But if we apply 100, we are in the middle range and we see an impedance that is the number of total R divided by 4.
So the bias current causes an error that is not constant with the input.

Moreover, watch transistor has its Ron, that is in series with a variable resistance due to R. Hence the actual i/o curve will be something like the red one, with a very bad INL and bad INL.



Then the offset of the opamp shifts rigidly this curve up or down.

WEIGHTED-R DAC

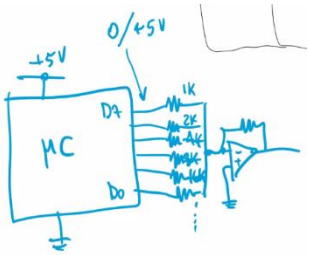


Components: n resistors (not quite…) and 2·n MOSFETs (p-MOS to V_{ref}, n-MOS to V_{low})
1 OpAmp

Problems: large silicon area tolerance of resistors OpAmp I_B

We are using resistances that are one the double of the other. There is a mistake in the image because the last resistance should be 2^(n-1) * R.
Then we use a deviator or just a switch. When we close the switch, we force a current to flow in the branch where we have the resistance. The highest current flows in the branch with the smallest resistance, so in the highest bit. Then the currents are fed to a current adder.

Another possible implementation of the upper circuit is the following one. If the uC is biased at 0-5V, the Di are either at 0 or 5V, so we can design the DAC directly out from the uC, creating it with resistors.
The resistance of the current adder in feedback should be half of the resistance of the LSB, so that the gain is half.



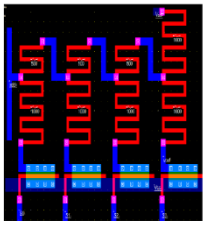
So the Weighted-R circuit seems good, because we have less switches, but the problem is that the resistors, to be multiple one of the other, should be a copy and paste several time of the smallest resistance, so we have a huge amount of resistances.

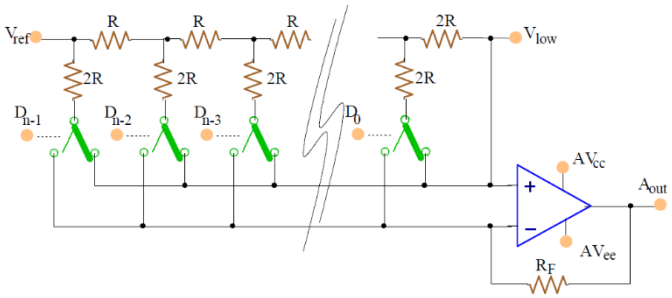
Conversely, switches are now very easy, because they should connect just +5V to the node, so it can be just a p-channel transistor. So we just need 8 p-transistors.

As for the offset voltage (at the + terminal), the gain of the circuit depends on the switches that are closed (and then we have the non inverting gain). So more or less if all switches are closed we see half R, so the gain is 2 → the bad thing is that the gain changes depending on the resistance and switches that are closed, so the offset depends on the resistance. This is the reason why we don't use just a p-channel transistor but a deviator, such that the resistor can be eventually connected to ground or power supply so that the offset sees always the same gain.

Moreover, each switch has its Ron, that is significant if the R is small, so the Ron causes an error proportional to Vref.

CURRENT SCALING DAC

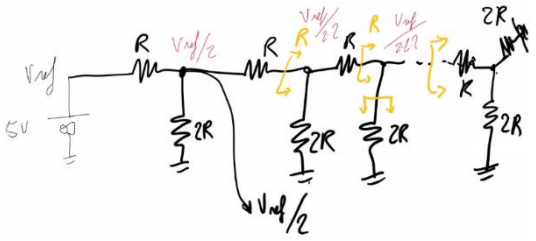




Components: 2·n resistors (3·n resistors) and 2·n MOSFETs (only n-MOS)
1 OpAmp

Advantages: easy scalability of resistors (just two different types)
easy driving of MOSFET switches

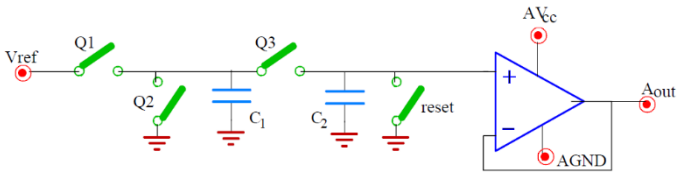
We start from Vref (e.g. 5V) and if we need half of it we take the middle node of a voltage divider with R and R. Then, to subsequently divide, we should do the following.



Then we take the current if we need (connect to – terminal, VG), or if we don't need it, we deviate to actual GND. Rf should be equal to…?
As for Ron, it doesn't play any effect because 2R is much larger, so we are just adding a constant error and we can add a resistance Ron in series to the horizontal R to compensate for this error. As for Ibias+ and Ibias– and Vos will be asked at the exam.

The switches can also be n-channel mosfets on both sides.

SERIAL INPUT DAC



Components: only 2 capacitors and 5 MOSFETs (1 p-MOS, 2 n-MOS, 1 passtransistor)
1 OpAmp

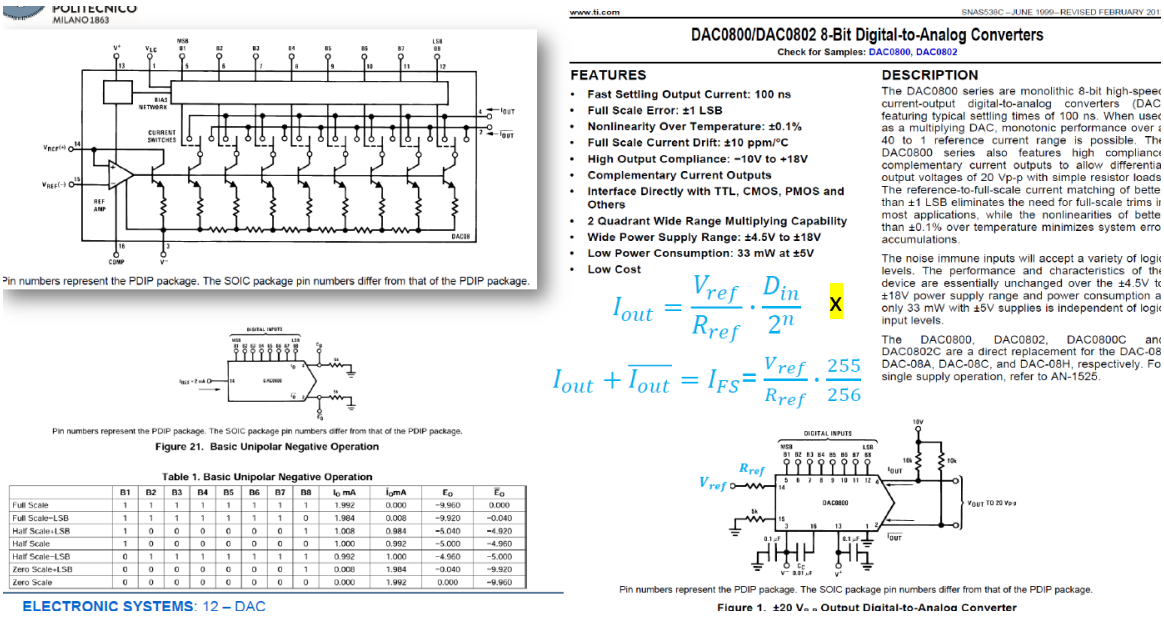
Advantages: extremely compact and easy, but needs serial input (bit by bit)
MSB first? How many bits?

This is a very smart solution. We apply one bit at a time, not the full parallel bus. With 0 we apply 0V, with 1 I apply Vref to the capacitor, and we have two capacitors performing charge sharing. At start, Q1 and Q2 are open and Q£ and Q4 are closed, because we reset.
Now we open the reset, Q3 and we apply the first bit. If 0, Q2 is closed and C1 charged to 0. Then we open and we close Q3 so that charge is shared with C2. Then Q3 opens. If the second bit is 1, Q1 closed and Q2 open, so we store 5V on C1 (and C2 was at 0V). Then we open and close Q3, so we share 5V on C2 and they go to the average value between the two, that is 2.5V. Then Q3 opens and if the next bit is a 0 we close Q2, C1 is charges to zero. Then we open Q1 and Q2 and we close Q3. 2.5V were on C2, C1 was at 0V so we have 1.25V on both.

In the end, the Aout is the analog conversion of the digital input stream. Because of this, the first bit we should apply is the one that experiences more divisions, so the LSB, while the last one is the MSB.

This DAC has 4 bits (4 switches) and two capacitors.

MULTIPLYING ADC



POLITECNICO MILANO 1863

Pin numbers represent the PDIP package. The SOIC package pin numbers differ from that of the PDIP package.

Figure 21. Basic Unipolar Negative Operation

Table 1. Basic Unipolar Negative Operation

	B1	B2	B3	B4	B5	B6	B7	B8	I_O mA	$I_{\overline{O}}$mA	E_O	$E_{\overline{O}}$
Full Scale	1	1	1	1	1	1	1	1	1.992	0.000	−9.960	0.000
Full Scale−LSB	1	1	1	1	1	1	1	0	1.984	0.008	−9.920	−0.040
Half Scale+LSB	1	0	0	0	0	0	0	1	1.008	0.984	−5.040	−4.920
Half Scale	1	0	0	0	0	0	0	0	1.000	0.992	−5.000	−4.960
Half Scale−LSB	0	1	1	1	1	1	1	1	0.992	1.000	−4.960	−5.000
Zero Scale+LSB	0	0	0	0	0	0	0	1	0.008	1.984	−0.040	−9.920
Zero Scale	0	0	0	0	0	0	0	0	0.000	1.992	0.000	−9.960

www.ti.com SNAS538C – JUNE 1999 – REVISED FEBRUARY 201

DAC0800/DAC0802 8-Bit Digital-to-Analog Converters

Check for Samples: DAC0800, DAC0802

FEATURES

- Fast Settling Output Current: 100 ns
- Full Scale Error: ±1 LSB
- Nonlinearity Over Temperature: ±0.1%
- Full Scale Current Drift: ±10 ppm/°C
- High Output Compliance: −10V to +18V
- Complementary Current Outputs
- Interface Directly with TTL, CMOS, PMOS and Others
- 2 Quadrant Wide Range Multiplying Capability
- Wide Power Supply Range: ±4.5V to ±18V
- Low Power Consumption: 33 mW at ±5V
- Low Cost

DESCRIPTION

The DAC0800 series are monolithic 8-bit high-speed current-output digital-to-analog converters (DAC featuring typical settling times of 100 ns. When used as a multiplying DAC, monotonic performance over a 40 to 1 reference current range is possible. The DAC0800 series also features high compliance complementary current outputs to allow differential output voltages of 20 Vp-p with simple resistor loads. The reference-to-full-scale current matching of better than ±1 LSB eliminates the need for full-scale trims in most applications, while the nonlinearities of better than ±0.1% over temperature minimizes system error accumulations.

The noise immune inputs will accept a variety of logic levels. The performance and characteristics of the device are essentially unchanged over the ±4.5V to ±18V power supply range and power consumption at only 33 mW with ±5V supplies is independent of logic input levels.

The DAC0800, DAC0802, DAC0800C and DAC0802C are a direct replacement for the DAC-08, DAC-08A, DAC-08C, and DAC-08H, respectively. For single supply operation, refer to AN-1525.

$$I_{out} = \frac{V_{ref}}{R_{ref}} \cdot \frac{D_{in}}{2^n}$$ X

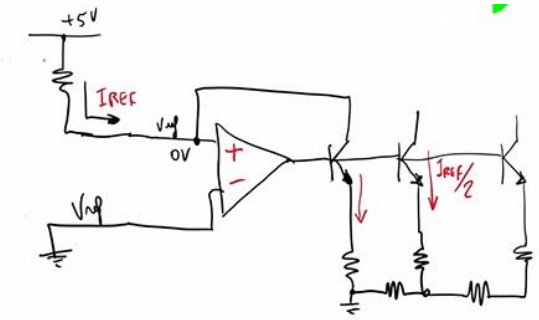
$$I_{out} + \overline{I_{out}} = I_{FS} = \frac{V_{ref}}{R_{ref}} \cdot \frac{255}{256}$$

Pin numbers represent the PDIP package. The SOIC package pin numbers differ from that of the PDIP package.

Figure 1. ±20 V$_{p-p}$ Output Digital-to-Analog Converter

We have an opamp that drives BJT transistors. The feedback must be connected to the + terminal, otherwise we get a positive feedback. On the – terminal we can place the Vref, e.g. ground, and so +

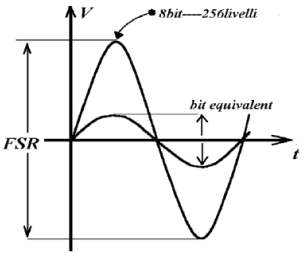
terminal is at ground. So we can set the current Iref that will flow in the transistor. Now we can take the same transistor and the same resistor and have the same current. If I want a Iref/2, we add a resistance. We have also to add deviators as switches. Then Iout will be the digital conversion of the currents.



The switches are current steerer. So we have a pin with the sum of the current, and also another, one is Iout, the other Iout-not.

If we buy this DAC, Vref- is connected to Vref, e.g. zero, and on the other pin we can connect whatever we wish to force Iref, that is Vin/Rref. Then Iout is given by equation x.
It's called multiplying DAC because Vref can vary, and if Vref varies, current varies and so all the currents vary. It performs the multiplication of an input signal analog voltage with an input digital bus.

DYNAMIC PERFORMANCES



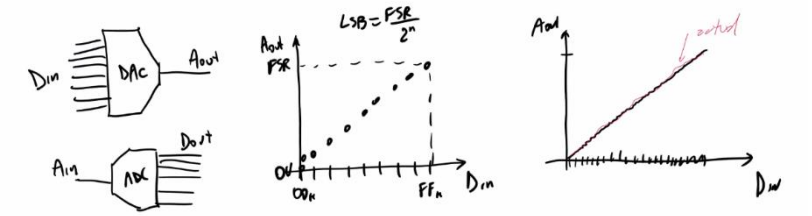
Dynamic range: $D = 20 \cdot \log \frac{FSR}{LSB} = 20 \cdot \log 2^n = 6.02 \cdot n$

Quantization Noise: $\sigma^2 = \int_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} \varepsilon_q^{\,2} \cdot \frac{1}{LSB} \cdot d\varepsilon_q = \frac{1}{LSB} \cdot \left[\frac{\varepsilon_q^3}{3} \right]_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} = \frac{1}{LSB} \cdot \frac{2}{3} \cdot \frac{LSB^3}{8} = \frac{LSB^2}{12}$

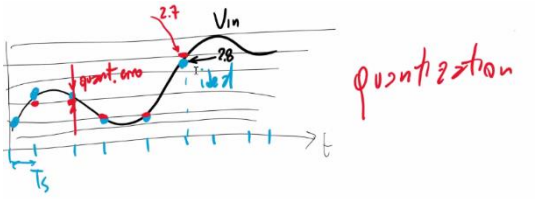
Ideal Signal-to-Noise Ratio: $SNR|_{max} = \frac{\text{max signal power}}{\text{min noise, just quantization error}}\Big|_{dB} = 10 \cdot \log \frac{(FSR/2\sqrt{2})^2}{(LSB^2/12)} = 6.02 \cdot n + 1.76$

EffectiveNumberOfBit: $n_{eff} = \frac{SNR - 1.76dB}{6.02dB}$

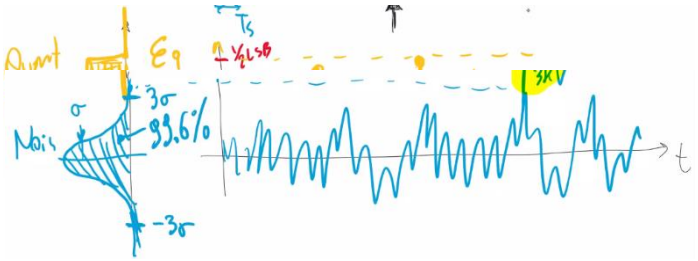
We want to know how much the converter distorces the input signal and the total errors on the input signal.
Given a converter, e.g. a DAC, we have a digital stream in input and an analog output. The idea is that we know that the digital input is quantized, so also the analog output will be quantized. If we have more than 8 bits (right), we have to compute the ideal i/o characteristic, so it is difficult to understand the errors.



So it is better to move to the digital domain instead than the frequency domain.
So we have the signal and the ADC is divided in bits. So we have the problem of quantization. Selected the sampling time Ts, **the quantization error is a deterministic error**.



We can plot this error, that can be either positive or negative or 0. If the ADC is good, so the INL is not too big, so the distance between the ideal staircase and the actual one is sufficiently limited, so lower than 1 LSB, then we can say that the quantization error is in between +1/2 LSB and -1/2 LSB.
We can also study the distribution of the deterministic quantization error. If Vin is constant it will accumulate in one position, but if Vin moves it changes, so we can assume the quantization error as randomly distributed between ½ LSB and -1/2 LSB.



We know how to deal with noise, given the gaussian curve and the sigma. In fact, between +- 3*sigma we have 99% of the total area.
The quantization error is different, but if Vin moves between different levels, I can consider this error as random and consider it as if it was a noise. So given the error that is at most +- ½ LSB, at which noise level corresponds?
The peak to peak value of the error is 1LSB, and the peak to peak value of the sigma is 6*sigma. So the quantization noise is equivalent to a noise sigma = LSB/6.

Let's compute now the power, not the peak to peak value, of the quantization error across the +- ½ LSB range. Then we also divide by the range.

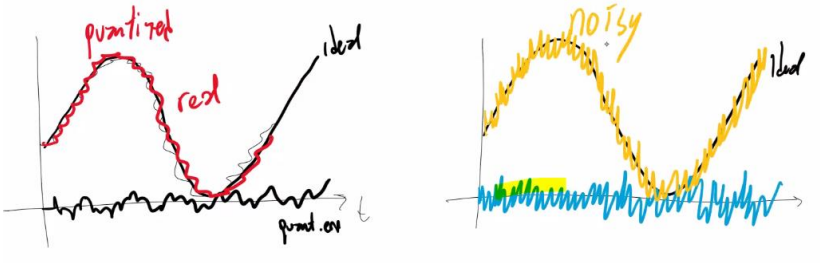
$$\frac{1}{LSB}\int_{-\frac{1}{2}LSB}^{+\frac{1}{2}LSB} \varepsilon_q(x)^2\, dx$$

The result I get is the one in the slide, so LSB^2 / 12. But we know that the power of the noise is sigma squared, so I can say that:

$$\frac{LSB^2}{12} = \sigma^2$$

So if the introduced error is LSB, the equivalent noise is LSB^2 / 13. Basically we get an ideal signal plus a quantization error that gives me the real signal, but this can be also seen as an ideal signal plus the noise (right).

Then noise has a gaussian distribution whose sigma is sigma = sqrt(LSB^2 / 12). If the quantization error is +- ½ LSB the two are the same.



Moreover, if a signal has a peak value that is Vp, the power of the signal is the following.

$$Power_{signal} = \left(\frac{V_p}{\sqrt{2}}\right)^2$$

I also know the power of the quantization error because I can describe it as equivalent noise.

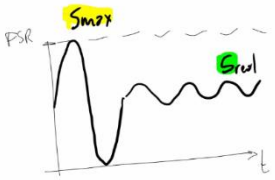
$$Power_{Bad\ stuff} = Power_{quant_g} = Power_{noise} = \frac{LSB^2}{12}$$

Lastly, I can also define the SNR. The ideal one is the one with the maximum signal, when it reaches FSR, and the minimum noise. The minimum noise is instead when the noise is just due to quantization, so LSB^2 / 12. But LSB = FSR/2^n.

$$SNR_{ideal} = SNR_{max} = \frac{S_{max}}{N_{min}} = \frac{\left(\frac{FSR}{2\cdot\sqrt{2}}\right)^2}{N_{quant_g}} = \frac{\left(\frac{FSR}{2\cdot\sqrt{2}}\right)^2}{\frac{FSR^2}{12\cdot 2^{2n}}} = \frac{12\cdot 2^{2n}}{4\cdot 2}$$

$$= \frac{3}{2}\cdot 2^{2n}$$

Given my dynamics, so the FSR, we can have a theoretical maximum signal but we have also an ideal signal. So we have to introduce also the concept of theoretical SNR, that is the real signal over the minimum noise.



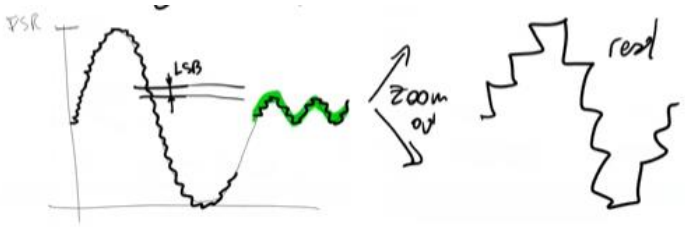
$$SNR_{theoretical} = \frac{S_{real}}{N_{min}}$$

This is nothing else than the SNR_ideal minus delta_S, where delta_S is the ratio between the maximum signal and the real one (S_max/S_real).

$$SNR_{theoretical} = \frac{S_{real}}{N_{min}} = SNR_{ideal} - \Delta S$$

If S_max = S_real, delta_S = 0 dB and the signal is the maximum one. If instead the signal maximum has a maximum value of 5V peak to peak and the S_real moves by just 50 mV, then I'm doing it wrong because I could have exploited the full dynamics of the DAC, and the quantization is worse because I'm using just a portion of the FSR, just 50 mV compared to 5V.
Hence delta_S = 5V/50mV = 1/100 = -20 dB, so I'm loosing -20 dB.

So we are applying the LSB to a much smaller sinusoid.



So the real quality is worsen with respect to the case in which we amplified the signal before to use the FSR.
Moreover, the electronics could be noisy, so on top of the real curve there may be also noise. Hence I also need to define a SNR real (Nreal is the real noise, that is the noise due to quantization plus extra noise).

$$SNR_{real} = \frac{S_{real}}{N_{real}} = \frac{S_{ideal} \cdot \Delta S}{N_{prima} + \Delta N} = SNR_{ideal} - \Delta S - \Delta N$$

So on top of the signal there is an extra noise and if we use for instance an 8 bit DAC, the SNR_ideal is:

$$SNR_{ideal} = 6.02 \cdot n + 1.76 = 50dB$$

But then the real is just 50 mV peak to peak, so the SNR_theoretical is less. Since there is also some real noise, I'm even more unlucky.

$$SNR_{theo} = SNR_{ideal} - \Delta S = 50dB - 20dB = 30dB$$
$$SNR_{real} = SNR_{ideal} - \Delta S - \Delta N = 30dB - 15dB = 15dB$$

This means that me could have entered the ADC or DAC with full dynamic, with an ideal quality that results in a SNR_ideal = 50 dB but instead no, because I'm using a smaller signal; the steps are equal but the quality is worse. Even worse, the actual real output has the extra noise.

