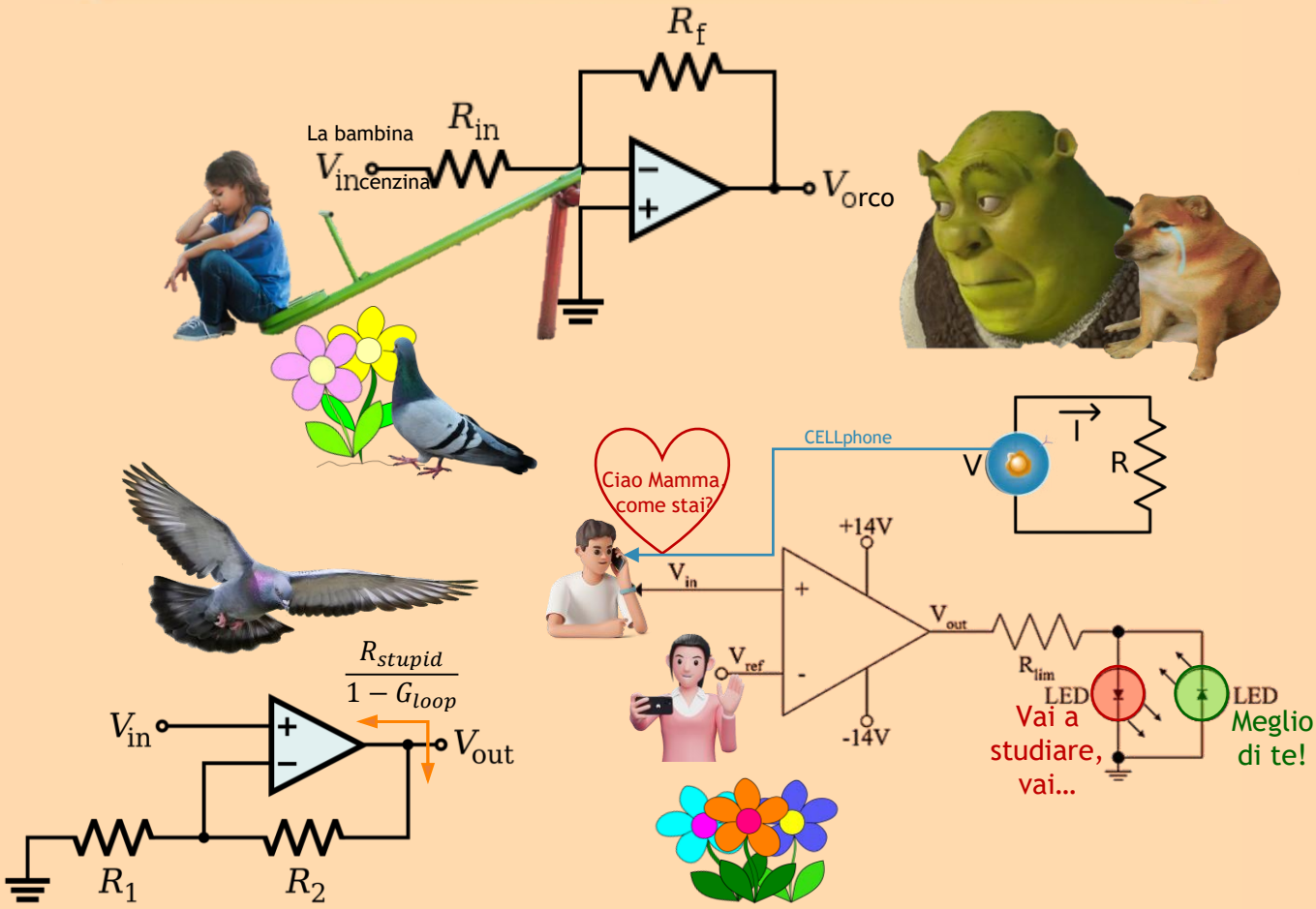
2022/2023



Electronic Systems Notes

Noise, advanced Operational Amplifiers and circuits, Sample&Hold circuits, advanced DAC and ADC converters,

Franco Zappa

SOCIETÀ EDITRICE CULAPIO

I took these notes after rewatching carefully ALL the lessons and included the book parts for missing explanations. But keep in mind there might be errors and inaccuracies.
Take a look also at the exercises and summaries collections.

« Everybody asks "Mamma, dov'è quello, mamma, dov'è questo?" , but nobody asks "Ciao Mamma, come stai?" »

«I liked this course so much that I took it 2 times » -average ES student

«ABBAndonate ogni speranza voi che vi iscrivete» -Dante

Honest, not asked for, REVIEW OF THE COURSE

A NEEDED PREFACE

I shall start mentioning that I'm an automation engineering student that took this course by choice. *Looking back, I wouldn't make that choice*.

Indeed the course it's difficult enough for electronic eng. students, so it's fair to say that my different background may have intensified the struggles. Anyway most of the problems come from the prof and not from the subject, that I personally found interesting.

LECTURES

Starting from the lectures, that are basically a compilation of repetitive sounds, ~~not funny~~ jokes, stupid and overcomplicated examples and schemes sketched at the speed of lights. The slides suck, the book too, the songs in it too.

EXERCISES

I think the prof also underestimated the exercise sessions, since, at least this year, he covered the majority of them. Indeed they're done too fast, the solutions are often dispersive and not focused on defined requests, this makes it difficult to understand the reasoning and the method behind the solution. The assistants are competent, I wished Zappa could have give them more space, and let them do most of the exercises.

EXAM AND PROJECT

The project, in my opinion, is kinda useless. I mean it's fine to learn how to use a new software, but the fact that in the end it's just used to solve a simple example make it useless. They should probably get rid of it or assign a more serious project that maybe takes up more point at the oral exam. Indeed, I think that, for the number of hours assigned to theoretical lectures wrt exercises, it would be more fair to value more the oral (not just ± 3), since the theory also requires a lot of time and the written test is 50% luck and 50% Shrek.

In the end, if you have the possibility to do some course ZAPPing, **don't** choose this course. (R.I.P. the ones who have it as mandatory).

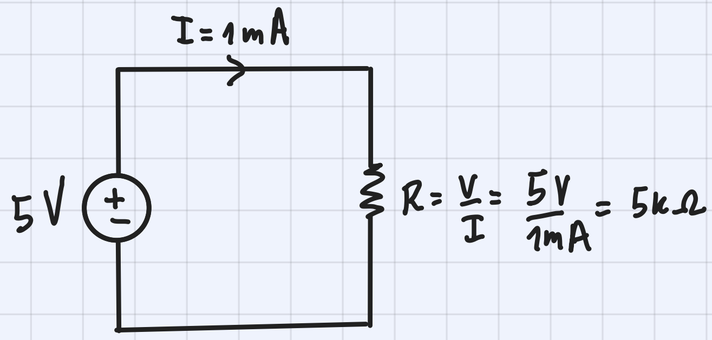
CONTENTS

ES01_ BASICS

Resistors

- Values: $0.1\,\Omega \div 4.7\,M\Omega$
- Tolerances: $10\% \div 1\%$
- Power ratings: $1/8, 1/4, 1/2, 1, 2, 5, 10$

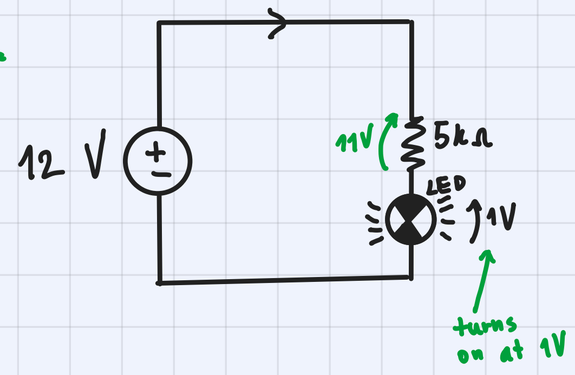
Ex.



Resistor:
- $R = 5k\Omega$
- toll = 10%
- $P = V \cdot I = \frac{V^2}{R} = RI^2 = 5mW$

Formulas valid only for the resistor

Ex.



The LED turns ON at 1V, meaning we want on the resistor a 11V voltage drop.

That means we'll have a higher power dissipation on the resistor: $P = \frac{V_R^2}{R} = \frac{11^2 V^2}{5k\Omega} = 24.2mW$

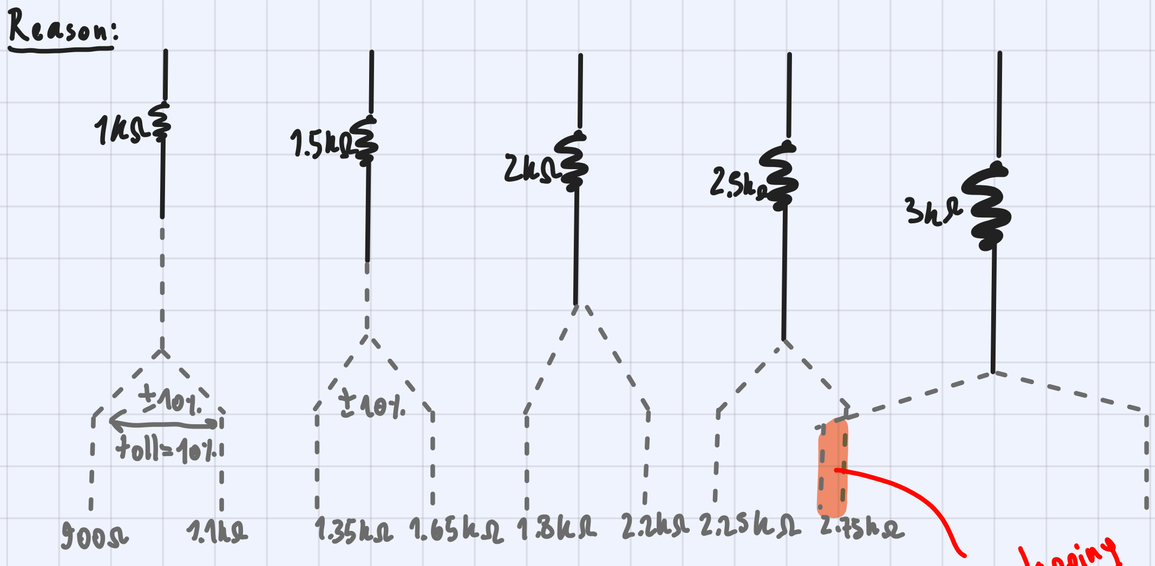
This higher dissipation may damage the circuit, the best thing to do in this case is buy a resistor with an higher power rating.