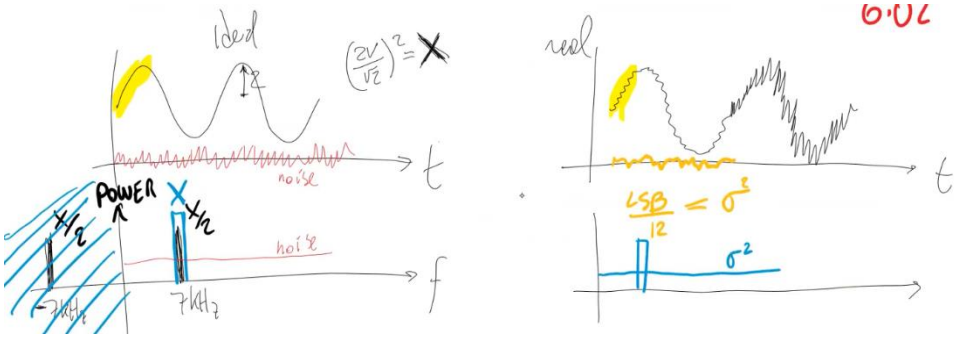
So if I use a 8 bit ADC and then a low amplification we are not properly using it.
So the question is: how many bits should an ideal ADC or DAC have to have a good SNR? When the ideal SNR is equal to the real one?
The equivalent number of bits will be the following.

$$SNR_{ideal} = 6.02 \cdot ENOB + 1.76 =)$$
$$ENOB = \frac{SNR_{real} - 1.76}{6.02} = 2.3$$

Moreover, usually the signal is in the time domain; the real signal will be quantized and maybe even noisy. How can we say how much the ideal signal is good with respect to the real one?

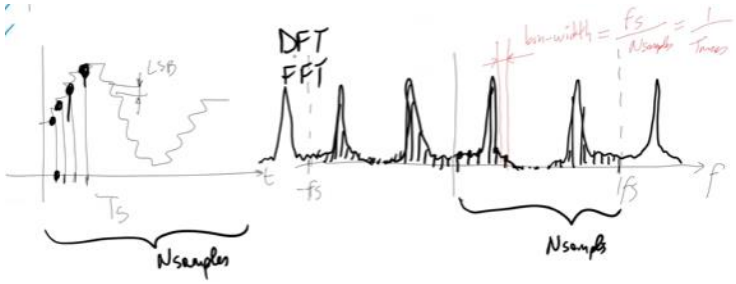
Instead of remaining in the time domain, let's move to the frequency domain. A sinusoid in the time domain is a delta in the frequency domain; if the frequency of the sinusoid is 7 kHz, we have two deltas at + and – 7 kHz.
If the peak voltage of the sinusoid is 2V, the power will be (2V/sqrt(2)^2. If we compute the spectrum of the sinusoid, the deltas have an height of power/2 (half at +7 kHz and half at -7 kHz).

We can consider just positive frequencies and saying the sinusoid has a power of x at +7 kHz.
If I have noise on the sinusoid, the spectrum of the noise is a constant plateau over all the frequencies.



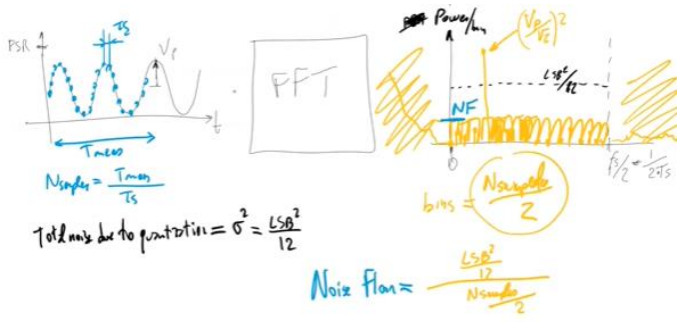
As for the real signal, I have quantization, and quantization error is deterministic (yellow one), but the quantization LSB can be considered as a noise whose sigma-squared is LSB/12.
Hence the spectrum of the real signal will be the spectrum of the sinusoid and then the power of the quantization error will be equal to sigma-square.

Moreover, if we have a time continuous signal, the spectrum goes to infinity. If instead we have a time discrete signal, the spectrum is periodic.

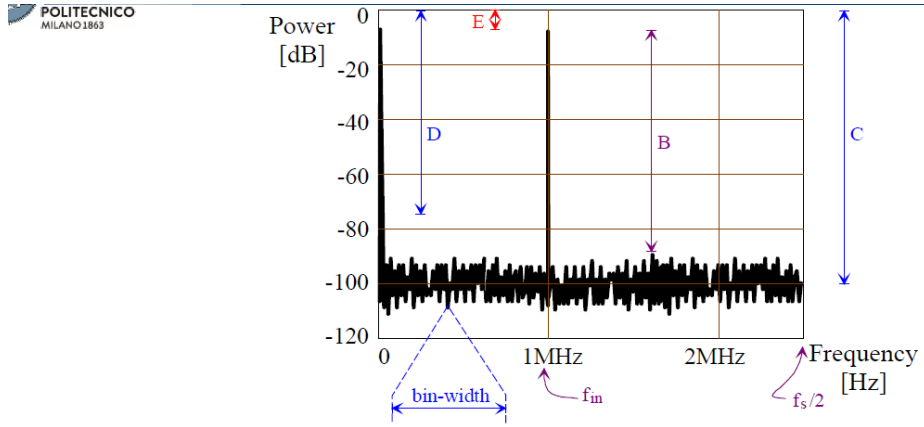


T_measure is the time we take to measure the signal. If we have a long T_measure, the histogram is very well defined.
If the signal is real (not complex number), then the spectrum is symmetric around fs/2, so it is enough to study the spectrum between 0 and fs/2.
Furthermore, it is better to consider the vertical axis in the frequency domain not the power but the power per bin.

Finally, given a signal with its FSR, we measure for a time T_meas and the distance between samples is Ts. Then let's use a FFT algorithm and in the frequency domain we plot from 0 to fs/2. The number of samples inside will be Nsample/2 because I'm not using the full fs. The peak in the frequency domain will be equal to the power of the sinusoid in the dime domain.
However, we have also the quantization error. The total noise due to quantization is LSB^2/12, and since we are plotting the power per bin, since the quantization error is a white noise and I want to compute the noise in each bin, and the bins are Nsample/2, the quantization noise will be (LSB^2/12)/(Nsamples/2).
So the noise is divided into many histograms and the height of each histogram is called Noise Floor NF.



SPECTRAL PERFORMANCES

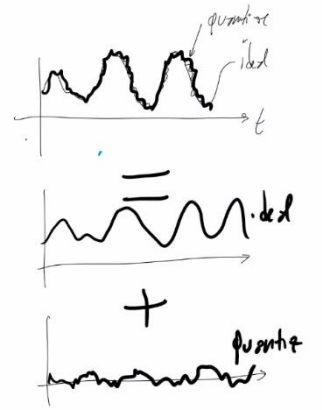


Total noise (only quantization error): $\sigma_q^2 = \frac{LSB^2}{12}$

Bin-width: $bin - width = \frac{f_S}{N_{samples}}$ **Noise in each bin:** $\sigma_q^2 / N_{samples} = \frac{LSB^2}{12} / N_{samples}$

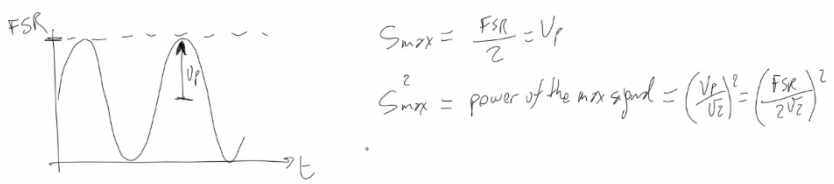
... hence $Noise\ Floor = -\left\{6.02 \cdot n + 1.76 + 10 \cdot \log\left(\frac{N_{samples}}{2}\right)\right\} = -SNR|_{teorico} - 10 \cdot \log\left(\frac{N_{samples}}{2}\right)$

We have a signal that we don't want to study in the time domain but in the frequency domain. The real signal can be seen as the superposition of the ideal signal and the quantization error.



We know how to deal with noise so the idea is to consider the quantization error as if it was due to noise. Once we know the LSB we can quote the variance of the equivalent noise.
Now we want to plot the power of the signal within each histogram, and the sinusoid is within a given bin width at the frequency of the sinusoid. Conversely, the total power of the noise is along all the frequencies. But since we want to describe the power due to all contributions, since the number of samples is given, the height of the sample is the noise floor, which is the full power divided by the number of samples.

If we change the FSR, the LSB changes and so maybe it is better to plot the power in terms not of V^2, but in terms of maximum possible signal.
We consider the maximum input signal we can apply within the FSR of the converter. The maximum peak value is half the FSR. If we want to compute the power of the maximum signal S_max, we have the following.



The idea is now to perform a sort of normalization, saying, from now on, that S_max^2 = 0 dBc (dBc = dB carrier).
If so, we can define the SNR_ideal that is S_max/Nmin. In terms of power, it is S_max^2/(LSB/12). If we substitute S_max:

$$SNR_{ideal} = \frac{S_{max}}{N_{min}} = \frac{S_{max}^2}{\frac{LSB^2}{12}} = 6.02 \cdot n + 1.76$$

This means that the noise minimum Nmin is S_max^2/SNR_ideal.

$$N_{min}^2 = \frac{LSB^2}{12} = S_{max}^2 / SNR_{ideal} = S_{max}^2 - \left(6.02 \cdot n + 1.76\right)$$

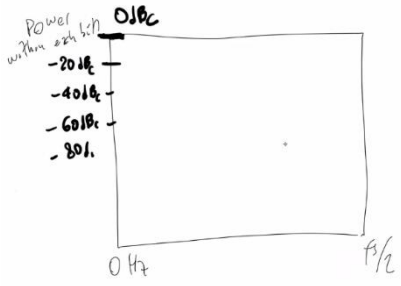
$$N_{min}^2 = \frac{LSB^2}{12} = -\left(SNR_{ideal_{dB}}\right)$$

If signal is in dBc and noise is in dBc, SNR = S/N means:

$$SNR = 0dB_c - (-50dB_c) = 50\ dB$$

So the power is in dBc but the ratio is in pure dB.
Now we can draw our spectrum, which is symmetric around fs = 1/Ts. If N_samples is the number of samples acquired in the time domain, they are also in the frequency domain, but since we plot only from 0 to fs/2, we will have just N_samples/2 in the range of interest.

The power of a signal should also be described in the negative frequencies domain, but we can forget about them, and double by two the height of the power in the positive frequencies to take into account both of them.
From now on we will use the following codification to plot.



$$0dB_c \equiv S_{max} \equiv \left(\frac{FSR}{2\sqrt{2}}\right)^2$$

If the real signal S_real != S_max, for instance 0.5Vp, we can define the attenuation delta_S = S_max/S_real. Then I convert delta_S in dB.

$S_{real} \neq S_{max} = 2.5V_p$

$S_{red} = 0.5V_p = 500mV_p$

$\Delta S = \frac{S_{max}}{S_{real}} = \frac{2.5V_p}{0.5V_p} = 5 = 7dB$

So my sinusoid won't have the maximum height, but a height that will reach – 7dBc, so 7 dBc lower than 0 dBc.
Now we have to plot also the power of the noise. Nmin = LSB^2/12 = -SNR_ideal because we defined S_max = 0 dBc.

$N_{min} = \frac{LSB^2}{12} = \ldots = -SNR_{ideal} =$

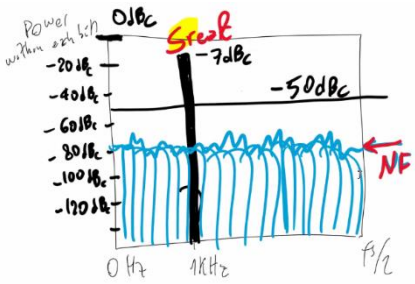
$SNR_{ideal} = 6.02 \cdot n + 1.76 = 50dB$

Hence Nmin = -50 dBc.
Moreover, we are using a given number of samples to acquire the signal, e.g. 1024. So the total power is the sum of many histograms and the height of each histogram describing the noise is called **noise floor**.

$NF = Noise\ Floor = \frac{N_{min}}{\frac{N_{samples}}{2}}$

Translated in dB: $-SNR_{ideal} - 10\log\frac{N_{samples}}{2}$

Eventually, the noise floor sets to -50dBc over 1024 samples, so -50 dBc – 27 dBc = -77 dBc. This is the value at which the noise floor sets from the peak. The blue is the noise floor.

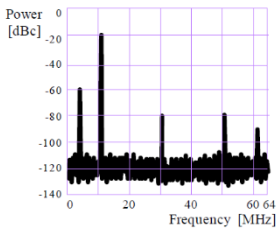


Now we can perform all the other possible computations. Given NF, we can compute the total noise reverting the computations.

$N_{real} = NF \cdot \frac{N_{samples}}{2} = NF_{dBc} + 10\log\frac{N_{samples}}{2}$

Example 1

The 12bit DAC has 5V FSR. The measured output spectrum has a 2kHz bin-width. There is a distortion tone at 5MHz.
a) Compute SNR_{ideal}, SNR_{theor}, SNR_{real}, and ENOB.
b) Compute the THD and the SiNAD.



Resolution

We know FSR = 5V, n = 12 bit, so we expect SNR_ideal = 6.02*n + 1.76 = 74 dB.
But from the plot I see that S_real = -20 dBc (dBc and not dB because I'm considering power).

$$FSR = 5V \qquad n = 12bit$$
$$SNR_{ideal} = 6.02 \cdot n + 1.76 = 74\,dB$$
$$S_{real} = -20\,dB_c$$
$$SNR_{th} = SNR_{ideal} - \Delta S$$

Delta_S = 20 dB.
Hence SNR_theoretical = 74 dB – 20dB = 54 dB.
Now I should check the Nreal (ideal noise) in the ADC (or DAC). I need to check also the NF intensity and if it corresponds to the ideal one.

The ideal noise Nideal (noise just due to quantization) is expected to be – SRN_ideal, so – 74 dBc.
The number of bins in the spectrum is Nsamples/2, and over the bins we spread the Nideal. The width of each histogram is 2 kHZ and the plot stops at 64 MHz, so fs/2 = 64 MHz, so fs = 128 Msps. In the range 128 Msps I find:

$$N_{sample} = \frac{f_s}{bin\text{-}width} = \frac{128MHz}{2kHz} = 64k\,samp$$

The total real noise is the noise floor real (in the plot) multiplied by the number of samples. From the plot, NF_real is more or less -110 dBc (in dB multiplication is the sum):

$$N_{real} = NF_{real} \cdot \frac{N_{samples}}{2} = -110dB_c + 10\log\frac{N_{samples}}{2} = -110dB_c + 45dB = -65dB_c$$

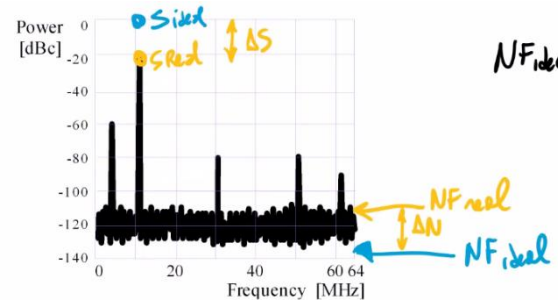
Now we can compute the real signal to noise ratio. The real signal is -20 dBc, the real noise is -65 dBc. Since we are in dBc, we are subtracting the two values, and the result is in dB (ratio between powers in dBc results in dB).

$$SNR_{real} = \frac{S_{real}}{N_{real}} = \frac{-20dBc}{-65dBc} = -20dBc - (-65dBc) = 45dB$$

The real SNR is not the ideal one because I loose delta_S and the extra noise, so I can also compute the extra noise.

$$SNR_{real} = SNR_{ideal} - \Delta S - \Delta N$$
$$\Delta N = 74dB - 20dB - 45dB = 9dB$$

So the DAC is not ideal because it has this extra noise. Theoretically we are in a situation like the following.

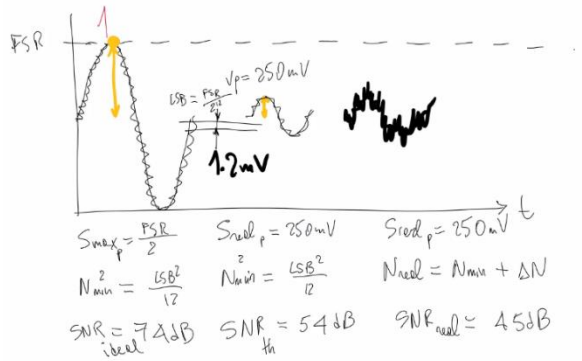


$$NF_{ideal} = \frac{N_{ideal}}{\frac{N_{samples}}{2}} = -74dB_c - 10\log\frac{N_{samples}}{2} =$$
$$= -74dB_c - 45dB = -119\,dB_c$$

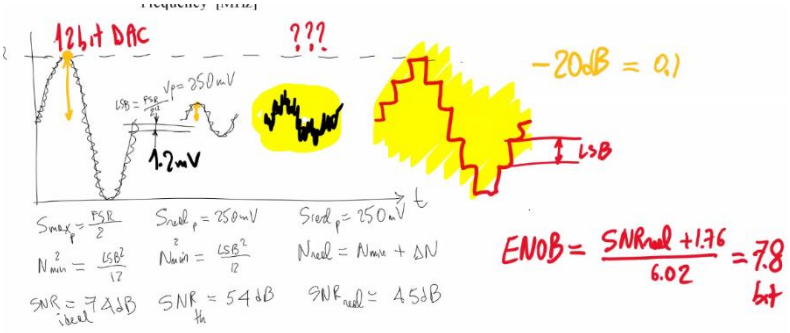
Thus I can confirm that the increment in noise is 9 dB, as already done with the previous computation.
In the time domain it means that we could have expected a signal with a certain quality, then due to quantization we have a 'staircase' signal. But in reality we enter with a much lower sinusoid whose peak is -20 dB with respect to the best one. When we shift from dB to normal amplitude we have to divide by 20, while when we work with power by 10. So 10^(20/20), that is a factor 0.1 less of the maximum peak.

Since I know the FSR of the circuit was 5V, the Vp (peak voltage) will be one tenth of the FSR. And since the LSB is constant (FSR/2^n), I can say that the quality is much poorer than before.
Because the real peak is only 250 mV and noise Nmin is the same of before, because the quantization is the same, so the theoretical SNR is smaller, 54 dB. But instead no, because the NF_real is higher than the ideal one by a delta_N of 9 dB.

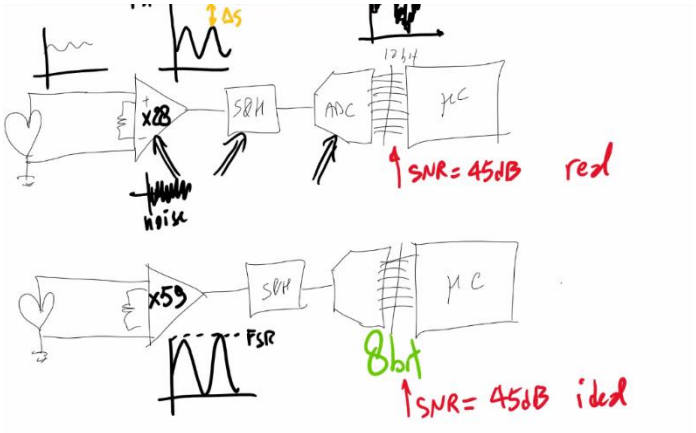
This means that the real signal is not the small one with just the quantization error, but with more noise. The real situation is the black one. Nreal will be the minimum noise plus the delta noise. This is the reason why SNR_real is less dB.



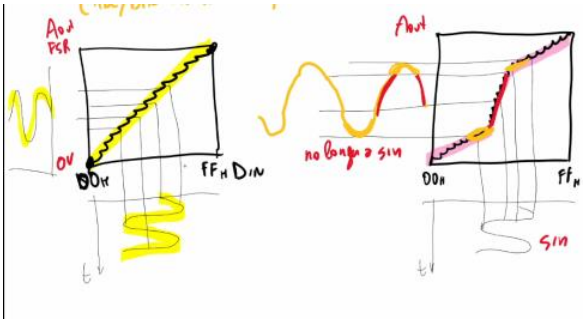
So we started with a 12 bit ADC and ended up with a quality that is bad, and that could eventually be reached with a smaller number of bits.
Hence we can define the ENOB, that is 8 bit.



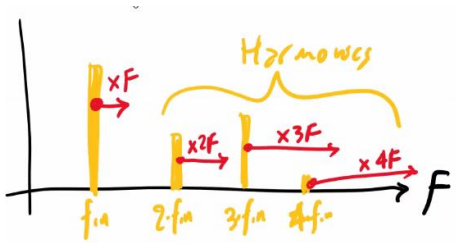
So we took the signal, we amplify it, then we go in the S&H circuit, then the ADC (e.g. 12 bit) and we reach the uC. If the amplification is not enough the SNR is poor and of 45 dB, so we don't reach the FSR. Moreover, at the end of the conversion we have a worsening due to the noise in the analog electronics and quantization error.



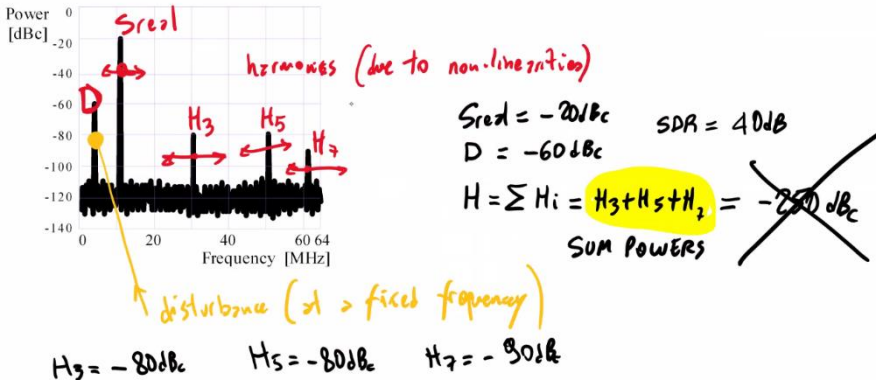
The concept of ENOB is to have an amplification higher to have a bigger signal reaching the FSR and the ADC could use just 8 bit instead of 12 because eventually what reaches the uC will have the same SNR.
Furthermore, looking given the plot, we know S_real and Nreal and SNR_real and we have also minor peaks. In the text I've told there is a disturbance at 5 MHz, that is on top of the signal and always at a fixed frequency.
The other peaks are due to **harmonic distortion, they are due to ADC or DAC non linearities**. If the converter has a i/o characteristic that should be a straight line of 45°, if we enter with a precise sinusoid, we have a perfect output sinusoid. But since the converters have some non linearities (INL and DNL), we may experience compressions and decompressions. So the output is no longer an ideal sinusoid. It is still periodic for sure. But at the same time some tones are created.



To differentiate between a disturbance and tones of the real sinusoid we have that if we change the frequency of the sinusoid, the disturbance stays fixed, while the harmonics change as below, coherently with the sinusoid.



Now we want to compute how big the harmonics are compared to the signal, so we have to compute the power of those signals.
S_real = - 20 dBc, and from the plot the distortion is at – 60 dBc. So the signal distortion ration is 40 dB.



When we sum powers, we don't have to sum dB directly, but 10^(-80dBc/10) ecc… .

$$= 10^{\frac{-80dBc}{10}} + 10^{\frac{-80dBc}{10}} + 10^{\frac{-90dBc}{10}} =$$

$$= 10^{-8} + 10^{-8} + 10^{-9} = 2.1 \cdot 10^{-8} =$$

$$= -77\,dBc$$

Now we can compute the total harmonic distortion THD, which is the harmonic divided by the S_real.

$$THD = \frac{H}{S_{real}} = \frac{-77dBc}{-20dBc} = -77dBc + 20dBc = -57dB$$

This means that given our sinusoid, the distortion is as such that the sinusoid gets impaired by 57 dB less than the signal.

Finally, we can define the SiNAD, that is the signal over the noise plus distortion and harmonics. All of them are the real terms.

$$SiNAD = \frac{S}{N+D+H} = \frac{S_{real}}{N_{real}+D+H} = \frac{-20dBc}{-65dBc - 60dBc - 77dBc} =$$

$$= -20dBc - \log\left(10^{-65} + 10^{-6} + 10^{-7.7}\right) =$$

$$= -20dBc + 59dBc = 39dB$$

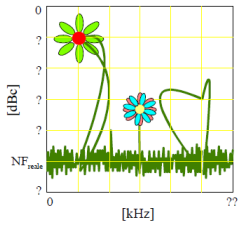
So coming back to understand what this means, we have the ideal case with a perfect sinusoid and quantization. But we didn't amplify the signal enough and we have a theoretical quality. To worsen the situation, we have also the noise, so we have SNR_real.
Then, even worse, there are also two other contributions: disturbance and harmonics.

Example 2

A 16bit DAC with FSR=5V receives a 2Msps stream of a sinusoidal signal at f_C=400kHz with 200mV_{peak} amplitude. A noise is overimposed to the signal, thus lowering the SNR by 20dB. The number of samples used for the FFT is 512,000.

a) Compute SNR_{ideal}, SNR_{real}, ENOB and real NoiseFloor.

b) Draw the spectrum, properly quoted in [Hz] and [dBc], adding also a harmonic at 3·f_C due to a THD=-70dB.



Resolution

We have a 16 bits DAC. We don't care about S_max^2 because we set it at 0 dBc. Then I can compute the SNR_ideal and the noise due to quantization Nquant.

$$SNR_{ideal} = 6.02 \cdot n + 1.76 = 98\,dB$$
$$N_{quant} = \frac{LSB^2}{12} = \ldots ? = N_{min} = -SNR_{ideal} = -98\,dB_c$$

As for the ideal noise floor:

$$NF_{ideal} = \frac{N_{ideal}}{\frac{N_{samples}}{2}} = \frac{-98dB_c}{\frac{512000}{2}} = -98dB_c - 10\log\frac{512000}{2} = -98dB_c - 54dB = -152\,dB_c$$

This is NF if the ADC was ideal. But if we have some extra noise, the NF_real = NF_ideal + delta_N. Indeed, we have noise overimposed of 20 dB.

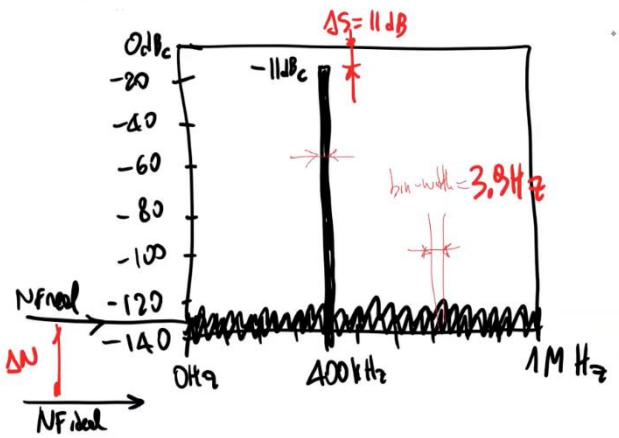
$$NF_{real} = NF_{ideal} + \Delta N = -152\,dB_c + 20dB = -132\,dB_c$$

So ideally the noise floor should be at -152 dB but in reality it's higher.
Let's now compute the SNR_theoretical, that is not the SNR_ideal because maybe we loose delta_S, and this is the case, because FSR = 5V but we are entering with a 200 mV peak sinusoid. Hence the delta_S is the total signal I could apply (5V peak to peak) divided by the signal I apply (400 mV peak to peak); in dB it is 11 dB. So SNR_theoretical is 87 dB.

$$SNR_{th} = SNR_{ideal} - \Delta S = 98dB - 11 = 87\,dB$$
$$SNR_{real} = SNR_{th} - \Delta N = 87\,dB - 20dB = 67\,dB$$

Moreover, SNR_real is not SNR_theoretical because we loose also delta_N because we have an extra noise added. Delta_N is 20 dB.

Moreover, the input signal is at 400 kHz = f_in. Then I'm sampling at 2 Msps = fs. So the plot will stop at fs/2 = 1 Msps. Let's plot the spectrum.



$$bin\text{-}width = \frac{f_s}{N_{samples}} = \frac{f_s/2}{N_{samples}/2} = \frac{2MHz}{512K}$$

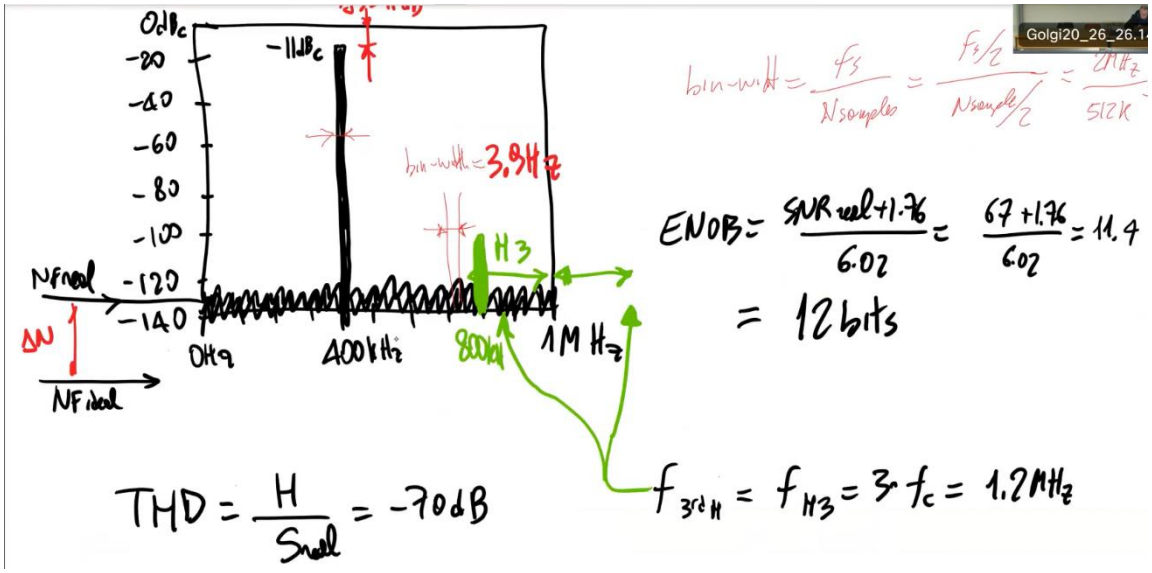
Furthermore, we know that the number of samples is 512'000, meaning that the bin width will be fs/512'000 = 3.9 Hz.

Let's now compute the ENOB = (SNR_real + 1.76)/6.02

$$ENOB = \frac{SNR_{real} + 1.76}{6.02} = \frac{67 + 1.76}{6.02} = 11.4 = 12\,bits$$

Hence we started from a 16 bit DAC but due to noise and extra noise the performances are the same of a 12 bit DAC.
The final request is to add a harmonic at 3 – fc due to total harmonic distortion of 70 dB. The third harmonic will be at 3 times 400 kHz.

The result would be at 1.2 MHz and in theory I should not see it if I plot up until 1 MHz, but the spectrum is symmetric, so I also see it at 800 kHz.



It's height will be 70 dB lower than the signal, since we have a distortion only due to this third harmonic. Since the signal is -11 dBc, the H3 will be at -81 dBc.

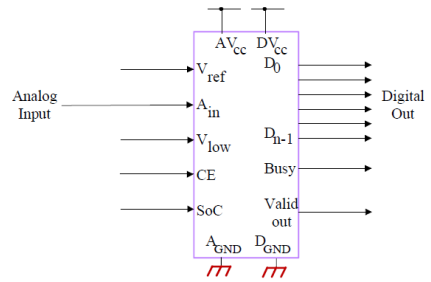
NOTES

$$\frac{N_{dm}}{Den} = \frac{Power}{attenuation\ factor} = \frac{70dB_c}{20dB} = 70dB_c - 20dB = 50dB_c$$

$$H_1 + H_2 + D + N = -80dB_c - 90dB_c - 85dB_c - 69dB_c = 10\log\left[10^{-8} + 10^{-9} + 10^{-8.5} + 10^{-6.9}\right] = -65dB_c$$

ADC

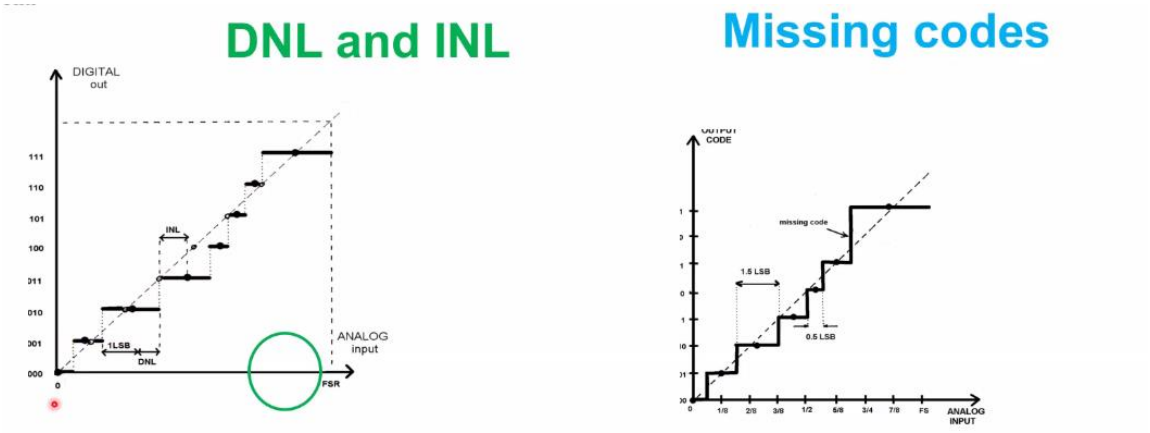
The idea is to start from an input analog signal and convert it into a digital one.



Number of bit:	n	8	16
Number of levels:	2^n	256	65536
$F_{ull}S_{cale}R_{ange}$	V_{ref}-V_{low}	5V	5V
Resolution:	$L_{east}S_{ignificant}B_{it}$ = FSR/2^n	19.5mV	76µV
	$1/2^n$	$3.9^0/_{00}$	15ppm

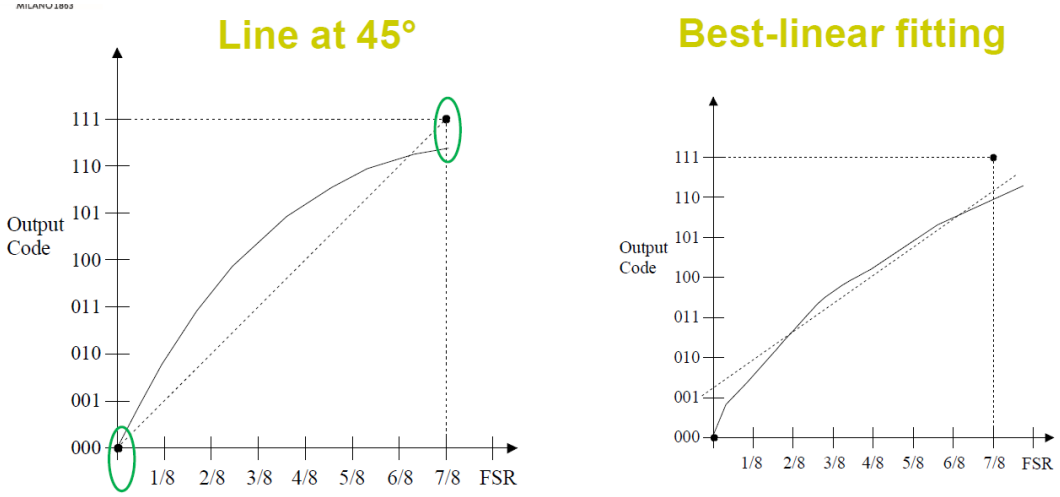
It requires some time to convert the signal, ranging from ns (SAR) to ms (dual-slope ADC).

As for the DAC, also in the ADC there are errors, that are the INL, DNL and missing codes. In the DAC the input was quantized and to the i/o characteristic was a set of dots; in the ADC the input is continuous, so the plot is a set of flat regions.



The difference between the actual step and 1LSB is the DNL. The first step has a width of 0.5LSB and the last one of 1.5LSB because the centroid starts from 0V.
As for missing codes, it may happen due to errors in the components inside the ADC.

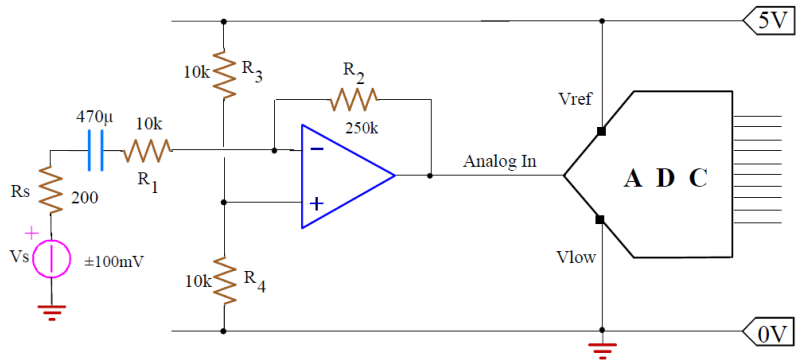
Moreover, we may have also gain, offset errors and also non-linearity errors.



The INL is the distance between the real curve and the ideal one, but sometimes the INL is not important, the important thing is to have a i/o characteristic that is as much straight as possible.

SIGNAL CONDITIONING

Signal conditioning: adjust amplitudes and impedances



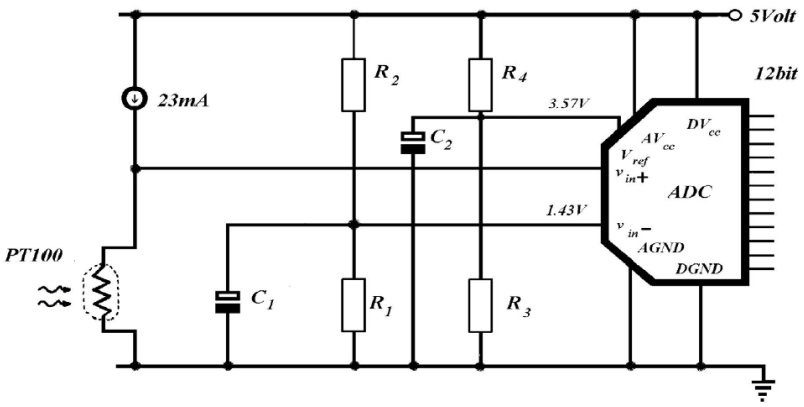
We cannot apply directly the signal to the ADC, because to exploit all the range we need some amplification. Moreover, in this example we cannot apply negative signal, so I want to shift 0V midrange and then shift all the input signals. R3 and R4 shift the range to 2.5V in input in DC, and above this we have the input signa coherently amplified.

Example of sizing

Product:	temperature meter	
Specs:	resolution	0.1°C
	temperature	–100°C ÷ +140°C
Components:	thermoresistance PT100	100 Ω at 0°C +0.385Ω every 1°C
	(61.5Ω ÷ 215.5Ω)	
discretization	240°C / 0.1°C = 2400 ≈ 4096 = 2^{12}	
bits	12	
LSB	38.5 mΩ	

The choice is to use a 12 bit ADC that discretizes the range in 4096 values. The LSB will be 38.5 mOhm. A first solution is the following.

12bit ADC and setting of V_{low} and V_{ref}

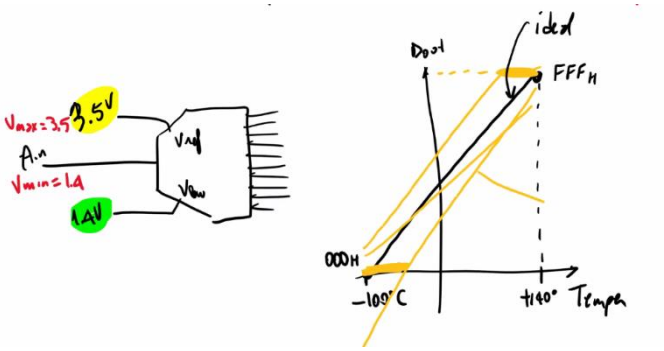


€ 0.5 + € 3
but low precision

Limitation due to resistors' tolerances: $V_{in-} = +5V\frac{R_1(1\pm toll)}{R_1(1\pm toll)+R_2(1\mp toll)} = +5V\frac{R_1(1\pm toll)}{R_1+R_2} \cong +5V\frac{R_1}{R_1+R_2}(1\pm toll)$

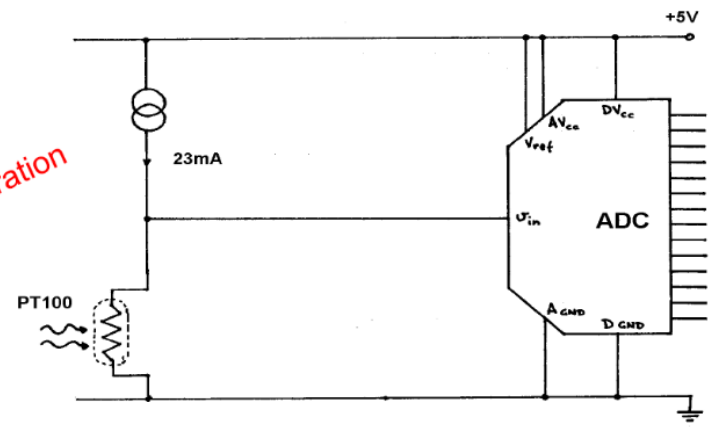
We pump e.g. a current of 23 mA and we will have a given voltage that will vary. Since we use a 12 bit ADC and we want 0x000 at the output when the input is -100 °C and 0xFFF when the input temperature is 140 °C, we could use Vref of 3.5V and a Vlow that is 1.4V. To extract those voltages we can use a voltage partition and use a capacitor to avoid a variation of the voltages due to ripple noise.

If for any reason R4 has its own tolerance (e.g. 10%), it may happen that the voltage in Vref or Vin- is not constant, but affected by the tolerance and change a lot.
The ideal characteristic I want is the black one, which occurs if Vmin = 1.4 V and Vmax = 3.5V (with the PT100 characteristics of the previous image). but if for any reason the reference values 3.5V and 1.4V change, the i/o characteristic changes as the yellow one and we have saturations.



If I want to increase as much as possible the range, I can connect directly the Vref to PS and the other to ground the following.

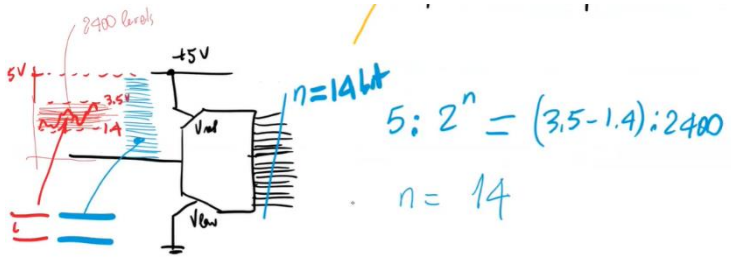
14bit ADC



€ 4
need of calibration

Limitation due to self-heating

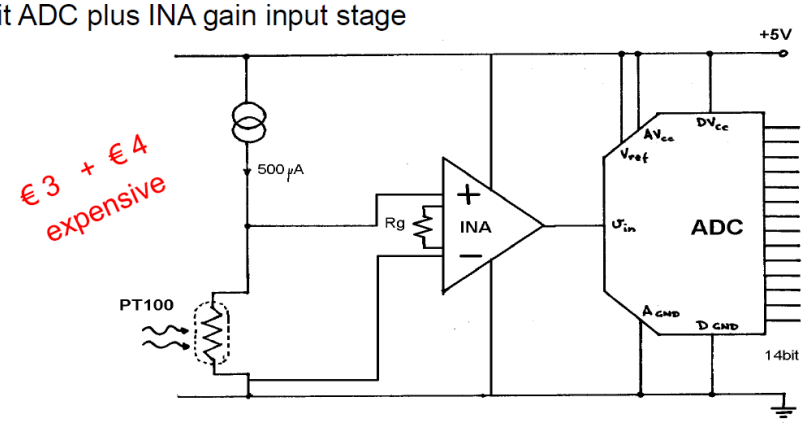
Vin still moves from 1.4V to 3.5V, so I cannot use a 12 bit ADC because the FSR is 5V.
In this approach I connected Vlow to 0V and Vref to 5V but the analog in varies between 1.4V and 3.5V. So I'm using just a little portion of the dynamics. To have a lot of levels to sample the signal I have to increase the number of bits to have the same LSB.



The issue is that we pump 23 mA in a resistor. But if so, there is the Joule effect and power dissipation. Since the value of the resistor is around 100 Ohm, there will be power dissipation P = VI = RI^2 = 53 mW.
It is not low, because we want to measure a temperature and the resistor is heating up the sample due to power loss. To reduce power dissipation and self-heating we have to reduce the current.

Let's use 500 uA.

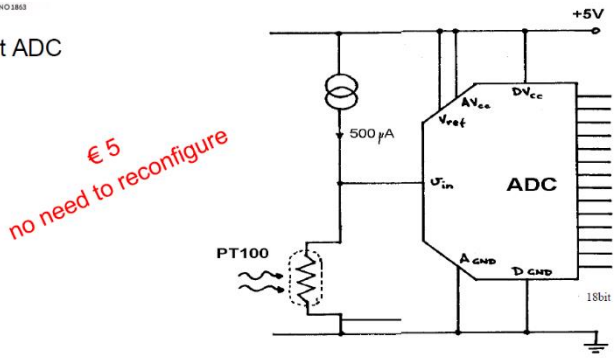
14bit ADC plus INA gain input stage



Expensive

Now power dissipation is very much reduced. But if we reduce the current, also the dynamics are reducing, so the 14 bit ADC is seeing a lower signal, so we need to put an INA in the middle to amplify the signal. The problem is that this solution is costly. So let's remove the INA and use another solution.