Rule of thumb for resistor's power rating choice: overrate the limit power value by a factor $\frac{3}{2}$

e.g. limit value > 5mW → Choose a resistor with power rating = $5mW \cdot \frac{3}{2} = 7.5mW$

- Preferred values:

Reason:



→ we should only produce some preferred values for resistors

Preferred values:

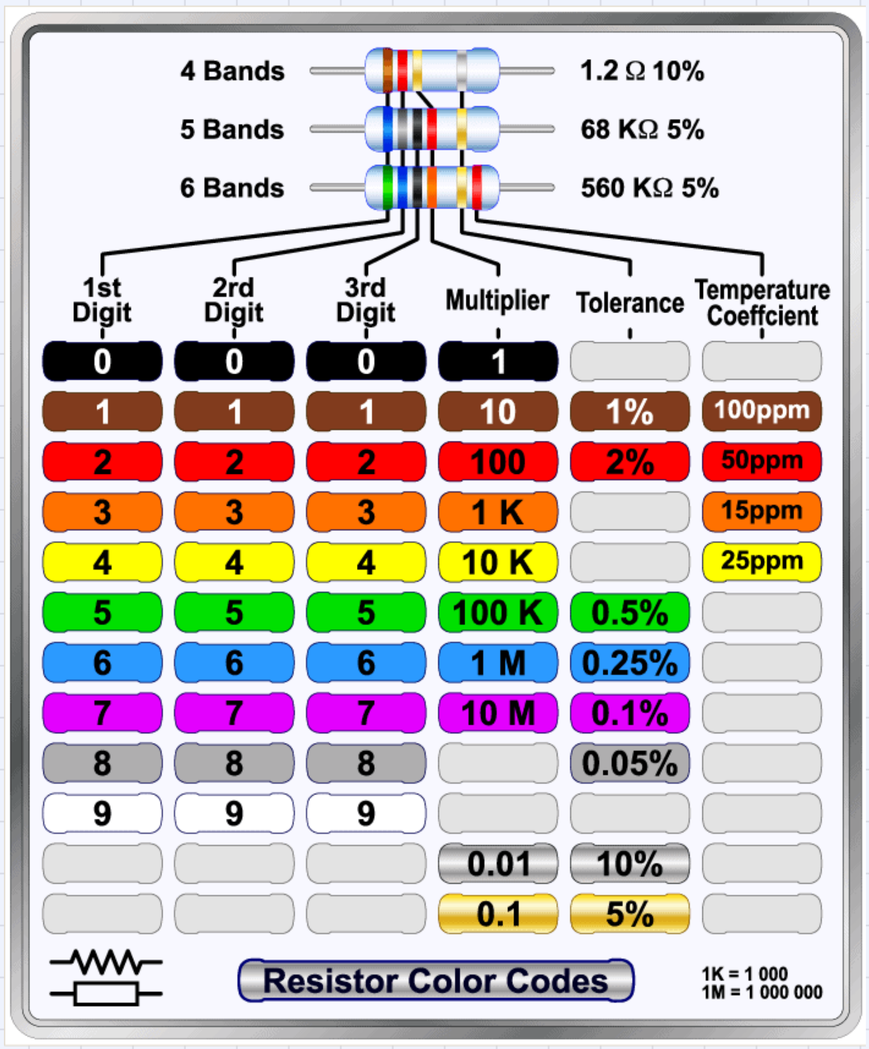
e.g. for 10%: 10, 12, 15, 18, 22, 27...

- for 5%/2%: 11, 13, 16, 20, 24,...

- for 1%: 100, 102, 105, 107...

higher the toll. higher the number of available values

Coding:



Capacitors

- **Definition:**

Capacitor:

In both digital and analog electronic circuits a capacitor is a fundamental element. It enables the filtering of signals and it provides a fundamental memory element. The capacitor is an element that stores energy in an electric field.

The circuit symbol and associated electrical variables for the capacitor is shown on Figure 1.