18bit ADC



Cheap and no need for calibration or re-redign in case of different sensor, temperature range, resolution

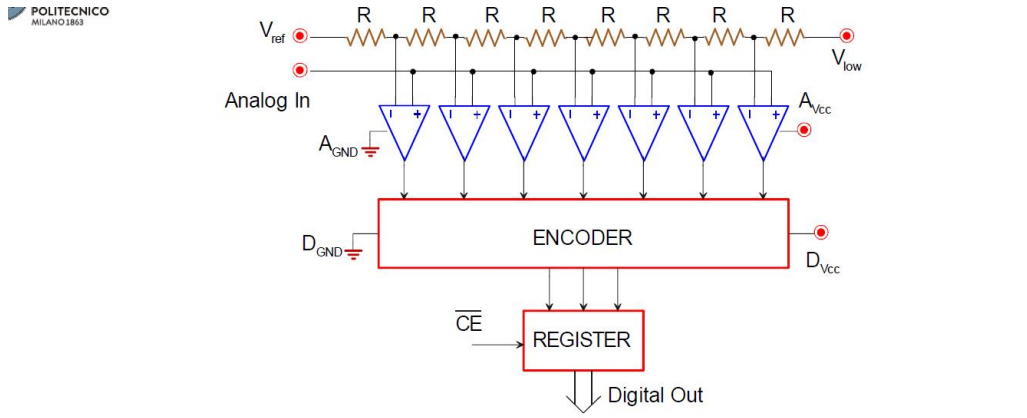
We directly put the signal in the ADC, but the signal is much smaller (we reduced the current), so we increase the number of bits of the ADC.

$5V : 2^n = \frac{(3.5-1.4)}{46} : 2400$

$n = 18\ bit$

The advantage of this solution is that we to pay just for a more expensive ADC and we spare the INA.

FLASH ADC



Components: 2^n resistors and 2^n comparators
1 "thermometric" encoder

Pros & Cons: very fast (T_C<50ns) few bits though (n<10) offset and IB of comparators

It is a fast ADC because it provides all the possible analog voltage and depending on the analog input it provides a digital output. We have an analog input, a Vref and a Vlow and we partition the FSR within these two levels. The number of resistors and comparators is 2^n. If the analog input is higher or lower than all the voltages, some comparators will be triggered and others not. Eventually we will have a digital code depending on the comparators that are active and we encode this digital code into a binary one with an encoder.

As we can see, we have a digital and analog Vdd and GND. It is a complex architecture in terms of number of components, even if it is very fast, we just need to wait for the comparator to trigger and the encoder to convert in a binary stream.
It is suitable for a low number of bits, otherwise we would need too many comparators.

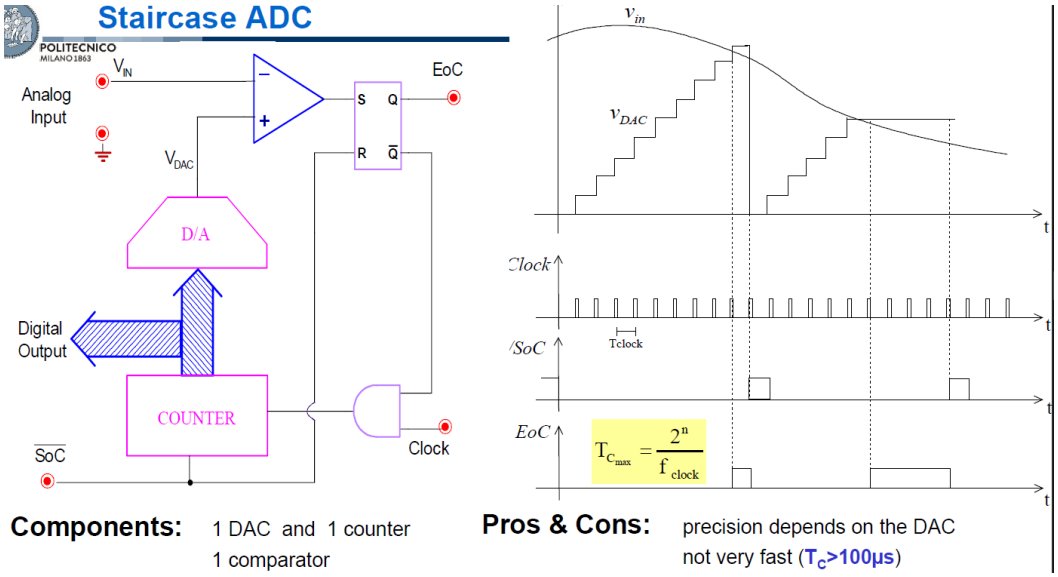
Typical errors are due to tolerances of resistors, offset and bias currents Ib of the comparators. Since Vos is limited and much lower than the LSB when Vref is high, this is not a problem, but if Vref is too small, the offset remains constant and the quality of the ADC is worsening, and we may have some missing codes. Typically, the datasheet provides INL, DNL gain and offset if Vref = Vdd, but if Vref is smaller, the errors increase.

Moreover, the comparators don't have a positive feedback, they are OL opamp, so it may happen that one differential pair is fully imbalanced on the + or – depending on the analog signal, so the bias current of the comparator might be on one pin and not on the other. So the actual voltage on the reference node of the comparator will change depending on the analog input signal. Hence the characteristic will have some non-linearities, and INL increases in the middle of the conversion range.

STAIRCASE ADC

We simplify the ADC and we don't use 2^n resistors. We apply a Vin and compare it with a voltage that we keep updating with an internal DAC and internal counter. When we apply a low level to SoC, the latch is reset, and Q goes low and Q-not goes high, so the latch is permanently reset. So if we apply pulses to Clock line, we are applying pulses to the counter. If SoC is high, counter is at 0 and output of the DAC is 0. To start the conversion, SoC goes low.
Only when Vdac > Vin, the comparator triggers and Q goes high, so we have the EoC.

Now we have only one comparator, one latch, an 8 bit counter (8 flip-flops) but also a DAC, which may be composed by resistors and opamp. So in terms of area it seems smaller than the Flash ADC but not always.



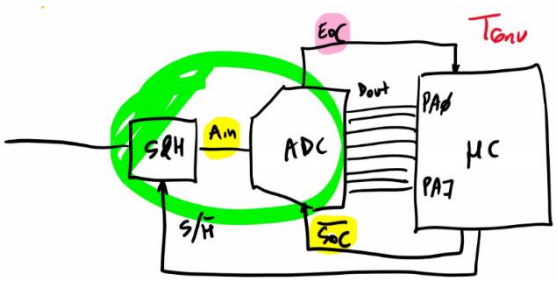
Components: 1 DAC and 1 counter
1 comparator

Pros & Cons: precision depends on the DAC
not very fast (T_C>100µs)

The precision depends on the DAC, if it is not precise, the EoC arrives with a wrong number of samples. Moreover, we still have the problem of offset and it is not so fast. In the worst case we have to wait 256 clock pulses if we are sampling the FSR.

Moreover, every time the conversion is over, there is another issue. The EoC goes high, the uC reads the data bus and re-applies a conversion to the SoC, and so on. But we cannot reapply SoC whenever we want, because we have to respect the Shannon theorem.
Furthermore, I give a certain SoC but I don't get the sample at the SoC time instant, but after some times, because I need to track the signal, and I run the risk that the values I convert are at different time durations because the delay between the SoC and the conversion depends on the variation of the signal.

To avoid this problem, we can place a S&H before the ADC, even if sometimes is already integrated in the ADC.



TRACKING ADC

To improve the previous solution, we allow the staircase to go both up and down. Once we have the conversion, we don't restart then from 0, but form the last converted value. Again we have a comparator, and a counter that now can move up and down. If the comparator is high, so the analog input is lower than Vdac, I need to decrease the voltage, so the counter goes down. And eventually up again. Once we reached Vin, the counter will go up and down chasing Vin.

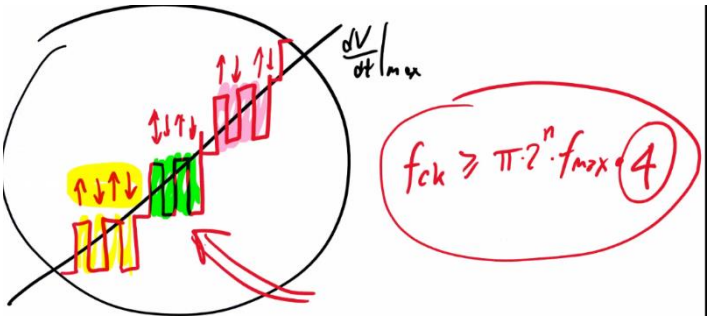
Of course, the clock should be fast enough to cope with the maximum slope of the input signal. So the tracking condition is the one aside, when the slope of the signal is smaller than the slope of the counter.

$$\left.\frac{dV}{dt}\right|_{max} = 2\pi f_{max} \cdot V_{pmax} \leq \frac{1LSB}{T_{ck}}$$

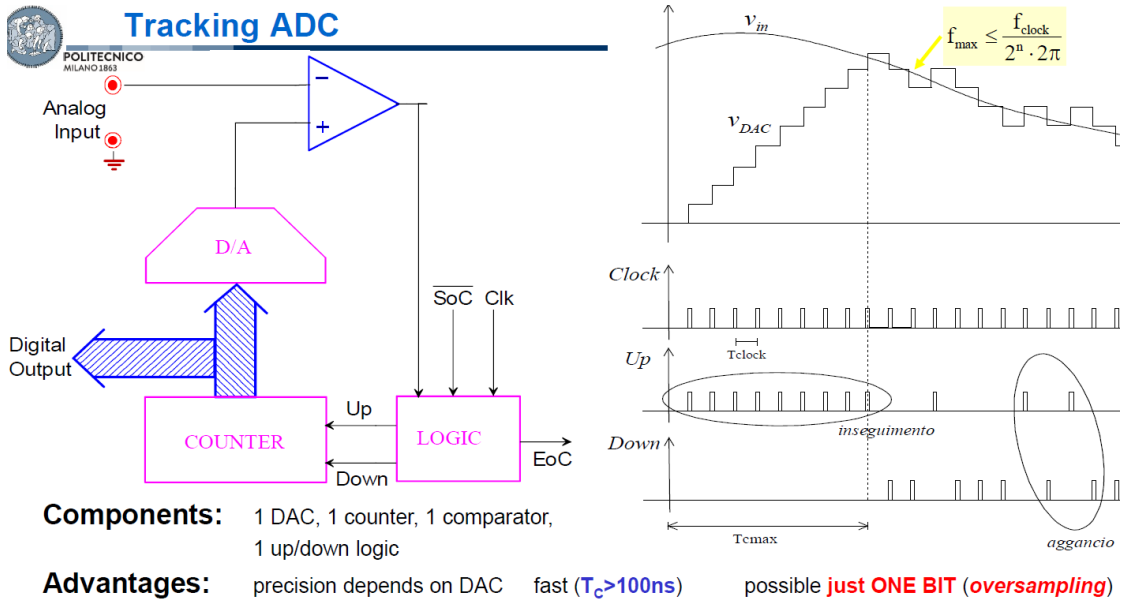
$$2\pi f_{max} \cdot \frac{FSR}{2} \leq \frac{FSR}{2^n} f_{ck}$$

$$f_{ck} \geq \pi \cdot 2^n \cdot f_{max}$$

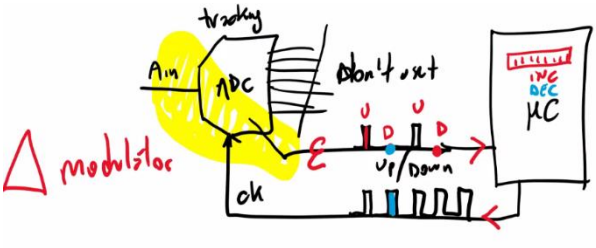
In a kind of oversampling, it is better to use a higher clock frequency.



The initial acquisition may be long but then, once we locked it, we can stay locked if fmax is the one computed before, and at every clock pulse we can have a conversion. We have a valid data at every clock pulse.



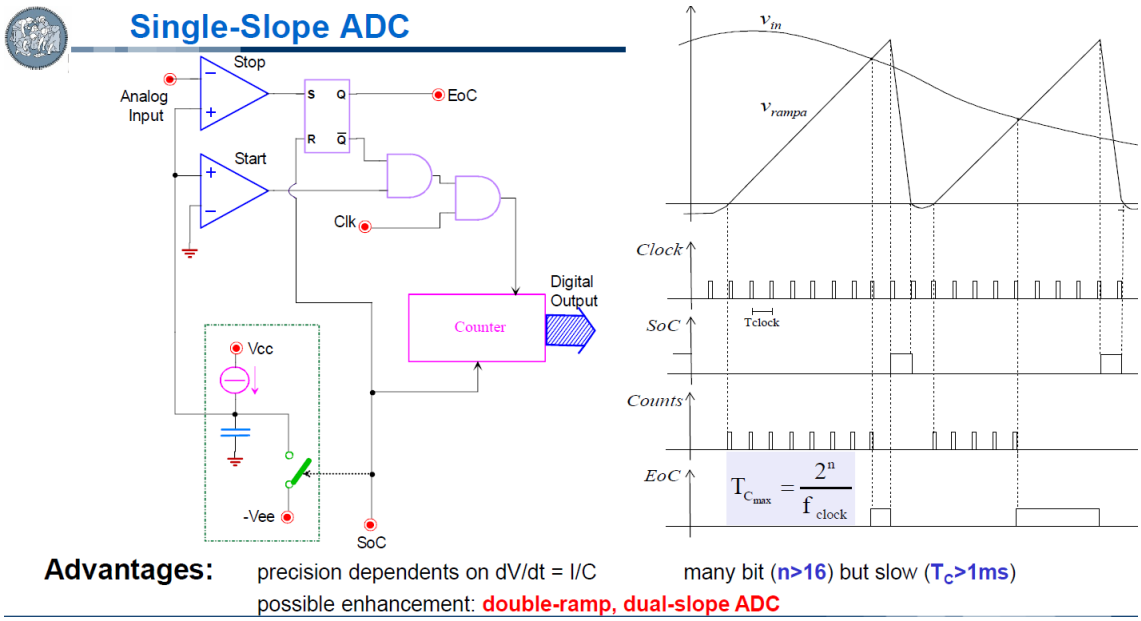
Moreover, we don't need to feed all the ADC output lines to the uC, because it s sufficient that the output signal up or down is given back to the uC in correspondence of the clock pulse it provides, so that it can update an internal register that the uC has.



This is also called delta modulator because it doesn't provide a digital bus to the output but simply the up or down pin.

Another possible solution is the following.

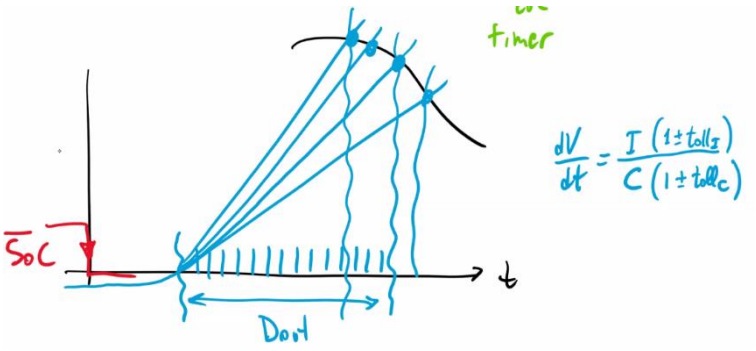
SINGLE SLOPE ADC



We have a constant current generator pumping current in the capacitor. When the voltage exceeds the analog input, then we stop the conversion. We start from a slightly negative value to compensate the Vos of the comparators. At the beginning we also reset the counter, which counts now the time.
Once SoC goes low, the switch is open and the capacitor starts to charge up, with a constant ramp. The start comparator is used so that we start counting only when the cross the 0V threshold. Then the counter is enabled and it counts.

The problem is that in the worst case scenario Tconv is high if we need to reach FSR.

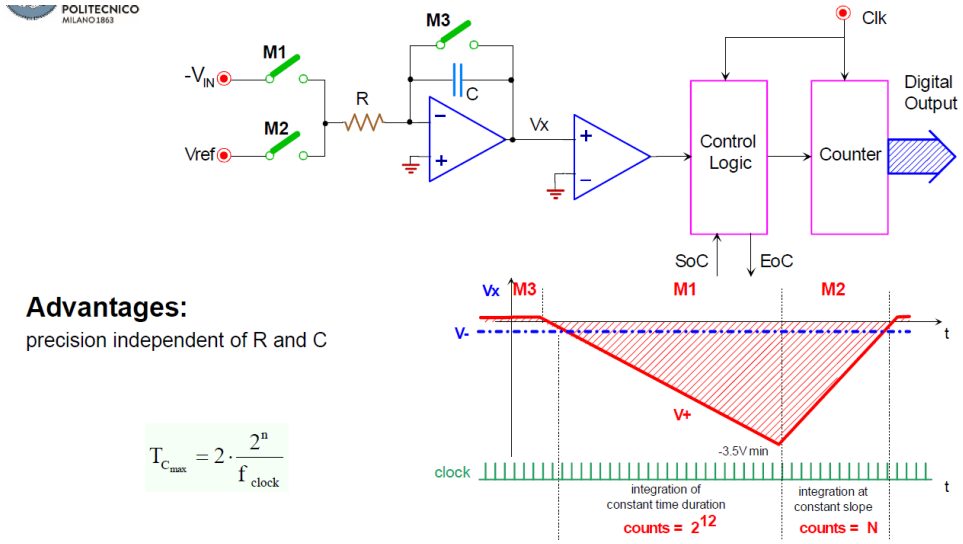
The issues of this ADC are the offsets of the comparator but also the tolerances in the analog part with the current generator and the capacitor, so the slope can be different and so also the crossing with the analog signal → the digital output may vary.



In this case we have a poor reproducibility and INL and DNL are very bad.

DUAL SLOPE ADC

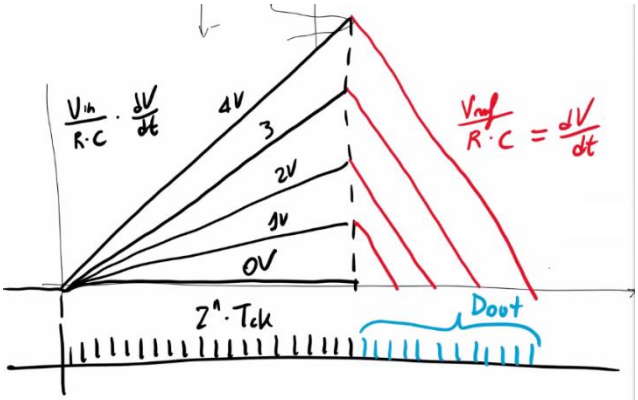
Since we know that we have tolerances in the components, we double them.
We start by integrating a current in the capacitor, and there is no current generator. We start by connecting Vref and we have an integrator that is reset if M3 is close. When we start the conversion, M3 is open, Vin applied, a current Vin/R is applied to the capacitor and we decrease the voltage with a



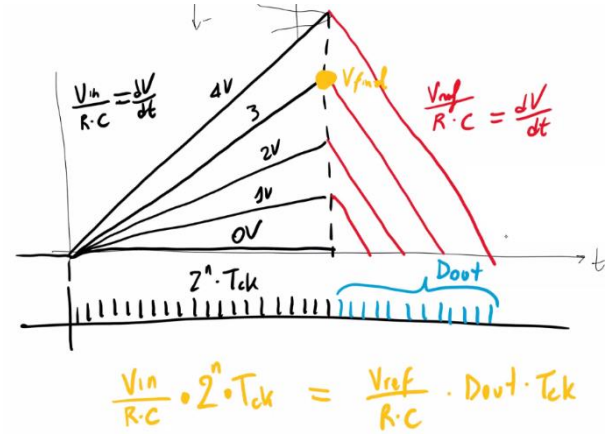
$$T_{C_{max}} = 2 \cdot \frac{2^n}{f_{clock}}$$

constant slope. Then I wait for enough time until I reach FF (2^n clock pulses) and the counter keeps counting up until then. Then M1 opens and M2 closes. The counter is reset (even if it is not needed because it is in overflow), now I integrate Vref (I must use a positive Vref and negative Vin or viceversa) on the capacitor. During this second phase we revert the direction of integration, the integration is faster because Vref = 5V and so it gets completed. The digital output is the number of pulses needed to bring the voltage Vx back to 0.

The slope during the deintegration is Vref/RC. So in the second phase the slope is steeper and the number of ck will be lower and it will be our digital output.



We can perform the following calculations.



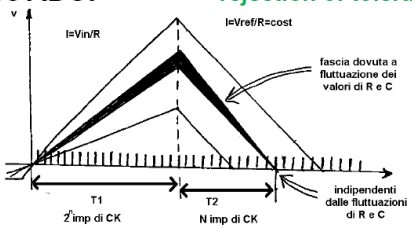
$$D_{out} = \frac{V_{in}}{V_{ref}} \cdot 2^n$$

$$\frac{V_{in}}{R \cdot C} \cdot 2^n \cdot T_{ck} = \frac{V_{ref}}{R \cdot C} \cdot D_{out} \cdot T_{ck}$$

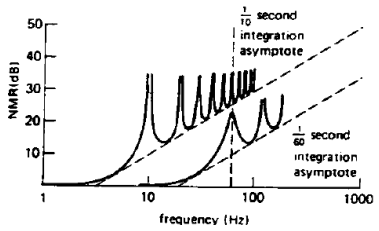
So the output conversion doesn't depend on C or on the clock frequency. This because the mismatches compensate in the two phases.
We can compute the maximum conversion time, that occurs when Vin = Vref, and the result is in the slide. Unfortunately, this type of ADC is slow, but at the same time I can use a large number of bits with a very fast clock.

Of course, the bias current of the opamp may impact the result, but since it has a finite and the same direction during both the integration and deintegration phases, its effect cancels out. But what about charge injection in the switches and Vos? Should we add a dummy cell to compensate charge injection?

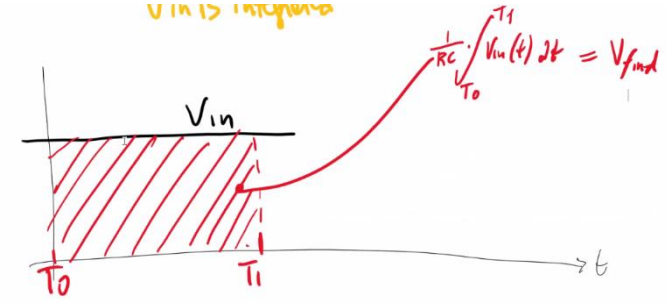
1st advantage of dual-slope ADC: rejection of tolerances of R and C



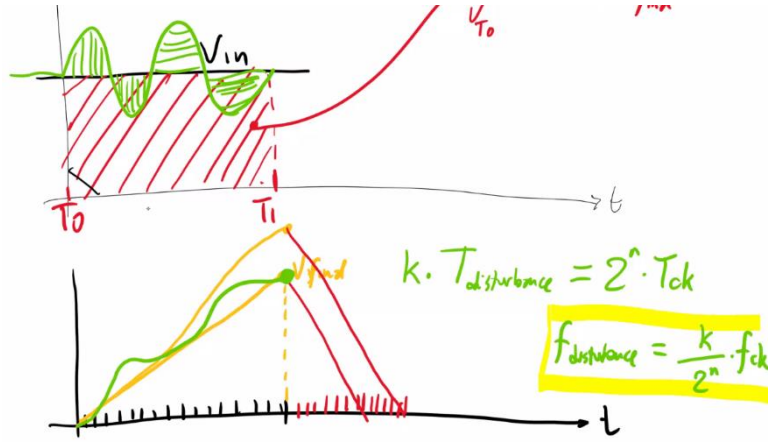
2nd advantage of integration ADCs: rejection of disturbance at given frequencies



The other major advantage of the dual slope ADC is that it is an integrating ADC, which means that during the first phase Vin is integrated, which means that the integral of Vin is Vfinal. But if Vin changes, the area (integral) changes. So if Vin is constant, during the first phase the slope will allow to reach a certain Vfinal, but if Vin varies, the final value will change and the result of the conversion will change.



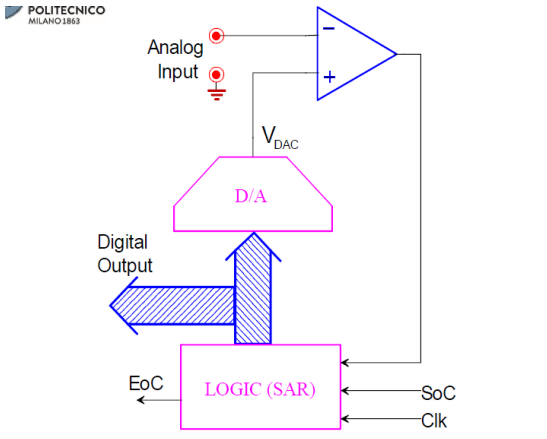
However, if on top of Vin we have a disturbance, the slope changes but the Vfinal is the same, because the areas of the disturbance cancel out. To reject the disturbance we need to obey the formula in the image.

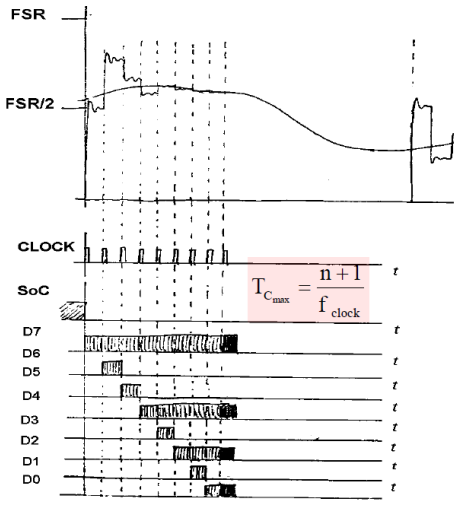


All the frequencies following this relationship are rejected.
NMR is the normal mode rejection factor and it increases the higher the frequency of the disturbance.

The higher the frequency gets, the smaller is the residual area, and this is the reason why the higher the frequency gets, the higher is the rejection factor.

SAR ADC





Components: 1 DAC and 1 sequential SAR logic

Advantages: precision depends on the DAC

many bit (**n=8..14**) and fast (**T_C>10µs**)

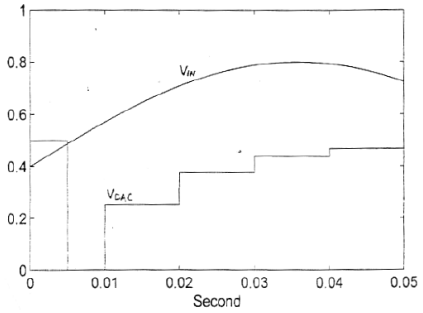
We consider the 0000, then we set just the MSB to 1 to have half of the range (half of FSR). Then we compare Vdac with the analog in. If Vin is still higher, we were right in setting the MSB, otherwise we set it low. Then we set the second MSB bit to 1 and make the same comparison up until I get the correct codification.
So we need a DAC but not a cunter, a logic component instead. It takes n+1 clock pulses to get the data (one last to fetch the data). So it is much faster than the staircase ADC.

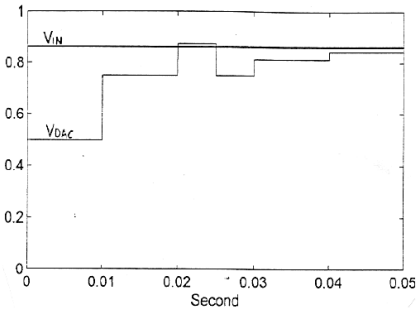
The logic of the SAR is composed by flip-flops.

Frequencies of the SAR

If Vin is constant, the SAR works perfectly, and after Tconv the result will be perfect. But if Vin changes it's a mess, because we are always comparing with a new different value. So better not to vary Vin, but this condition is pretty though to fulfill.

Maximum input signal frequency:





With NO S&H at the input:

$$f_{in,max} \le \frac{f_{clock}}{2\pi \cdot 2^n \cdot (n+1)}$$

with S&H at the input:

$$f_{in,max} \le \frac{f_{sampling}}{2} = \frac{f_{clock}}{2 \cdot (n+1)}$$

We want Vin to move less than 1 LSB; the maximum slope of Vin 2*pi*f_max*V_peak. If I consider 8*Tck, we have to respect:

$$2\pi f_{max} \cdot \frac{FSR}{2} \le \frac{FSR}{2^n \cdot 8} \cdot f_{clock}$$

Which in turns becomes:

$$f_{clock} \ge \pi 2^{n+3} \cdot f_{max}$$

If for instance I want to use 10 bits and f_max = 20 kHz:

$$f_{clock} \ge \pi \cdot 2^{10+3} \cdot 20kHz = 500MHz$$

It is a too high frequency for f_clock.
But Tconv = (n+1)*T_clock, so the conversion time is 11*T_clock. Moreover, 1/T_conv = 1/Ts = f_samlp.
Since f_max should be lower than f_sampl/2, we have the following.

$$T_{conv} = (n+1) \cdot T_{ck} = 11 \cdot T_{clock}$$

$$\frac{1}{T_{conv}} = \frac{1}{T_s} = f_{sampling} = \frac{1}{11} \cdot f_{clock}$$

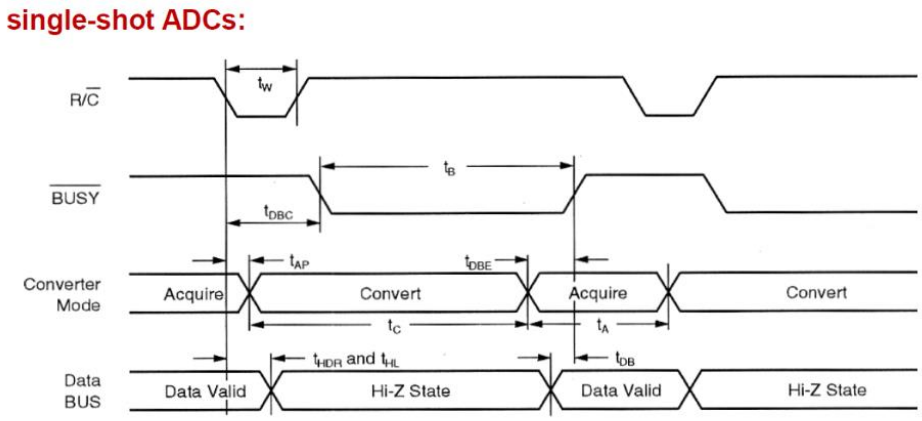
$$f_{max} \le \frac{f_{sampling}}{2} = \frac{f_{clock}}{22}$$

$$f_{clock} \ge 22 \cdot f_{max}$$

So the condition from Shannon is f_clock > 22*f_max, and on f_clock I have also another condition from the SAR ADC.
So the condition for the SAR that outputs 500 MHz is the correct one. The other condition is valid if Vin is constant, so with a S&H in input. The limitation to be considered is the most stringent one, so if Vin changes the first one.
But Vin should not change during the conversion, so better to use a S&H in between the source of the signal and the SAR ADC to keep Vin constant during the conversion.

TIMINGS

Single-shot ADCs



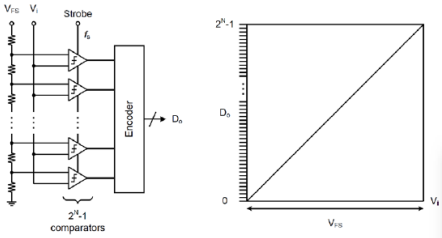
We can have single shot ADC or free running. A single shot ADC is in a condition, e.g. ready, that is le first line and it is high. If the ADC is not busy, the BUSY line is high. the data bus provides a valid data of the previous conversion. Then SoC (first line) goes low and we have the start of a conversion. Inside the ADC the S&H opens the switch and we enter the hold mode and the ADC starts converting. During conversion the output data bus is left in tri-state, which means that the output levels are in output impedance, so floating. Only at the end of the conversion, once the internal logic is done, then the logic says that the conversion is done and the digital output bus is provided to the external bus. Hence the during the conversion BUSY is low and once the data is outputted it returns to a high level.
So some ADC instead of having the EoC pin have the BUSY pin, that has the same functionality.
In the single shot ADC the uC asks for the data and then it is given to it.

ADVANCED ADCs

The basic idea is to improve the previous architecture to increase the speed of the ADC or to reduce the complexity. To increase the speed we can improve a serial pipeline or a parallel architecture.

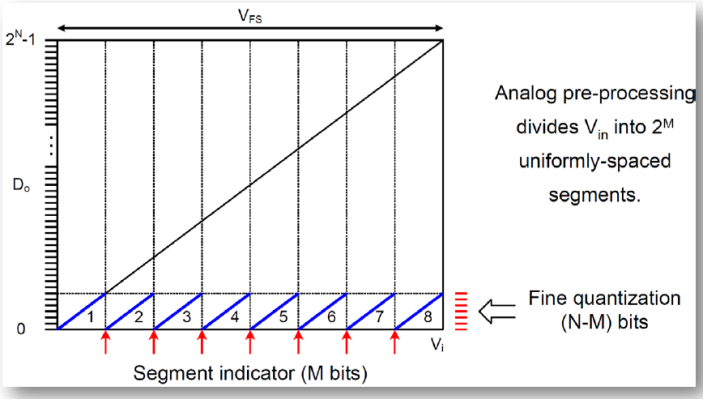
SUBRANGING ADC

In the Flash ADC we have to put in input all the possible voltages. It is not an efficient architecture because we also have a lot of comparators and we run the risk of having a lot of comparator keep consuming power and so silicon area. So the basic idea is to implement segmented quantization, a caurse quantization with less bits to locate the range where the signal is and then apply the fine conversation.



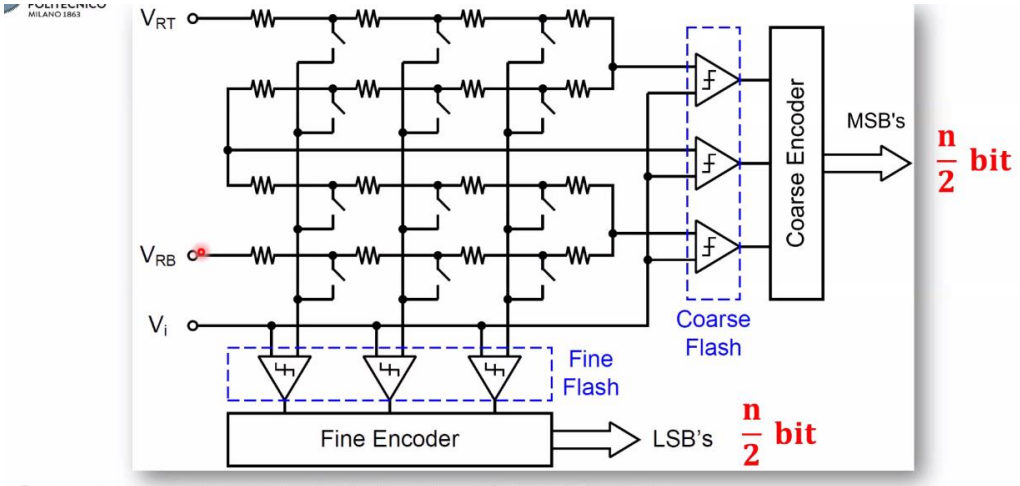
In standard Flash ADCs, comparators near V_{in} are important just for the "fine" conversion, hence they are useless for the "coarse" one, thus the architecture wastes resources and is not efficient

Therefore… **segmented quantization**



This is the idea of the subranging ADC. The idea is that in a standard flash ADC we would need 2^n resistors and the same comparators. Now we group the resistors in 4 groups, so that we need 3 comparators that are used for the coarse definition. Once the two MSB bits are known, we close the respective set of switches depending on the thermometric code on the coarse flash encoders.
Once the switches are closed, we move to the fine flash.

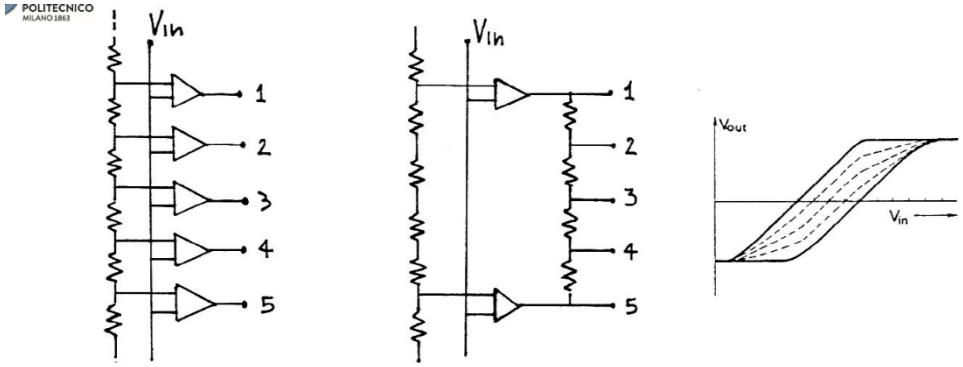
So we have developed a 4 bit ADC without needing 15 comparators, but just 6 comparators and digital electronics to drive the switches.



Coarse comparators are connected to the coarse reference ladder taps
After getting MSBs, the fine ladder taps are enabled and fine LSB are computed

This ADC consumes less power and area.

INTERPOLATION FLASH ADC

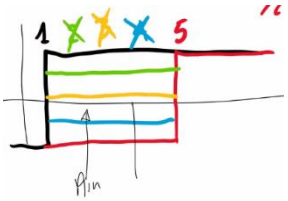


Example, for an 8 bit interpolation ADC:

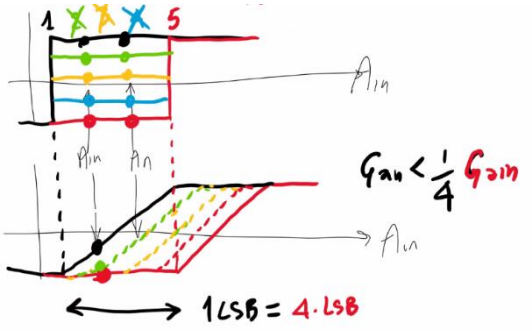
- instead of 256 comparators, it needs just 64 (¾ are removed)
- huge reduction of silicon area, power dissipation, input stray capacitance, layout
- improved dynamic performances (settling time, speed, …)

Another improvement of the flash ADC. This improvement relies on the interpolation of the output voltage. In the classical ADC we have the thermometric output. The other possibility is to reduce the number of comparators, keeping only the first and last one. So just ¼ of all comparators. Now the output of each comparator should be digital.

The upper comparator won't have a digital output, so a very high gain and sharp transition, because we want to use comparators with lower gain.
In the classical flash ADC we used a comparator for each bit and the output should have been digital, so we need a high gain, to have a vertical sharp transition.
The concept of interpolation is that we remove some comparators and we add resistors. In this way we reduce the area occupation and we have not only the digital output of the comparator but also the intermediate values, that won't be digital of course. So the output will be like below (red and black are comparators 5 and 1).



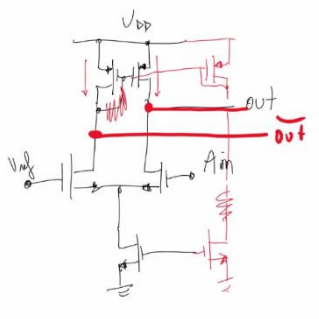
If we vary Vin, the interpolating value will just be the average between the levels high and low of the comparators. The problem is that in this way we have the same output value with different analog in.
So the idea is to have not an infinite gain, but to reduce it in the comparators. Since the gain is lower, it requires a lower number of transistor so it occupies less area on the chip.
Now the averages follow the slope of the comparator, and with the value of analog in, the intermediate point can be above or below 0. With this information we can reconstruct the position of analog in (now the LSB is 4 times the LSB of the traditional flash ADC).



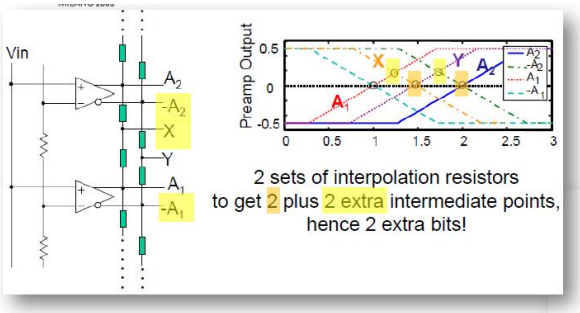
After the resistors we need to place some electronics so that we can see which analog output is above zero, and we can regain in this way the 4 bits that were lost removing the comparators.
The drawback is that we have a larger area occupied by the resistors.

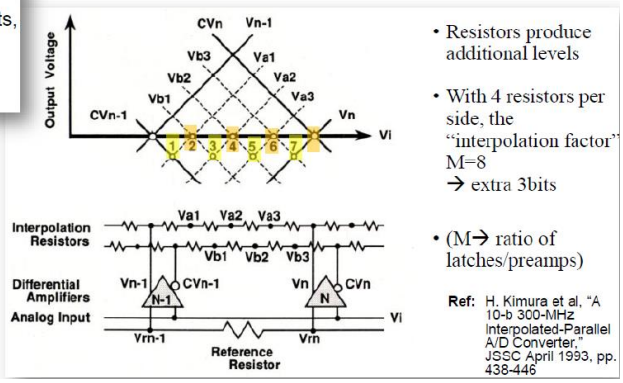
Higher order interpolation

Using a differential comparator. In its internal architecture, instead of using a current mirror and employing a single mode output, we remove the mirror and have a differential output. The advantage of this configuration is that when we apply a Vdiff between the inputs of the opamp, instead of having just one commutation we have also the one on the other output. Thus, we can duplicate the network of resistors at the output of the comparator.

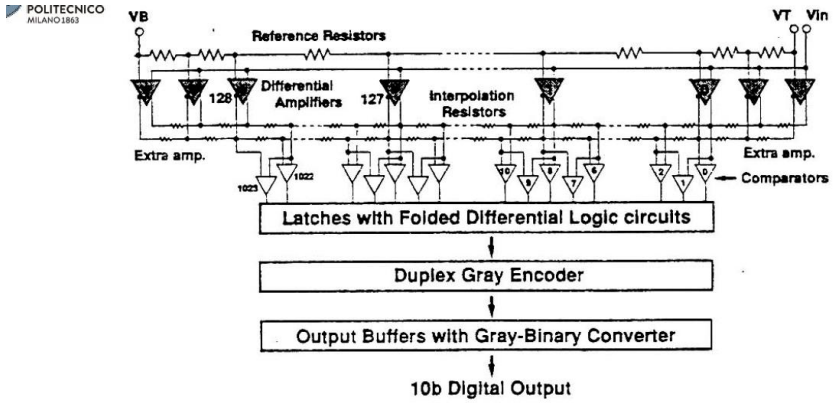


We will have the intermediate values X and Y and, in this way, we have more input signal available to understand where we are.



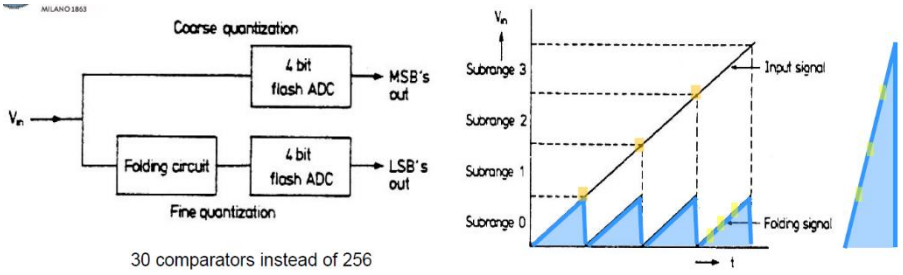


To find more commutating threshold we need to take adjacent points on the two lines of resistors.



Example of a 10bit ADC with 300Msps:
- 128 differential OpAmps (G=10), each with 4+4 resistors and 8 comparators
- same total number of comparators (128·8=1024), hence same area and dissipation
- but advantages for a reduced C_{in} (128/1024), better DNL ($V_{osOpAmp}$/8 e V_{osComp}/Gain)

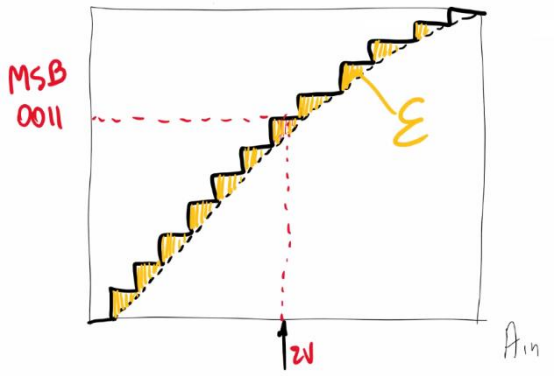
FOLDING FLASH ADC



Analog pre-processing aimed at obtaining separately:
- the MSBs through a low-resolution flash ADC)
- the LSBs through a "folding" analog circuit followed by (amplification and) ADC

Let's suppose we want to develop a 8 bit ADC and we have just two 4 bit ADC. Maybe we can use the two in parallel. Of course not, because the i/o characteristic is split in just 16 levels for both, and it is the same, not divided in 256 levels. So we cannot do this.
But we can use two comparators but on the path to one of them we need to introduce something more. A Vin is applied to one ADC, so we can get the 4 MSB of the conversion. Then we need to check in which range we are and the residual error on the range.