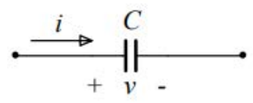


Figure 1. Circuit symbol for capacitor

The capacitor may be modeled as two conducting plates separated by a dielectric as shown on Figure 2.

When a voltage v is applied across the plates, a charge $+q$ accumulates on one plate and a charge $-q$ on the other.

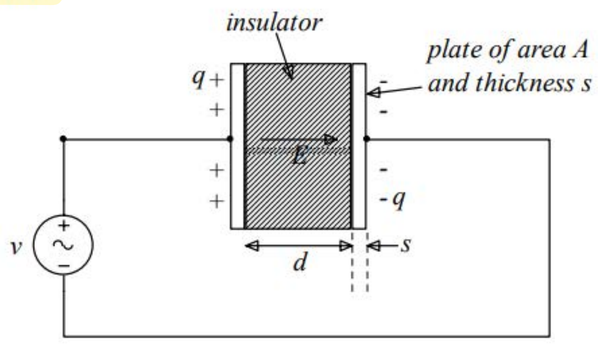
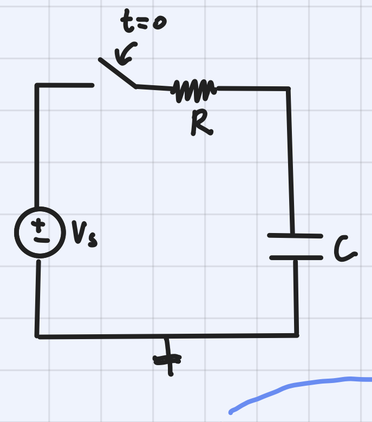


Figure 2. Capacitor model

Formulas

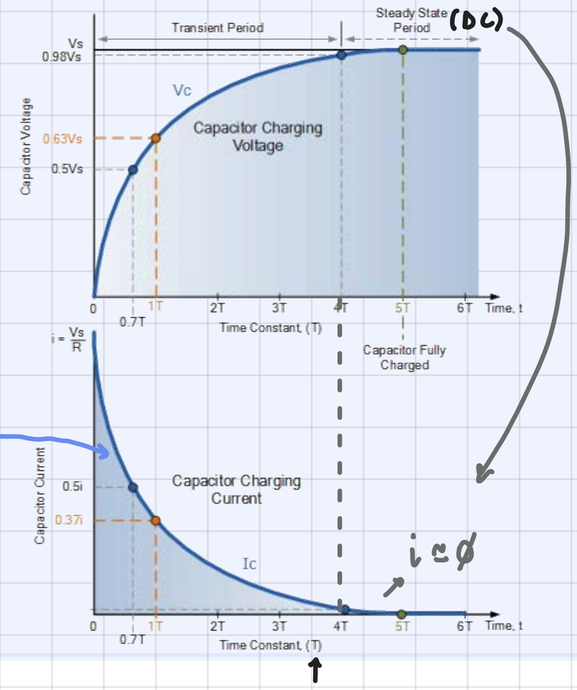
- ELECTRIC FIELD generated across the plate: $E = \frac{q}{\varepsilon A}$
- CAPACITANCE: $C = \frac{\varepsilon A}{d}$
- VOLTAGE across the capacitor plates: $v = E\,d = \frac{q\,d}{\varepsilon A} = \frac{q}{C}$
- CURRENT flowing into the capacitor: $i = \frac{dq}{dt} = \frac{d}{dt}\left(\frac{\varepsilon A}{d} v\right) = \frac{\varepsilon A}{d}\frac{dv}{dt} = C\frac{dv}{dt}$

Ex. RC circuit

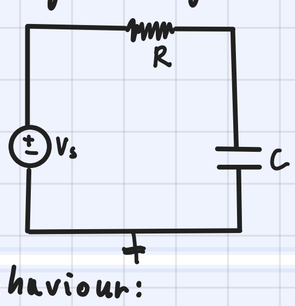


area:

$$\int_0^{+\infty} i(t)\,dt = Q = C \cdot V$$



- **Values:** $1\,pF \div 1000\,\mu F$
- **Tollerances:** $20\,\% \div 5\,\%$
- **Voltage ratings:** $5\,V \div 400\,V$



The capacitor DOESN'T dissipate power → It just charges and discharges (reactive component)

↳ We must instead specify the max voltage than can sustain (with 150% rule of thumb on specifications ratings)

- **Behaviour:**

Let's now consider the circuit shown on Figure 3 where a capacitor of capacitance C is connected to a time varying voltage source $v(t)$.

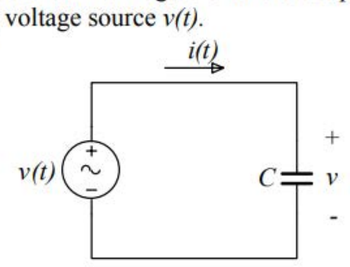


Figure 3. Fundamental capacitor circuit

If the voltage $v(t)$ has the form

$$v(t) = A\cos(\omega t) \tag{1.8}$$

Then the current $i(t)$ becomes

$$\begin{aligned} i(t) &= C\frac{dv}{dt} \\ &= -C\,A\,\omega\sin(\omega t) \\ &= C\,\omega A\cos\left(\omega t + \frac{\pi}{2}\right) \end{aligned} \tag{1.9}$$

Therefore the current going through a capacitor and the voltage across the capacitor are 90 degrees out of phase. It is said that the current **leads** the voltage by 90 degrees.

The general plot of the voltage and current of a capacitor is shown on Figure 4. The current leads the voltage by 90 degrees.

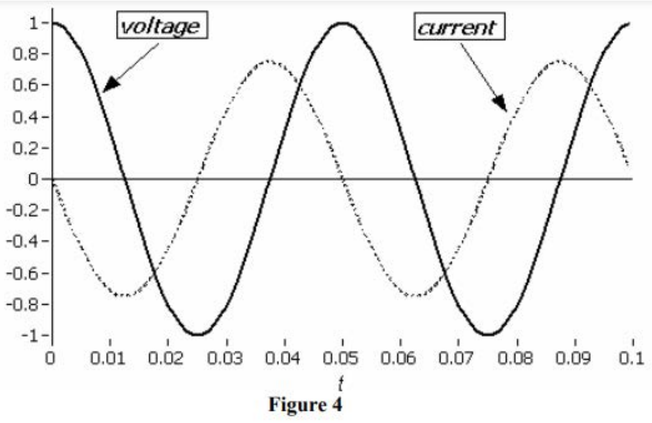


Figure 4

If we take the ratio of the peak voltage to the peak current we obtain the quantity

$$Xc = \frac{1}{C\omega} \tag{1.10}$$

Xc has the units of Volts/Amperes or Ohms and thus it represents some type of resistance. Note that as the frequency $\omega \to 0$ the quantity Xc goes to infinity which implies that the capacitor resembles an open circuit .

Capacitors dislike to pass current at low frequencies

As the frequency becomes very large $\omega \to \infty$ the quantity Xc goes to zero which implies that the capacitor resembles a short circuit.

Capacitors like to pass current at high frequencies

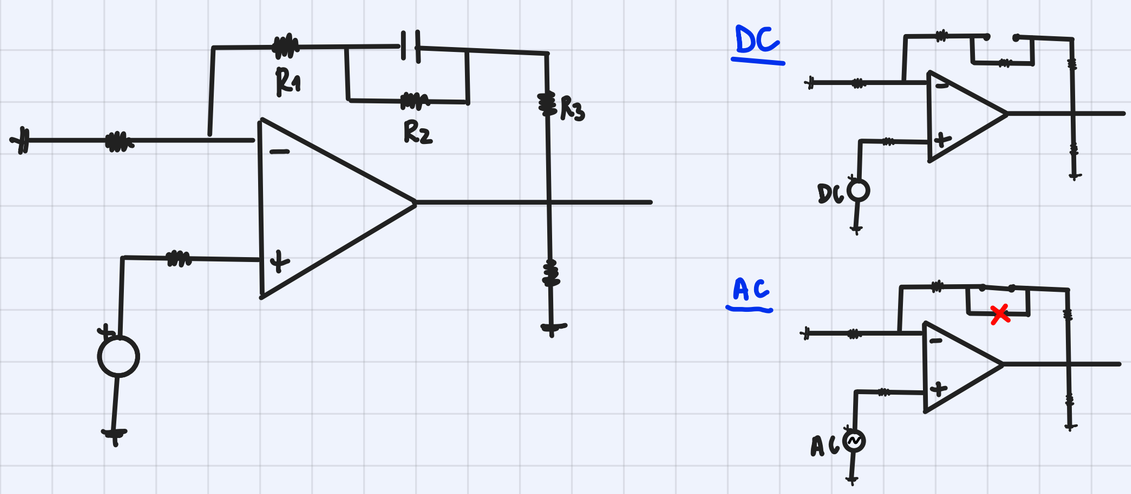
DC regime

- $\omega \to 0$ (steady state)
- DC voltage supply
- C ≡ OPEN CIRCUIT

AC regime

- $\omega \to \infty$ (high freq.)
- AC voltage supply
- C ≡ SHORT CIRCUIT

Ex. (with Op Amp)



Time constant: $\tau = C \cdot R_{eq}$

Equivalent resistance: $R_{eq} = (R_1 + R_3) \,//\, R_2$ (in parallel)

Frequency of the pole (for Bode diagram): $f_{pole} = \frac{1}{2\pi \cdot \tau}$

(freq. at which the capacitor changes from DC to AC condition)

- $f << f_{pole} \longrightarrow \omega = 2\pi f \longrightarrow 0 \longrightarrow$ DC
- $f >> f_{pole} \longrightarrow \omega = 2\pi f \longrightarrow \infty \longrightarrow$ AC

Inductors

- Definition:

Inductors

The inductor is a coil which stores energy in the magnetic field

Consider a wire of length *l* forming a loop of area *A* as shown on Figure 11. A current *i(t)* is flowing through the wire as indicated. This current generates a magnetic field *B* which is equal to

$$B(t) = \mu \frac{i(t)}{l} \qquad (1.23)$$

Where μ is the magnetic permeability of the material enclosed by the wire.