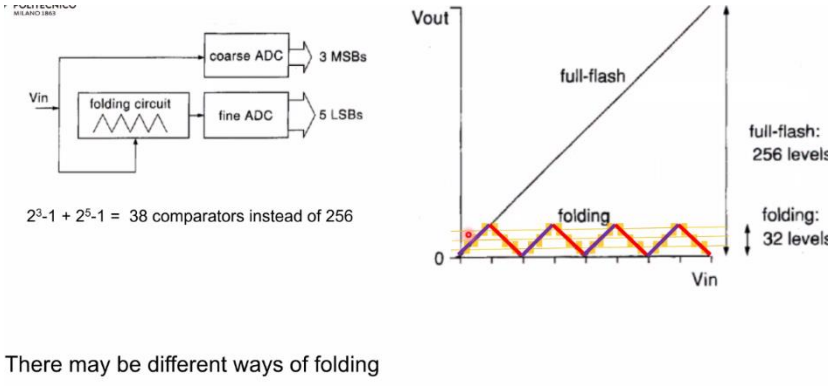
So at first we convert the analog in in 16 levels, but we have also a very big error. So we need a circuit to compute the residual error that we have (yellow). If I'm able to predict the error with a certain circuit, from that I can retrieve the LSB.



The folding circuit tells the residual error of the 4 bit ADC. Then the error is fed to an identical ADC used for the MSB conversion and we retrieve the LSB. Thus we need less comparators but a folding circuit.
The folding circuit is not a comparator, but it tells us the error that a 4 bit ADC provides.

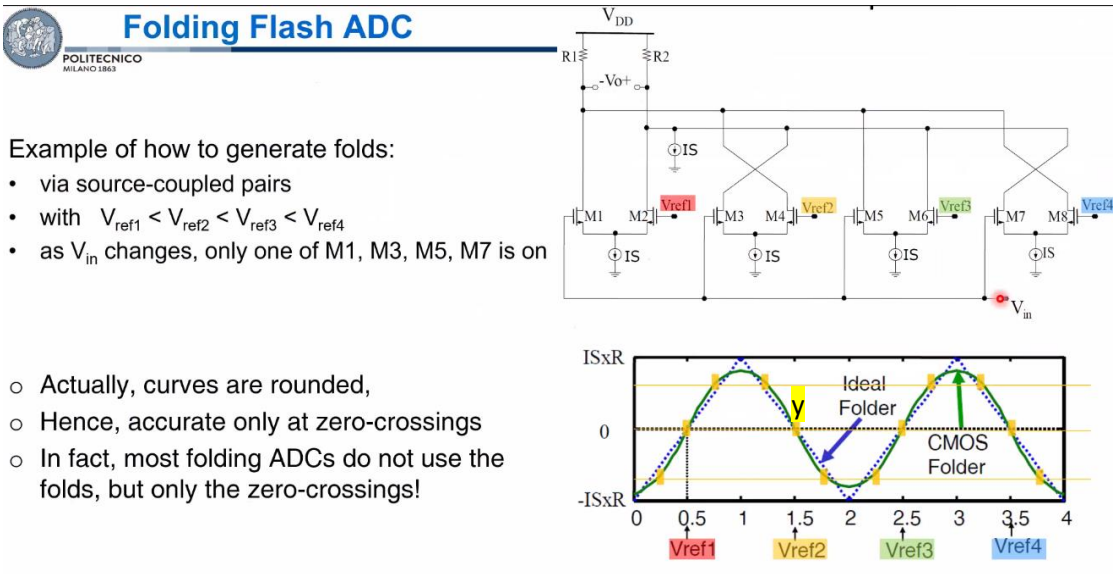
There can be different kind of folding ADC. Usually the following is the better transition, because having sharp transition over the folding one because in the sharp one we should be able to create very vertical transitions, and when we do this usually it implies very high gains and smoothing at the edges.

Then if for instance we need to regain two bits in the folding circuit, what we can do is taking the output of the folding circuit and compare the output and feed it to a two bits ADC (3 comparators). So the folding should be compared with 3 levels introduced by 3 comparators.
So if we enter with a certain analog input, the coarse ADC will give us the 3 most significant bits and the folding ADC will provide a residual analog voltage that will feed the fine ADC which will be composed by a given number of comparators. If the fine ADC has to be of 2 bits, then we would require 2^2 – 1 comparators, so just 3.



So the fine ADC provides the least significant bits. So if e.g. we want to convert in 8 bits, we can have 5 bits for the coarse ADC and 5 bits of fine ADC. In the end we require less comparators.

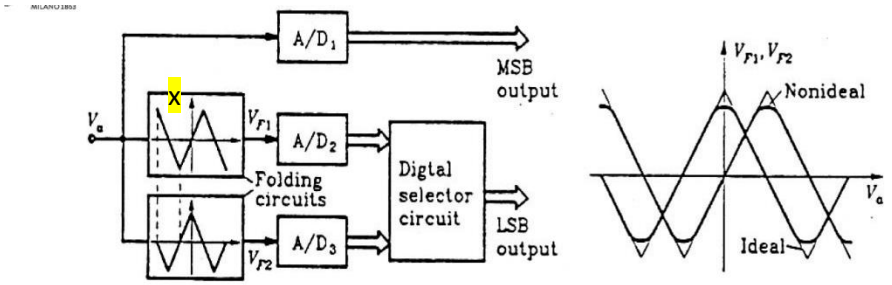
An analog implementation of the folding circuit is the following (not asked at the exam).



It requires 4 different Vref and differential stages, so that we can provide an analog signal that varies as the green one. It is similar to the ideal blue one (triangular) but not actually it, due to the linear behaviour of the transistors.
Now the output of this folding circuit should feed the fine ADC. But if the fine ADC is just a 3 bit one, the green curve should be partitioned by 3 comparators. Depending on the Vin we apply we will see which threshold we are exceeding.

Even with a triangular i/o curve like x, or almost sinusoidal, we run the risk of having a linear transition in y, while at the top of the curve we might suffer a saturation. This is the reason why in many real implementation of the folding ADC architecture we don't use one folding circuit and one fine ADC, but two folding circuits delayed by a proper amount so that one folding circuit operates in the linear region and the other one not, so it will suffer from smoothing. Vice versa, when the Vin requires the first folding circuit to operate in the worst region, then the second will work in the linear region. This requires a duplication of the folding circuits and of the ADC. Hence the digital selector electronics is more complex.

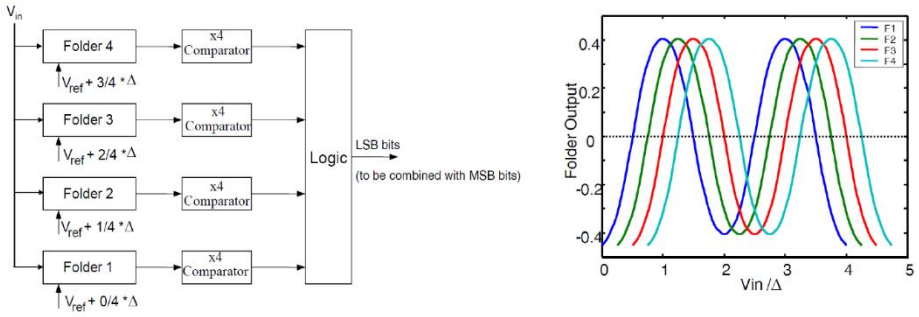
However, the overall performance of this duplicated flash ADC is better.



Example of a double folding circuit, in order to avoid non idealities of folded edges

… other folding shapes can be used…

There are also other architectures, like the one below (not asked at the exam). Instead of using one folding ADC and a fine ADC and a coarse ADC, we could use an architecture with 4 different folding ADCs and each folding ADC drives a comparator such that we can provide the required LSB.

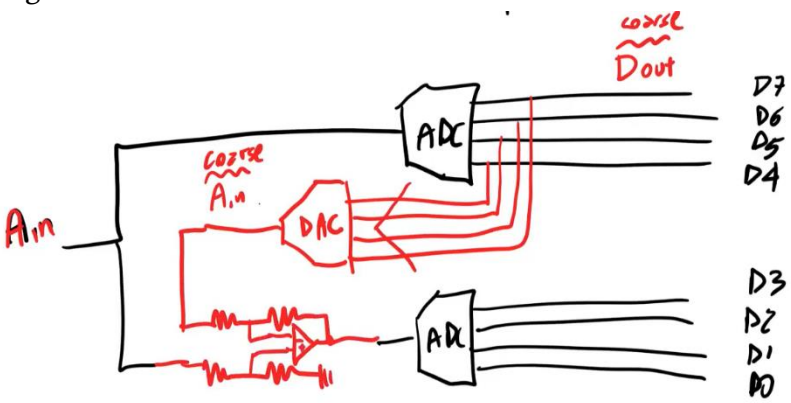


Example of 4 folders with 4 folds each, so 16 zero-crossings, hence + 4 LSB bits

… anywat upper limit… due to added complexity.

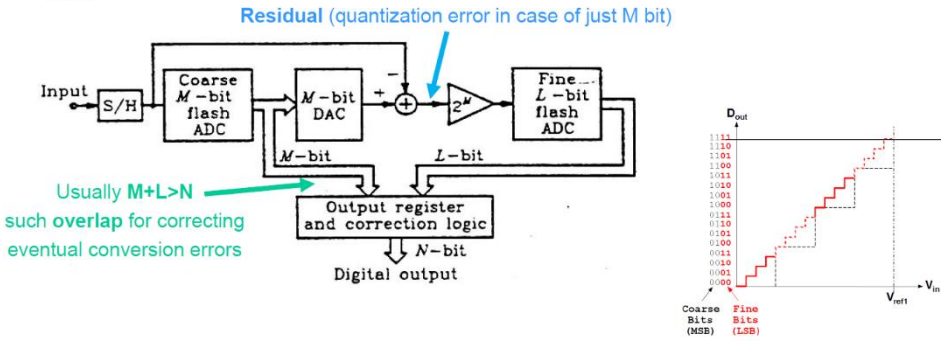
So the folding ADC requires a folding circuit that should be able to foresee which is the error that the coarse ADC causes. So it should provide the folded signal (triangular shaped). But how can we create this folding circuit? We need to check the digital output code of the coarse ADC, reconvert it in an analog signal, make the difference between the reconverted Vin and the original one to provide the residual error.

The one above is the coarse ADC, and the folding circuit could be done by taking its output code, putting a DAC to convert it into an analog signal and then we compute the difference between the signals with a standard analog stage.



This circuit works, but I would like to have a folding circuit with just analog components (without the DAC). So the correct name for this circuit is half flash ADC.

HALF FLASH ADC



Conversion is a two-steps process through two cascaded stages
Such "half-flash" (or "two-step flash") approach drastically reduces:
components count power dissipation input capacitance

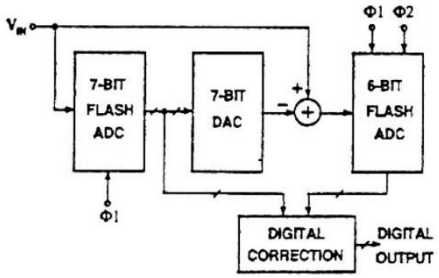
We apply an analog in, we use a coarse n bit ADC, we have a certain output composed by the most significant bits and then we convert these bits in analog with a n bit DAC, we do the difference to get the error and then we multiply the error by 2^n. Then we can use a L bit fine ADC.
Hence e.g. in the case of a 8 bit ADC I would require two 4 bits ADC, so 30 comparators instead of 255.

So the coarse ADC divides the input range in a given number of levels (very large steps) and the fine LSB will describe the smaller steps.

In reality, if we do simulate the circuit, something bad happens every time we have commutation of the coarse ADC. When we are very close to the commutation of the coarse ADC (e.g. from 01 to 10), we run the risk that, due to errors introduced by the DAC, by the voltage difference stage and in the voltage gain stage, the MSB is still 01 but the fine LSB sees 10. For this reason, if we require 8 bit, we don't convert 4+4 but we add some overlapped bits. E.g. we convert 5 bits, we compute the error and then 4 bits in the fine ADC, and we use them to have 8 bits at the output and avoiding jumps in the commutating regions. This could have been avoided if all the components in our analog part would have been ideal.

The problem of this configuration is that the conversion time is prolonged because we have the input, S&H, and during hold we need t_conv for the coarse ADC; then we require another t_conv for the DAC, some analog settling time to have a proper error subtraction and then again t_conv for the fine ADC. Hence a 8 bit flash ADC requires 10 ns, while the half-flash 40 ns.

It is called half-flash because we step the conversion in two steps. The one below is another implementation for a 12 bit ADC.



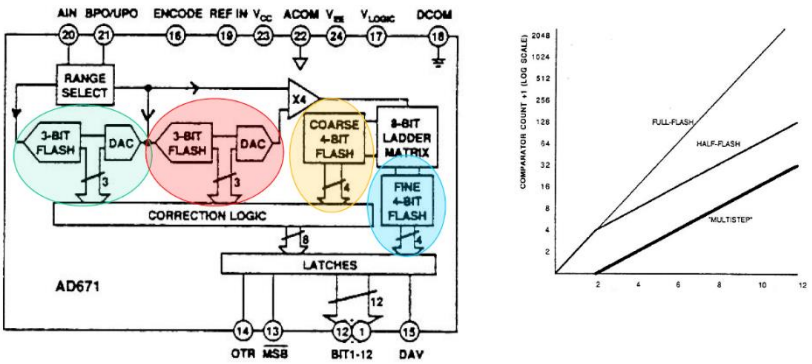
PERFORMANCE OF A/D CONVERTER

Differential Linearity	12 b
Conversion Rate	5 MHz
Peak SNDR	65 dB
Input Range	5 V
Power	200 mW
Power Supply	5 V
Input Capacitance	15 pF
Active Area	1.2 mm × 3.0 mm
Technology	1-μm CMOS

Example 12bit ADC:
a flash ADC should require 4096 comparators
the half-flash ADC instead needs 27+26=192 comparators (1bit overlap)

MULTISTEP ADC

Instead of splitting the conversion in two steps, we use a multistep approach and we split the conversion in multiple steps.



Example of a 4-step ADC:
First 3bit and second 3bit flash ADCs, then first 4bit and second 4bit half-flash
Overall just 48 comparators !

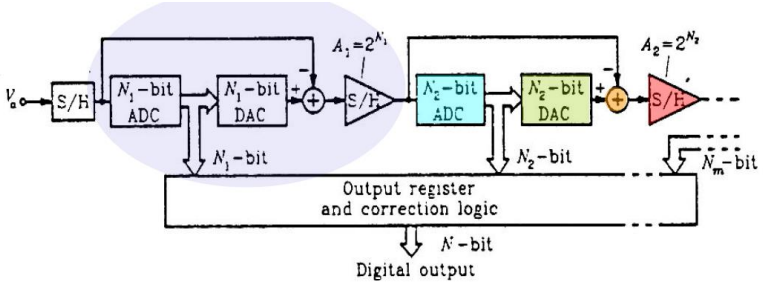
In the example that we have in the image, we have a 3 bit flash ADC (blue), then we convert them with the DAC, we compute the error, we amplify the error and use another 3 bit ADC (red one), we get the 3 bits, we reconvert them, we compute the difference, amplification and then reconverted back with 4 + 4 bits (yellow and blue). The total 12 bit output is achieved with multiple steps.

Given the idea of the half flash ADC, we copy that idea and implement how many conversion we wish. Ideally, we can use how many conversions we want → **pipelined ADC**.

PIPELINED ADC

So we use an ADC and then a DAC and a subtraction node, many times iteratively. The problem with this configuration is the long conversion time.
It is a multistep ADC where we add a S&H circuits in the middle before each conversion stage. If we do so, we apply Vin, we close the first S&H and open it in the first hold phase. The first ADC will start the conversion. After t_conv we get the MSBs. Then the second S&H stores the error related not to the Vin, but to the previous sample. So the second ADC can convert the error. Hence every time we apply our analog input, it takes t_conv to get MSBs from the first ADC and also the other bits, because of the S&H circuits. So from when I apply analog in to when we get the digital output code we need, it takes n*t_conv, where n is the stages I use. But once analog in has propagated, every time we apply a new clock pulse so we close and open the S&H circuits, the digital output of the pipelined ADC will be valid. So we have improved the conversion time of the ADC by a factor N, where N is the number of stages.

The only complexity to add are the additional S&H circuits.

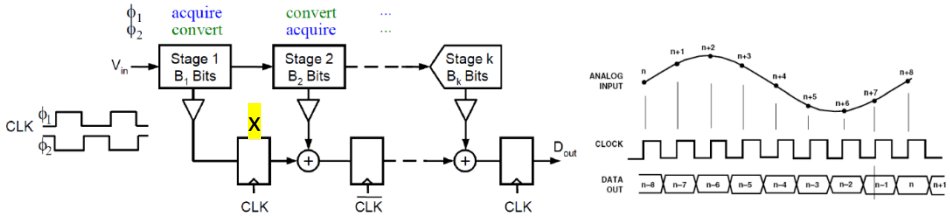


m-stages each one composed by:

ADC Flash (low resolution 1÷4bit), DAC, analog adder and amplifying S&H

Compared to multistep ADC, now S&Hs allow parallel-pipelined processing (like bucket-brigade)

Latency



New output data every clock, but a latency of 8 clock cycles
Need for data alignment

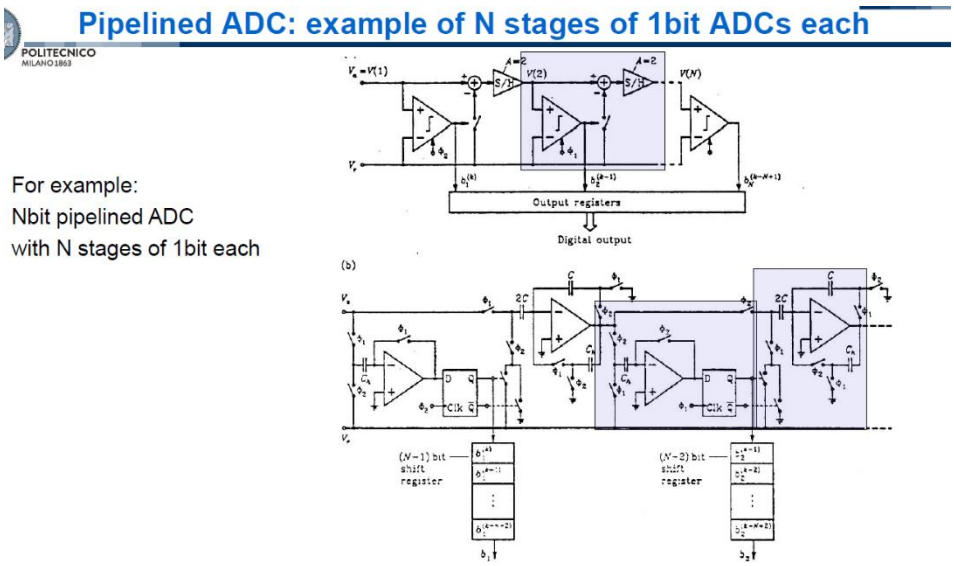
Let's exaggerate the reasoning and let's suppose that we convert the analog in with many stages. If we want a 8 bit converter, instead of using 2 stages of 4 bits each or 3 stages of 3 + 3 + 2 bits, I could use 8 stages of 1 bit each. Each stage could be a 1-bit ADC that converts the signal and gives e.g. the MSB, then we have a 1 bit DAC, the output will be amplified and fed to the second stage and so on. After 8 stages I have converted the 8 bits. Of course the first bit is ready after one clock pulse, the second after two clock pulses, and so on. In the end, after 8 clock pulses I can reach all the digital output code.

An additional thing we can do, instead of reading the bits in parallel, is to read them in a serial way, when also the final bit will be ready. To do so I use a serial bus and a FIFO.

Thanks to the pipelined architecture, once we want to acquire analog in, we don't have to wait 8 clock pulses to get the full digital bus of the conversion, because thanks to the x latches and the analog sample and hold before each stage, every time we give a clock pulse it will be converted in 8 clock pulses but the following clock pulse will allow the MSB to be reconverted, while the second MSB will be converted taking into account the previous data and not the present MSB and so on. Hence the conversion time of a single shot acquisition is 8*t_conv, but if we keep converting, it becomes t_conv, we have a valid data at every clock pulse.

Architecture

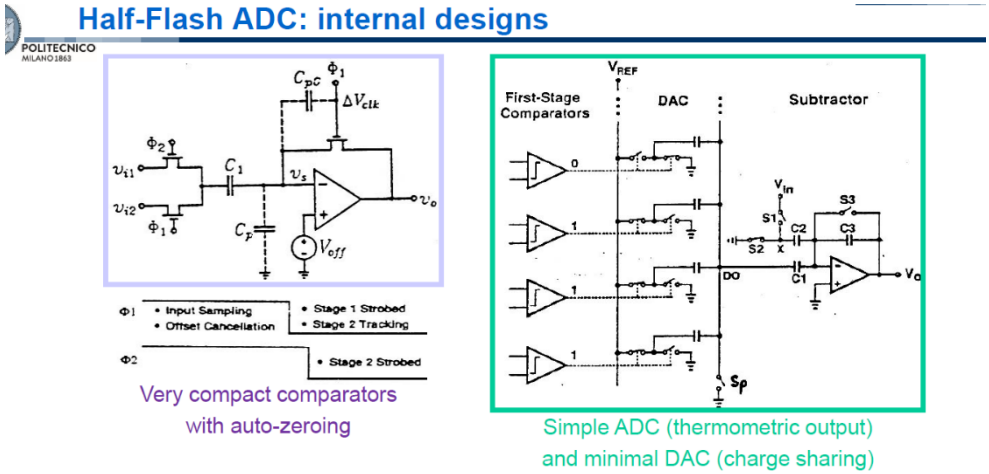
For 8 bits I need to use 8 stages, not just 2 as in the previous approach, but each stage is very simple because it deals with 1 bit.
The one below is the full schematic of the electronics.



A 1 bit ADC is a comparator with a certain threshold voltage equal to FSR/2. Then we want a 1 bit DAC, that is a switch. Then we have a subtractor, that is basically analog electronics (e.g. 4 resistors and an opamp) and then we have the S&H. In the output register we need to introduce shift registers because the value in output from each bit is at every clock pulse.

The voltage subtractor can be even simpler than a classical subtractor. It can be made by switch capacitors, instead of resistances. This is good because in this case the opamp is capable of performing both the subtraction and the hold phase of the S&H.

When we introduced the half flash, we performed a first conversion and then a second one (coarse and fine). In output of each conversion we have a line for each bit. But if we buy the components separately, we would need 3 ICs, 2 ADCs and one DAC. The thing is, we don't need to design a full ADC and a full DAC. In principle we can have the circuit below (not asked at the exam).