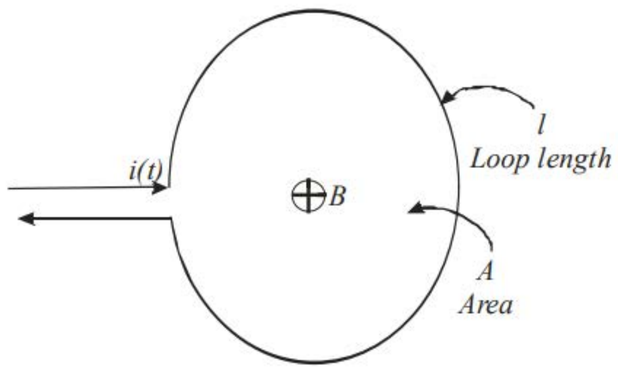


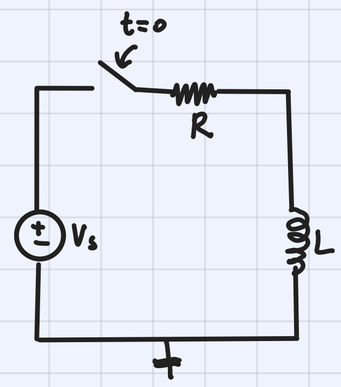
Figure 11. Current loop for the calculation of inductance

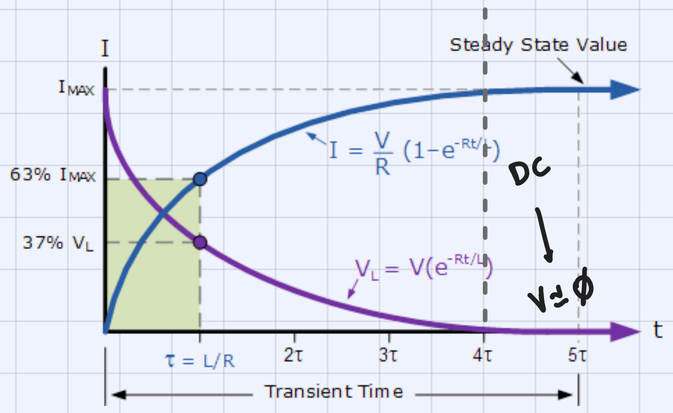
Formulas

- INDUCTANCE: $L = \frac{A\mu}{l}$
- MAGNETIC FLUX through A: $\phi = AB(t) = A\frac{\mu}{l} i(t) = L\, i(t)$
- VOLTAGE across the inductor: [Maxwell Eq.] $\frac{d\phi}{dt} = v(t)$

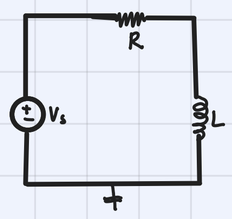
$$v = L\frac{di}{dt} \longleftarrow \frac{dL\,i(t)}{dt} = v(t)$$

Ex. RL circuit





- Values: 1 nH ÷ 100 mH
- Tollerances: 20 % ÷ 5 %
- Current ratings: 1 mA ÷ 1 A



The inductor DOESN'T dissipate power → no voltage ACROSS the inductor

↳ We must instead specify the max current than can sustain (with 150% rule of thumb on specifications ratings)

- Behaviour:

Let's now consider the circuit shown on Figure 13 where an inductor of inductance L is connected to a time varying current source *i(t)*.

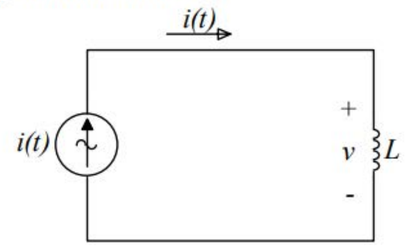


Figure 13. Fundamental inductor circuit

If we assume that the current *i(t)* has the form

$$i(t) = I_o \cos(\omega t) \qquad (1.30)$$

Then the voltage *v(t)* becomes

$$\begin{aligned} v(t) &= L\frac{di}{dt} \\ &= -L I_o \omega \sin(\omega t) \\ &= L\omega I_o \cos\left(\omega t + \frac{\pi}{2}\right) \end{aligned} \qquad (1.31)$$

Therefore the current going through an inductor and the voltage across the inductor are 90 degrees out of phase. Here the voltage **leads** the current by 90 degrees.

The general plot of the voltage and current of an inductor is shown on Figure 14.

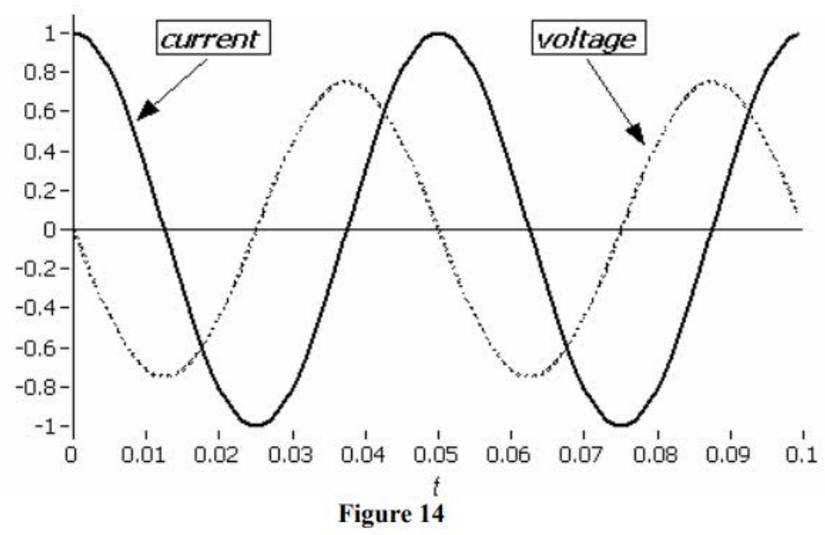


Figure 14

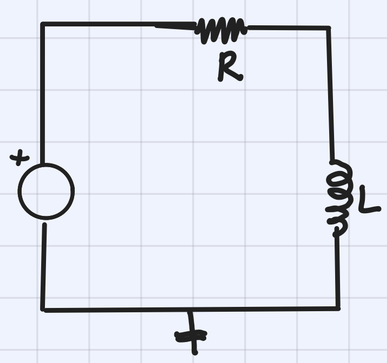
DC regime

- ω → 0 (steady state)
- DC voltage supply
- L ≡ | SHORT CIRCUIT

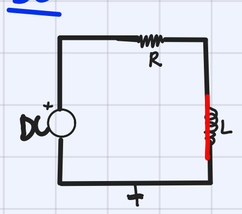
AC regime

- ω → ∞ (high freq.)
- AC voltage supply
- L ≡ ¦ OPEN CIRCUIT

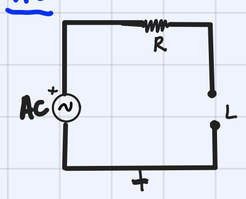
Ex.



DC



AC



$X_L = L\omega$

- $\omega \to 0 \Rightarrow X_L \to 0$ (short circuit)
- $\omega \to \infty \Rightarrow X_L \to \infty$ (open circuit)

Time constant: $\tau = \frac{R}{L}$

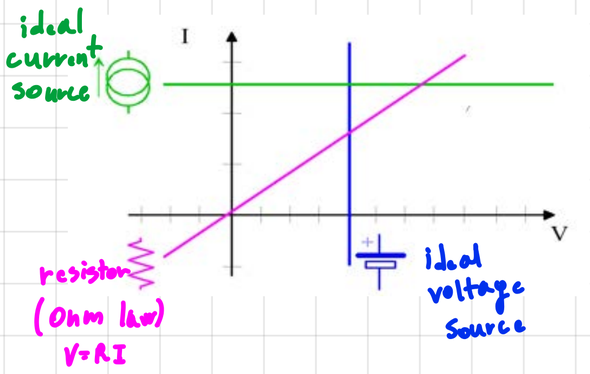
Frequency of the pole (for Bode diagram): $f_{pole} = \frac{1}{2\pi \cdot \tau}$

(freq. at which the capacitor changes from DC to AC condition)

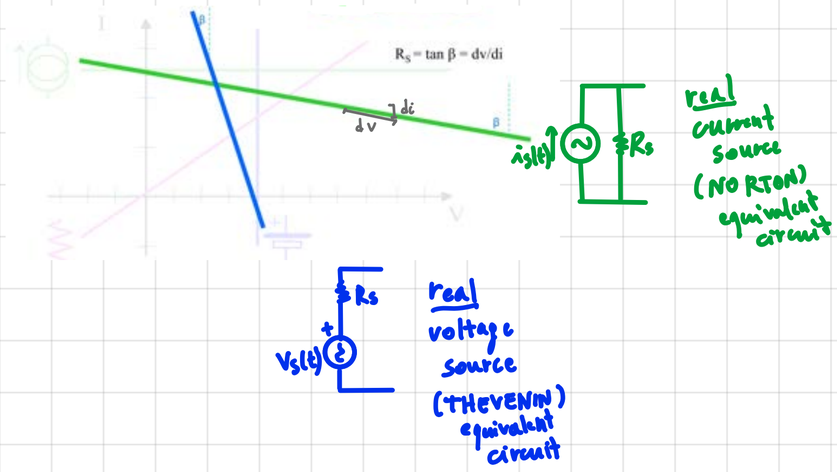
↳
- $f << f_{pole} \longrightarrow \omega = 2\pi f \to 0 \to$ DC
- $f >> f_{pole} \longrightarrow \omega = 2\pi f \to \infty \to$ AC