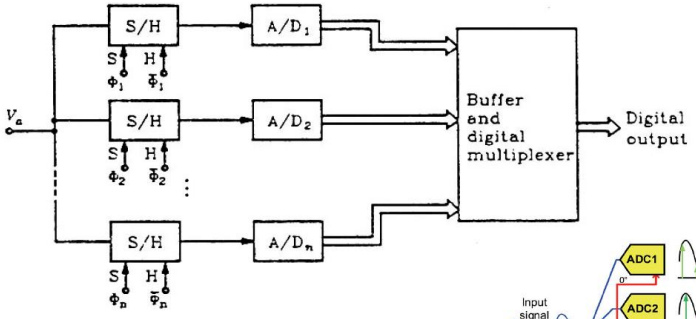


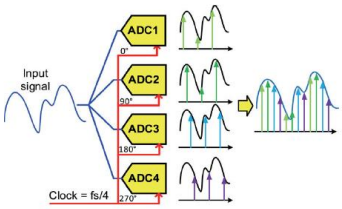
Compared to flash ADCs, half-flash ADCs minimize area occupation and power dissipation, though with slower conversion time

We need comparators. The comparator is the one in the purple box, and I connect it to the analog in. The other pin of the comparator should be the Vref.
But the comparator has its own offset, so we want to use the commutating auto-zeroing comparator. The same – terminal is used to connect either to analog in or to Vref. The capacitor C1 is used because firstly I store Vref and Vos closing the feedback switch, so that we have a buffer. So I close the feedback switch and store Vref on the capacitor. Then I close the switch of Vref and C1 charges up to Vref – Vos. Now the Vref switch is opened and the analog in is closed. The new voltage I apply is applied to C1, which was charged to Vref – Vos. The output is high or low depending if analog in + Vref + Vos is higher or lower of Vos. Vos cancels out and hence the output is the comparison between analog in and Vref, because the two are in series.

Then the comparator provides a thermometric code, so we need an encoder after the comparators to get the final 7 bits. But why should we provide a binary code to the DAC that then has to reconvert it in analog signal? We can directly use the thermometric code to feed the DAC. So the outputs of the comparators are used to drive the following DAC, that should not be a 'state-of-the-art' DAC. We use capacitors and n-channel and p-channel switches. The output after the capacitor is the analog output of the DAC. Then the next stage computes the subtraction.

TIME INTERLEAVED ADC





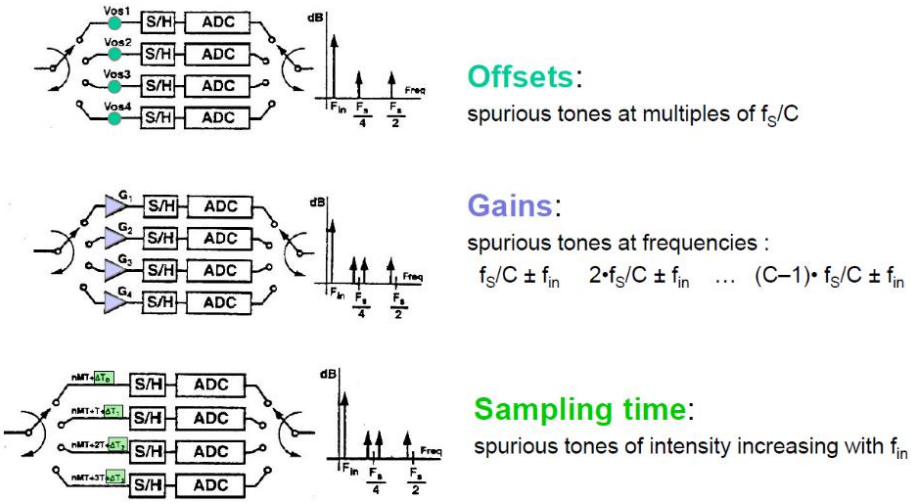
For high f_S more ADC can run in parallel on different samples:

- C channels with identical Nbit ADCs each, with individual S&H
- Each channel samples the input signal at a rate $(C \cdot T_{clock})^{-1}$
- The channel operate in sequence
- Overall sampling rate is N-fold that of a single channel

Instead of using different ADCs one after the other in a series approach, this approach relies on the parallel acquisition from multiple ADCs in parallel.
We have the analog in, then we implement different channels, e.g. 3 in the example. Each channel has its S&H, ADC and the digital output feeds a digital electronics that multiplexes one of the outputs to the digital output bus.

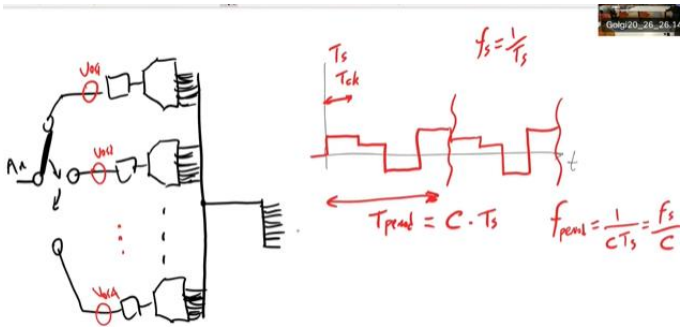
Let's suppose that we have a HF signal in the range of MHz; to properly acquire it we should need an ADC that should run in the giga-sample per second regime, so one sample every ns, so an ADC with 1 ms t_conv. Even the fastest ADC doesn't provide such speed, so we could use 4 fast ADCs and drive one or the other by time shifting the opening and closing time of S&H such that each channel acquires a sample in a specific time instant. In this way the digital output electronics will contain the conversion of each channel delayed in time and so we can provide the digital output code every $1/4^{th}$ of ns (from the bottom image).

Mismatches



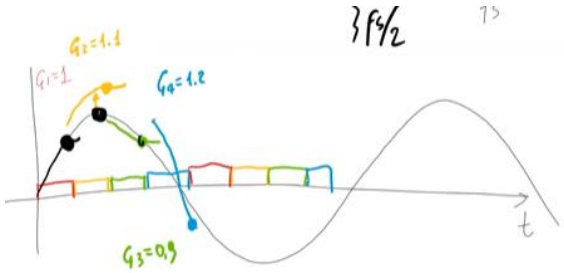
Errors now must be treated in a more complex way than in the standard case. Let's suppose that each channel has its Vos. They are DC errors, so the output error that we can find is a DC value. So theoretically we can subtract a constant number from the digital output but it is not like this.

We have different channels and in the time domain, even if the input is 0, the output will be the offset voltages of the different channels, repeating in time. The error repeats with a periodicity equal to the number of channels per the sampling time Ts.



This means that the digital output code won't have a constant error, but a periodic one. Because of this offset error, new harmonics appear in the spectrum, with a fundamental tone given by Fs divided by the number of channels.

Instead, in case of gain error, we know that each S&H has its own gain. So if we have the signal and each channel has its own gain, the input might suffer from a magnification or a decrease, so we would convert the wrong value.

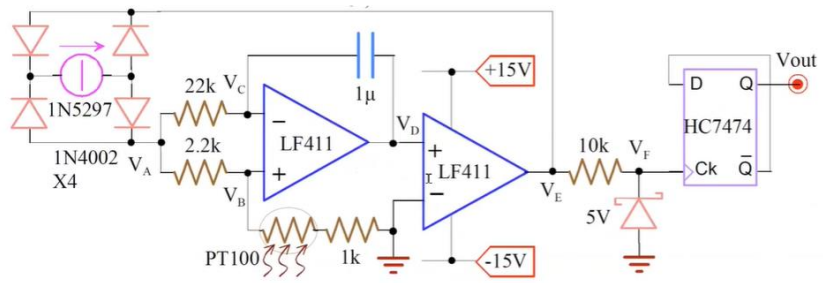


This is for sure not a constant error, because it depends on the Vin. So to model it, we see that every time a channel acquires the signal, the error we cause is an amplificated version of the input signal. We are suffering from an amplitude modulation, that happens every 4 times T_clock. So on the top of fs we modulate the amplitude, so we have f_in that appears on the left and right of each harmonic of fs, fs included.

Eventually, there is a third error that is the time jitter, which can also be modelled as an amplitude modulation of the input signal.

EXAMPLE OF WRITTEN TESTS

Ex. 1



Il 1N5297 è un generatore di corrente integrato che genera 1mA.
Il sensore termico è una PT100 che vale 100Ω a 0°C e varia di circa 0.4Ω/°C.
Studiare il funzionamento del circuito a 0°C.

a) Spiegare il funzionamento del circuito. (Si suggerisce di partire dall'istante in cui VE=+15V, per poi ricavare le correnti nelle resistenze da 22kΩ e 2.2kΩ e le tensioni VB e VC.)

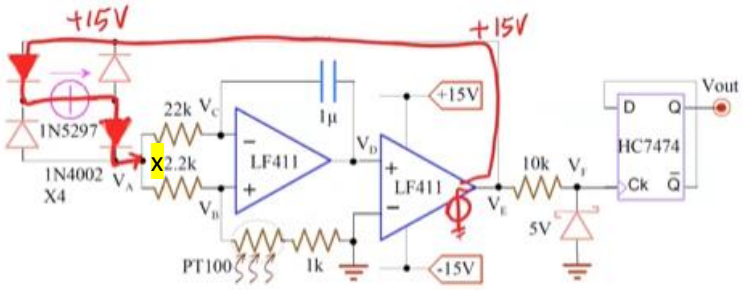
b) Disegnare le forme d'onda quotate delle tensioni ai nodi e ricavare la frequenza di commutazione di Vout.

c) Dire come si modificano le forme d'onda al variare della temperatura misurata dalla PT100.

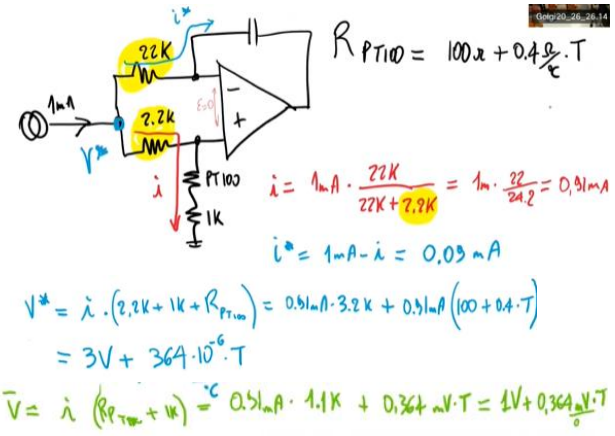
Resolution

The first opamp has a capacitor in feedback that should be an integrator. Then the second opamp could be a comparator. But we have also an overall feedback, so maybe we have an overall negative feedback. Because of the integrator, I can suppose that the signal in the forward branch changes slowly. Hence it is not a negative feedback circuit, but maybe an oscillator because of this delay.

If I still suppose that the circuit is a comparator, the output of the comparator can be one of the two PS. Let's suppose that we have +15V in output when we connect the PS. Let's suppose the C is not charged (0V across it). If this is the case, Vb is free to move, while Vc not because it is VG, to it is the copy of Vb. The current should go from high voltages to low one, so in the upper left diode and bottom right one.



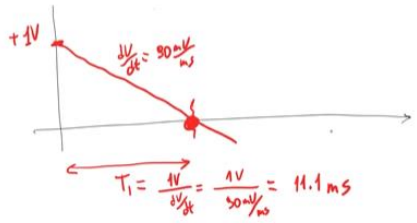
This current must be 1 mA because we have the current generator. At node x we have two paths. Thanks to virtual ground, node Vb and Vc are like connected together, so in parallel.



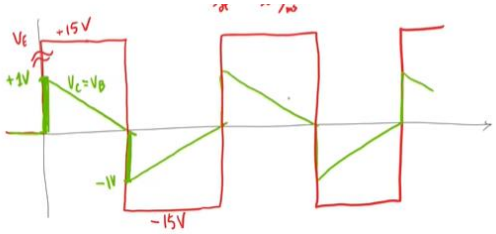
V_bar is the voltage at the + terminal of the opamp.
Then the i* flows in the capacitor.

$$\frac{dV}{dt} = \frac{I}{C} = \frac{i^*}{C} = \frac{0.09\,mA}{1\,\mu F} = 90\ \frac{V}{s}$$

If so, the voltage Vd is decreasing. It starts from 1V when the capacitor is discharged and then it decreases. We are interested in when the 0 is crossed to trigger the comparator. It is the delta in voltage divided by dV/dt.



Once we crossed this voltage, Vd goes below 0 and the output commutes to -15V. If so, the other diodes are activated, and still we have 1 mA, but in the other direction into node x.
Hence when we commute V_bar goes to – 1V (and also Vd) and the capacitor will start charging in the opposite direction, up to the point where the comparator is triggered again.



V_bar varies with temperature, so the period eventually will be longer → the oscillation frequency depends on the temperature.

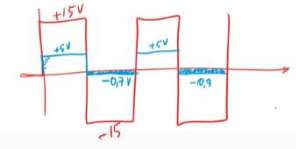
$$f_{osc}(0°C) = \frac{1}{2 \cdot 11.1\,ms} = 45\,Hz$$

$$f_{osc}(T) = \frac{1}{2\,\frac{1V + 0{,}364\,mV \cdot T}{90\,V/s}} = \frac{45\,V/s}{1V + 0{,}364\,mV \cdot T}$$

$$T_{osc}(T) = \frac{1V}{45\,V/s} + \frac{0{,}364\,mV}{45\,V/s} \cdot T =$$

$$= 22.2\,ms + 8.1\,\mu s \cdot T$$

Then we have the Zener diode, a 5V one. When the voltage is positive, it sets to +5V. When negative it is off and so it will be more or less 0.7V.



The flip flop has the D connected to the Q_not, so it behaves like a toggle flip flop.

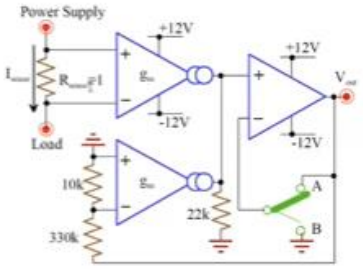
Ex. 2

Es. 1

Il circuito impiega due OTA aventi transconduttanza g_m=1mA/V e monitora la corrente in una R_{sense}=1Ω.

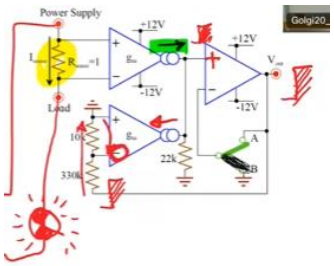
a) Quando l'interruttore è in posizione B, determinare la funzione del circuito ed il legame tra V_{out} e I_{sense}.

b) Quando l'interruttore è in A, dire **quantitativamente** cosa cambia rispetto al caso precedente.



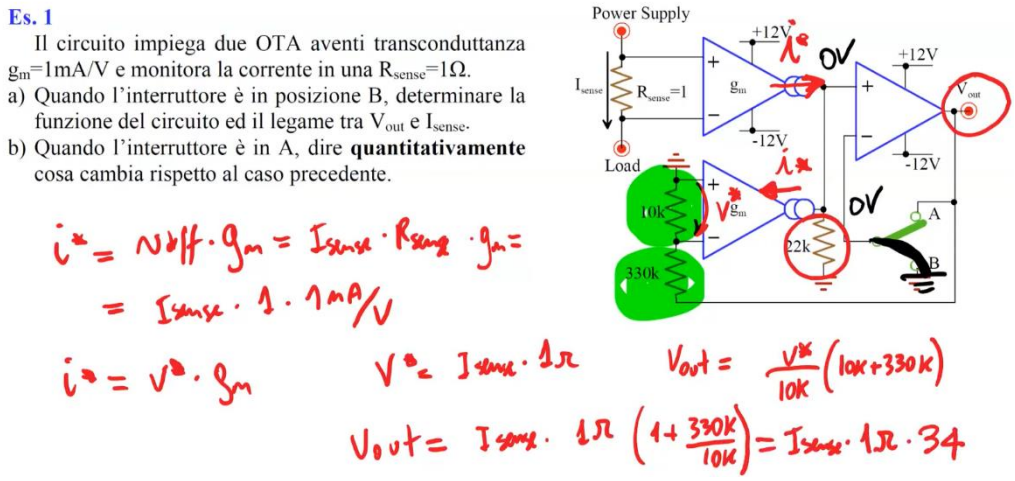
Resolution

The switch has two different positions, let's start from position B and study the schematic. We have the PS and we need to check the negative feedback.

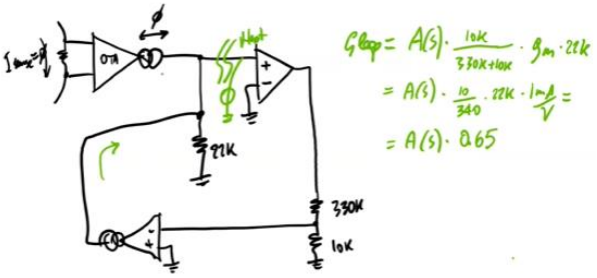


If the feedback is negative we can consider the circuit as ideal, with Gloop = inf and then we check later if it is infinite or not. If Gloop is infinite, we should have 0V on the + terminal of the third opamp due to negative feedback.

This means that the two currents i* must be equal if we have a negative feedback. Given this, we can compute Vout.



This is in the ideal case, but we need to check how big Gloop is. To compute Gloop we switch off the PS and we have the following.



Since Gloop is very high we can consider it ideal.

Let's now study the case when the switch is connected in position A, so the third opamp is a buffer. The circuit has a negative feedback.

The relationship between out and Isense should remain the same, but if we compute Gloop, it is ony 0.65.

$$G_{loop} = 1 \cdot \frac{10k}{330k+10k} \cdot g_m \cdot 22k = 0.65$$

This means that it is not strong enough, so the stage cannot be considered ideal and the relationship between Vout and Isense of the case B is not valid.
Given this, we should compute the real gain.

$$G_{real} = \frac{V_{out}}{I_{sense}}(F) = \frac{G_{ideal}}{1 - \frac{1}{G_{loop}}} = \frac{34\ V/A}{1 + \frac{1}{0.65}} = \frac{34\ V/A}{1+1.5} = 13.6\ V/A$$

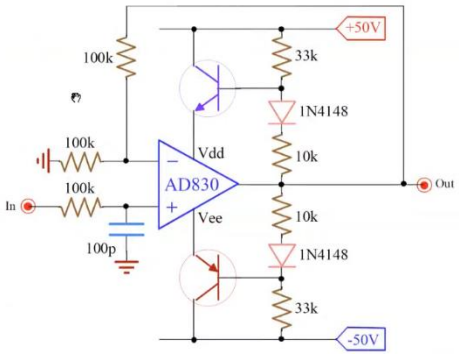
Ex. 3

Es. 3

Usando un OpAmp con $V_{supply,max}$=±14V, per potere erogare in uscita delle tensioni tra ±50V, ossia oltre i ***maximum absolute ratings***, si impiega il bootstrap.

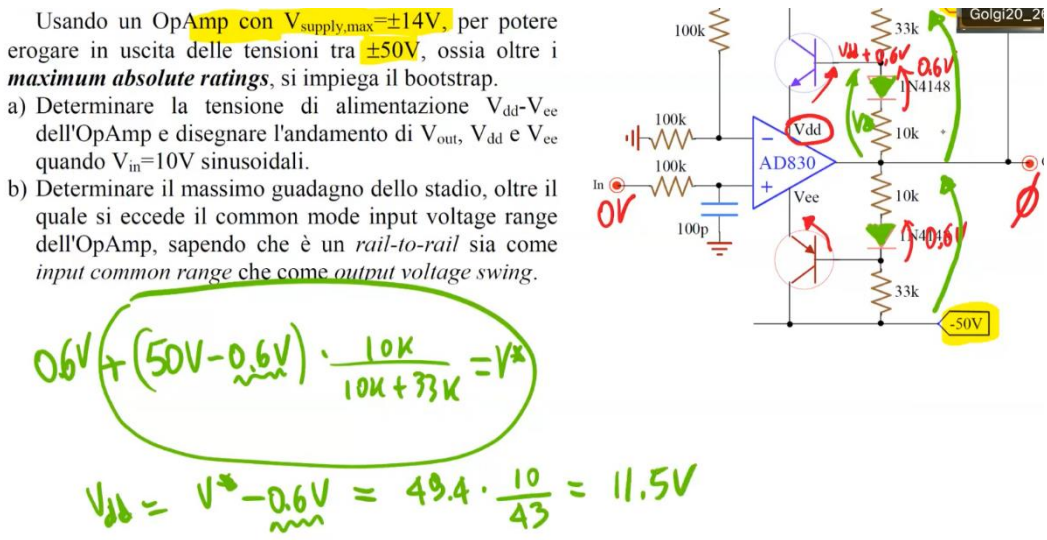
a) Determinare la tensione di alimentazione V_{dd}-V_{ee} dell'OpAmp e disegnare l'andamento di V_{out}, V_{dd} e V_{ee} quando V_{in}=10V sinusoidali.

b) Determinare il massimo guadagno dello stadio, oltre il quale si eccede il common mode input voltage range dell'OpAmp, sapendo che è un *rail-to-rail* sia come *input common range* che come *output voltage swing*.



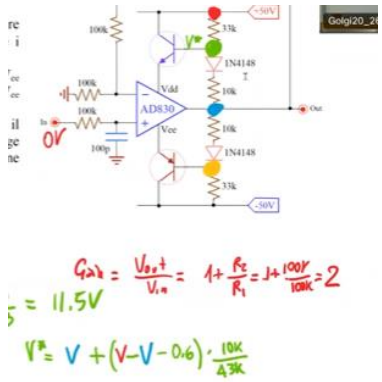
Resolution

The PS to the opamp is given by two transistors, a npn on the top and pnp on the bottom. The network made by resistors and diodes senses the Vout and provides the voltage to the bases of the transistors. The maximum PS of the opamp is +-14V, but I'm applying +-50V.
Let's suppose the input signal is 0, and the circuit has an overall negative feedback, so VG applies. The voltage drop on the two diodes is of 0.6V, that is similar to the one we have on the base emitter junction. Voltage of the base will be Vdd + 0.6V. Voltage across the top 10k resistor will be V = (50-0.6)*(10/(10+33)).

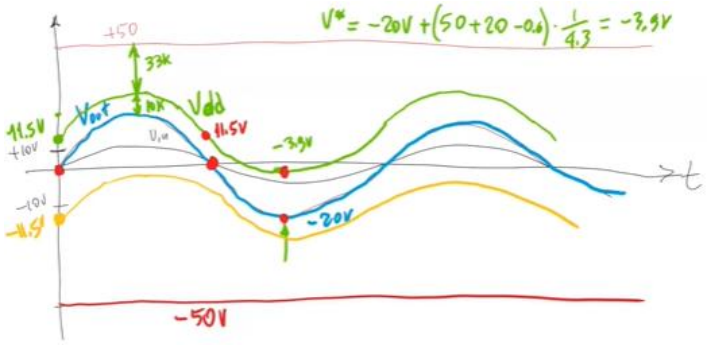


The same can be computed for Vee, and it is -11.5V. Hence the resistive network we have is needed just to provide the correct PS to the opamp → the opamp works correctly.
But let's now see what happens if we apply Vin. The Gain is the one of a non inverting stage, so 1 + R2/R1 = 1 + 100k/100k = 2.

We should recompute the voltage at the bases of the bjts.



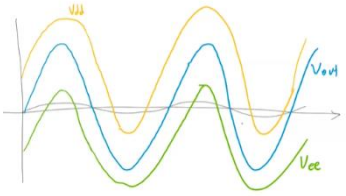
Let's plot the i/o relationship, or even better the Vout and Vin. Vout should be Vin but amplified because of the gain.



The final result is that the net accommodates the PS so that it is in between the acceptable ranges for the PS of the opamp.

The next point is to compute the maximum gain of the stage beyond which the common mode input voltage range of the opamp is no more correct.
The thing is, every time we apply a certain Vin, the PS moves accordingly to the output that is 2*Vin. But we run the risk that if Vin is no longer between the two PS, then we are applying to the opamp an input voltage that is no more between Vdd and Vee and we destroy it.

This happens if the gain is large, and the situation is the one below.



We see that in some points Vin is below Vdd and in some other points is above it, and this is not good because the opamp could be damaged.
By using one of the previous equations we can compute the maximal Vout value.

$$Vdd = (50 - G \cdot V_{in} - 0.6) \cdot \frac{10K}{43K} + G \cdot V_{in} \geq V_{in}$$
$$49.4 - G \cdot V_{in} \geq V_{in}(1 - G) \cdot 4.3$$

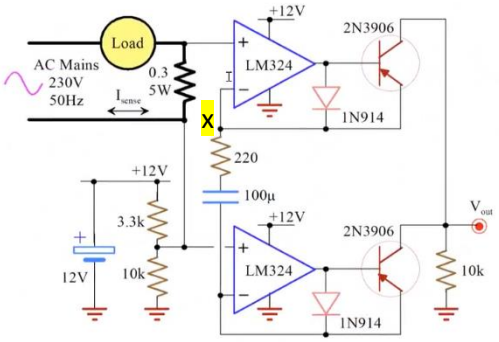
Ex. 4

Es. 2

Il circuito monitora l'assorbimento di un carico in alternata a 50Hz. L'alimentazione è flottante a batteria.

a) Disegnare la forma d'onda quotata di V_{out} nel caso sia collegato un carico resistivo di 50W e determinare il guadagno V_{out}/I_{sense}.

b) Realizzare l'alimentazione continua del circuito (circa 12V) partendo dalla tensione di rete, senza usare alcun trasformatore.

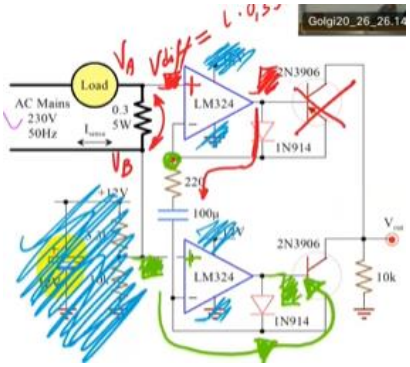


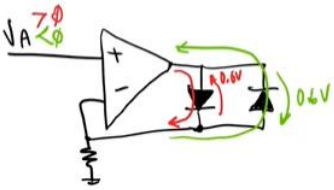
Resolution

PS is floating thanks to a battery, and let's forget it for the moment. A current flows through the bold part of the circuit and it is connected to the mains of the sinusoidal PS. So it can be clockwise or counterclockwise in direction. Voltage Vdiff across the 0.3 Ohm resistance is Vdiff = i*0.3. Then shit voltage is applied to the two opamps and let's see if they have or not a negative feedback.

If something positive is applied to the + terminal of the upper opamp, the output goes positive and if so the current flows through the diode that is on, and the pnp transistor is off, because the base is at higher voltage than the emitter and it is n doped.

The same concept applies at the bottom opamp. the + terminal goes down, and the output goes negative. But the voltage at node x is positive, current would like to flow to bottom, so it must enter in the transistor and not in the diode.

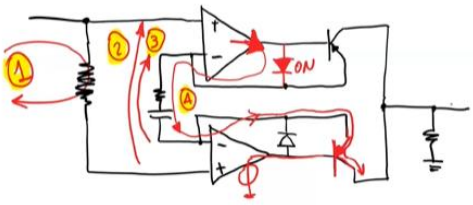




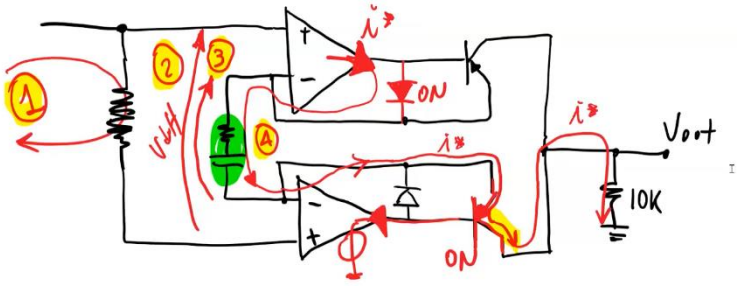
It's like having a situation on the left. Hence we have a buffer.

When the current from the mains flows clockwise in the sensing resistor, then the voltage Vdiff is connected as below.

Current cannot flow in the diode of the bottom opamp, it should go in the bjt but it cannot if the transistor is connected like in the image.



Almost the full current i* flows in the output resistor. In reality, in a BJT, Ic is beta*Ib and Ie = Ic + Ib.

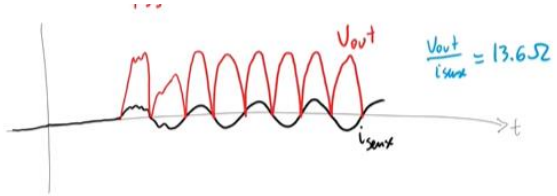


$V_{diff} = i_{sense} \cdot 0{,}3\Omega$

$i^* = \frac{V_{diff}}{Z_{in}} \simeq \frac{V_{diff}}{R} = \frac{i_{sense} \cdot 0{,}3\Omega}{220\Omega} = \frac{i_{sense}}{733}$

$V_{out} = + \frac{i_{sense}}{733} \cdot 10K = + i_{sense} \cdot 13{,}6$

Even when the current in the load changes direction, the reasoning is the same but the direction in the output resistance is the same. We can plot what is happening in the circuit, that is the fact that the input signal is magnified and brought positive. The gain is a transimpedance gain.



To the opamp, we should apply the proper power supply.

Ex. 5

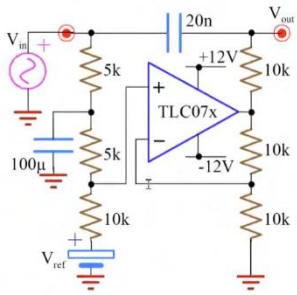
Es. 1

L'OpAMp ha A_0=120dB, GBWP=10MHz, I_{OS}=100pA, V_{OS}=0.5mV. L'uscita è lasciata aperta.

a) Determinare V_{out}/V_{in} in continua e ad alta frequenza ed il guadagno V_{out}/V_{ref}. Disegnare il trasferimento $V_{out}/V_{in}(f)$ e commentare il ruolo del circuito.

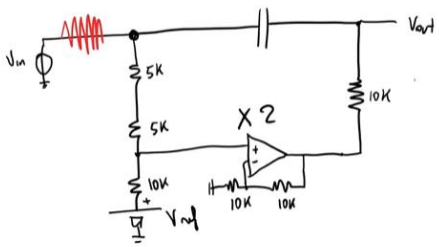
b) Calcolare l'impedenza di uscita del circuito a 10Hz quando V_{in}=+2V.

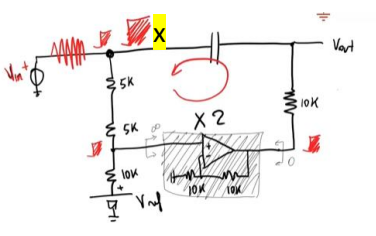
c) Calcolare l'errore in uscita dovuto all'offset di corrente (NON alle I_B) ed a quello di tensione.



Resolution

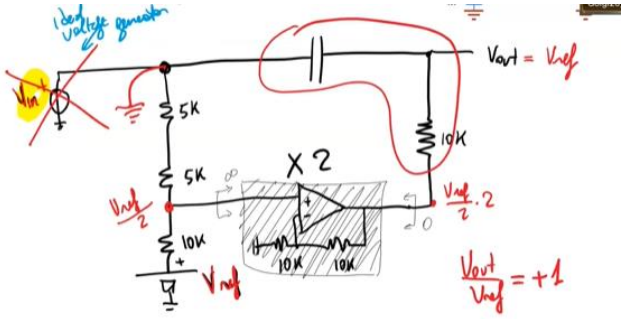
For sure there is a local negative feedback, but in the outer loop we don't have any overall negative feedback loop because we have a voltage generator that is touching the feedback loop, and whenever this happens we kill the loop. We can rewrite the schematic as below.



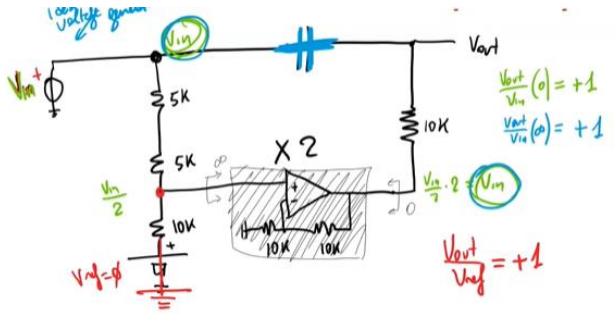


If there was the red resistance it would have been true that we have an outer feedback loop, that however would have been positive and I would have had a Schmitt trigger. But since there is no resistor, we apply a Vin to the loop and we cannot move node x because it is an ideal voltage generator, so there is no feedback at all, just an open circuit.

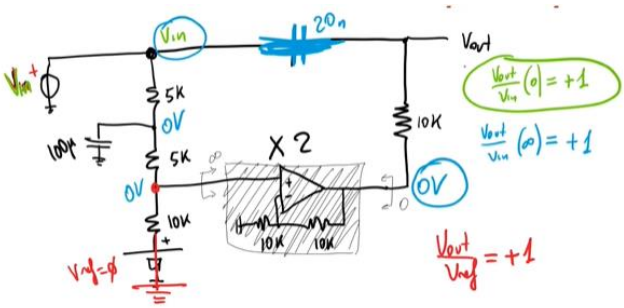
Let's kill Vin and compute Vout due to Vref, so we ground Vin. Vref is in DC so the capacitor is open and Vout = Vref.



Let's now consider Vin, so we switch off Vref (grounded). Across the capacitor and the 10k resistor we have Vin and Vin, but Vin has a frequency dependance. At DC, the capacitor is open so the output is open. At HF, Vout/Vin is still Vin.

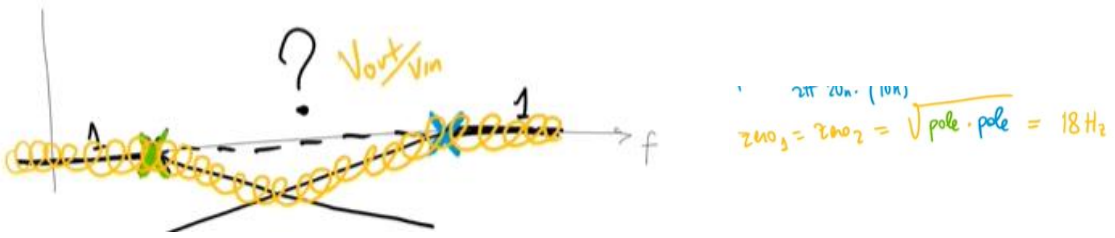


The thing is, in the circuit of the exercise there is another capacitor of 100u. At DC nothing changes for Vout/Vin, while at HF the relationship is still 1 because Vin feeds the output.



Let's study the i/o relationship for Vin. We saw that at DC the gain is 1 and also at HF. But is it always 1? We have to check the poles and who enters in shortcircuit.
The pole due to the 100u capacitor is 1/(2*pi*100u*R), where R is 5k||(5k + 10k). The pole is at 0.42 Hz. Maybe this capacitor also introduces a zero. Every time we enter and we have a capacitor in the path of the signal, the capacitor is not introducing a zero.

As for the 20n capacitor, at HF, we have a HP filter whose pole is 1/(2*pi*20n*10k) (Vin must be grounded). The frequency of the pole is 796 Hz, and then we have a zero at 0Hz. So the circuit is like having a LP stage in parallel to a HP stage. When they are both working, we have the product of them.

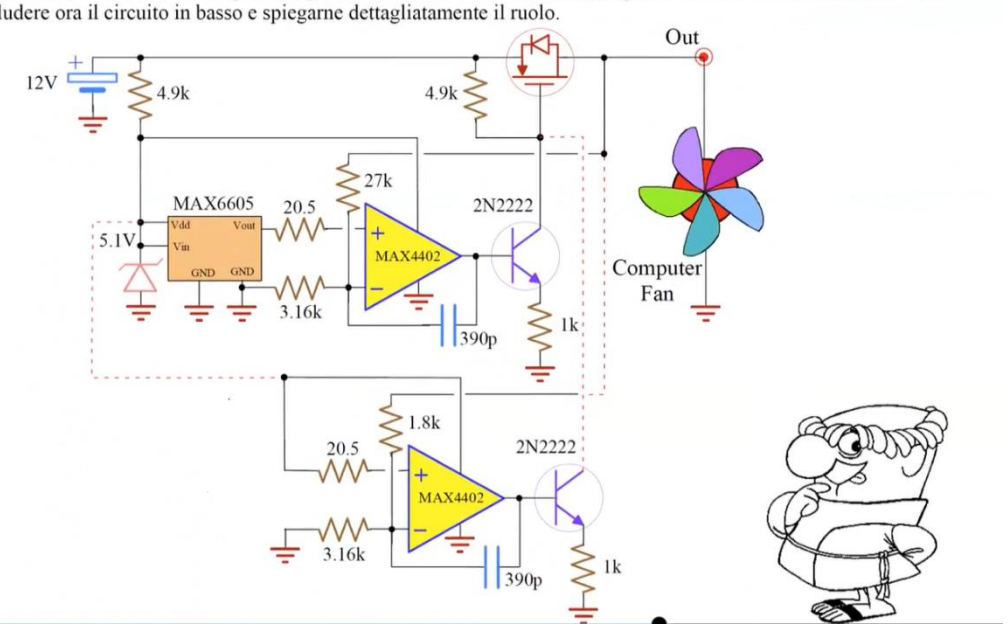


We can see that the circuit has two coincident zeros at the intermediate frequency (geometric average).

Ex. 6

Il sensore di temperatura MAX6605 fornisce $V_T(T)=744mV+T\cdot 11.9mV/°C$.

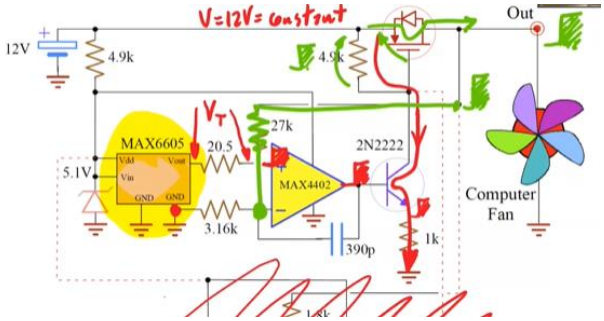
a) Scollegando il circuito in basso, disegnare il grafico quotato di $V_{out}(T)$ nell'intervallo 0°C÷50°C.
b) Determinare l'intervallo di temperature per cui il circuito è in zona lineare, sapendo che il MAX4402 è *rail-to-rail*.
c) Includere ora il circuito in basso e spiegarne dettagliatamente il ruolo.



Resolution

We have a temperature sensor that is the MAX6605. At first let's not consider the bottom circuit and see how the circuit behaves.

We have a +12V battery and a Zener diode. So if the battery is higher than the voltage of the Zener diode, the Zener diode is clamping the voltage to 5.1V. This voltage is used to bias the temperature sensor and also the opamp. The 3.16k resistor is connected to ground, while Vt varies with temperature. V+ of the upper opamp is equal to Vt because no current can flow in the input terminal. If something positive is applied to the + terminal, the output goes positive; if so, then also the emitter goes positive and the current increases, and if so the voltage across the 4.9k resistor is increasing. Then the Vgs of the mosfet increases, so also the current in it increases. Then Vout increases to the load. If so, also the – terminal of the opamp increases → negative feedback.



If Gloop is infinite, we can consider the ideal feedback and the concept of virtual ground, so on the two terminals I should have Vt. V- is given thanks to a voltage partition with Vout as top voltage.

$$V_{out}\cdot\frac{3.16k}{3.16k+27k}=V_T$$

$$V_{out}=\left(744mV+11.9\frac{mV}{°C}\cdot T\right)\cdot 9.5=$$
$$=7.1V+113\frac{mV}{°C}\cdot T$$

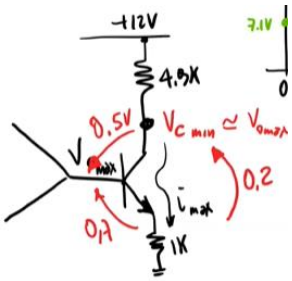
So Vout is proportional to the temperature plus a constant offset of 7.1V.

Theoretically, if we apply the maximal temperature of 50°C, the voltage should reach 12.8V, but this is a problem for the bias. But Vout cannot reach 12V because the transistor requires something between source and drain.

We want to compute the range of temperatures withing which the circuit can operate in the linear zone. The maximal voltage in output of the opamp is 5.1V because the opamp is rail-to-rail. So we can compute the current in the bjt.

$$i = \frac{5.1V - 0.6}{1K} = 4.5mA$$

Then for the mosfet, Vgs = i*4.9k = 4.5mA*4.9k = 22V. This means that the opamp will never saturate to the power supply because the current in the BJT would be so high that the Vgs of the MOSFET would have been too high and the collector would decrease too much.



The limiting factor of the BJT is when the collector is so low that we don't have anymore a reverse bias junction between base and collector.
When Vo in output of the opamp is max, we have the maximum current in the BJT and the minimum Vc of the bjt. When Vc is almost Vo,max, the transistor enters the saturation regime, when Vce is almost 0.2V and the base collector junction is forward biased with 0.5V.

The condition for not having saturation is the following.

$$12V - \left(\frac{V_{omax} - 0.7}{1K}\right) \cdot 4.9K \geq V_{omax} - 0.5$$

Performing the calculations:

$$12 - V_{omax} \cdot 4.9 + 3.43 \geq V_{omax} - 0.5$$

$$V_{omax} \leq \frac{16}{5.9} = 2.7V$$

We know hence the maximal Vo in output of the opamp.

Let's now include also the circuit at the bottom of the schematic. Also this part senses the same voltage as the circuit above, and they drive the same nodes. The only difference is the resistor of 20.5 that is not connected to Vt but to PS. So if we perform the voltage partition with Vout as top voltage as before, we have:

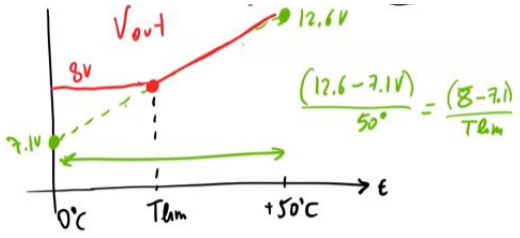
$$V_{out} \cdot \frac{3.16k}{3.16k + 1.8k} = 5.1V$$

The result is that Vout = 8V.
In the end the bottom circuit sets Vout to be 8V, the top one to 7V + 113mV/°C * T.
If the bottom BJT is capable of providing the proper current the output will reach 8V, but if for any other reason there is another stage that pumps a different current to set a different voltage, one of the two transistors will not operate properly. If e.g. T = 0°C, the top stage tries to force 7V, the bottom stage 8V. If the voltage on the collector of the bottom BJT is 7V, it means that the voltage on the – terminal of the bottom opamp will be lower than 5.1V on the + terminal. Hence the output of the opamp tends to increase, but if so, the current in the BJT increases, so the voltage drop on the Vgs of the mosfet increases and the current in the mosfet increases and so Vout increases. So Vout cannot stay at 7V because the bottom stage increases the current so that the output voltage is brought to 8V.
If the output is 8V, these are for top and bottom stages. The bottom stage is ok, but on the top opamp the voltage on the – terminal will be higher than the voltage on the + terminal and so the output of the opamp will go low decreasing the current in the BJT. So the Vgs of the mosfet decreases and so on.

At the most the upper stage can decrease its current, eventually even putting it to 0, but there is no problem because the current that sets the Vgs of the MOSFET is set by the bottom circuit.

So the second stage forces a constant 8V contribution to the output. Due to the second stage, the output will be 8V even for temperatures lower than a given amount.
If we change the values of the resistors of the bottom stage, e.g. by introducing a trimmer, we can change the Vout,min, and so the speed of the fan.

The i/o relationship we obtain will be the following.



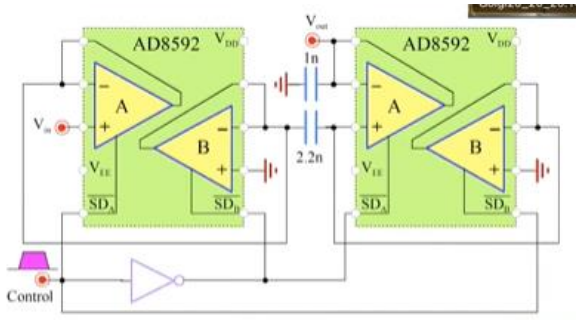
Ex. 7

Es. 2

I piedini di ShutDown *active-low* mandano in *three-state* la rispettiva uscita. Il segnale Control è a 1kHz.

a) Determinare la funzione svolta dal circuito e il legame tra V_{out} e V_{in}. Si consiglia di studiare separatamente il caso Control=*low* e poi Control=*high*.

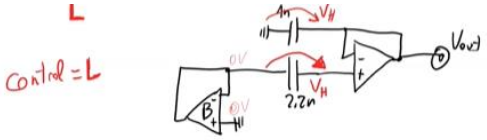
b) Dire cosa accade quando Control=*low* per 1ms e si considerano tutte le I_B=100pA.



Resolution

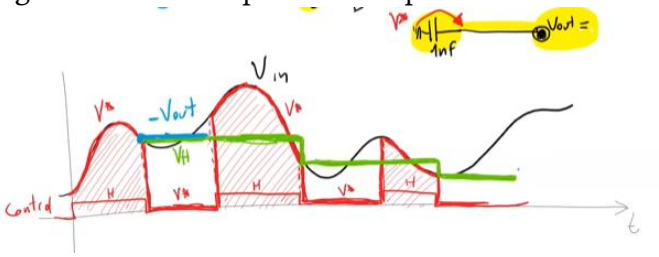
We have two ICs with two opamps inside, that can be turned on or off through a control pin SD, that is active low.
When the opamps are not active, the output is in tristate, so it is fixed.

Let's start from the control pin set to low level. Opamps A1 and B2 are off.



As for when control is high.

Hence during control = H we store Vin on the capacitor. When control = L, Vin is no more attached to the capacitor and the voltage is fed to the output. In the L phase Vout = -Vin.



Vh is the voltage on the 1n capacitor. This is a smart S&H because the hold phase lasts longer than in the normal S&H. Moreover, this stage also performs a voltage inversion.

x x