Sources

Ideal



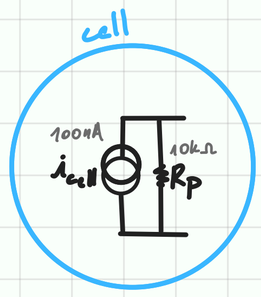
Real



ex. current source

Remember to read the V/I source with the according V/I instrument.

We have a cell ◯. Suppose that we know that the cell is a current source

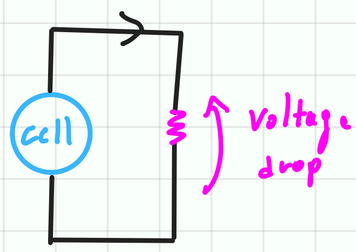


.

If:

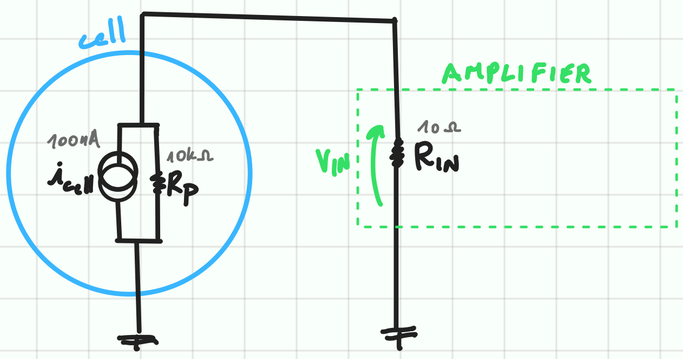
- we have just a voltmeter we have to measure the current via a resistor.

But this causes a voltage drop on the resistor, but this developed voltage across the cell can damage it.



(the cell, to properly work may need a voltage drop!)

- we want to measure the current, we have to use a specific amplifier:



→ with $R_{IN} << R_P$ → almost no voltage drop

$V_{IN} \cong \phi$ → no current flowing in R_{IN}

↓

in this way we can measure almost the entire I provided by the current source (100 nA)

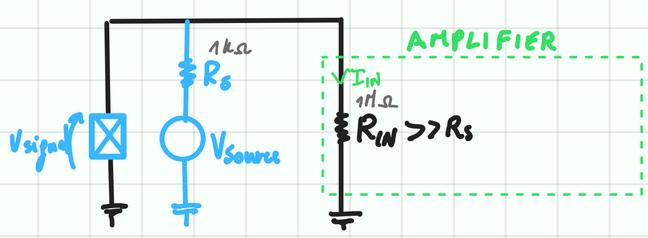
indeed in the ideal case

$\beta = tg^{-1}(R_P) = \frac{\pi}{2} \rightarrow R_P \rightarrow \infty$ (high)

ex. Voltage source

Let's suppose now to have a microphone ⊠ that generate a voltage signal from the acoustic wave.

In this case it's best to measure a voltage and not a current through an amplifier of this type:

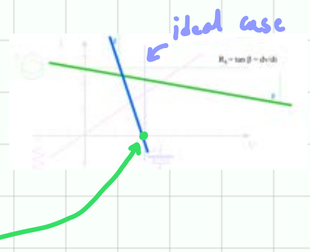


→ with $R_{IN} >> R_S$ → almost no current flowing into R_{IN} (or into the circuit)

$I_{IN} \cong 0$

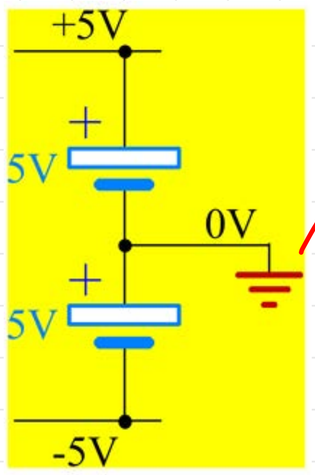
indeed in the ideal case

$\beta = tg^{-1}(R_S) = 0 \rightarrow R_S = 0$ (small)



Power supply

In almost all the circuit we'll see we'll have a power supply



Ground:

- local 0V reference
- usually it is the copper plate of Printed Circuit Board
- it can also be the chassis connection and possibly EARTH

Frequency Response

To precisely evaluate a frequency response, Laplace methods with complex impedance should be used.

HOWEVER, an **ASYMPTOTIC ANALYSIS** gives good results with easier computations.

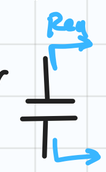
This asymptotic analysis evaluate the circuit response at:

- **DC** [$\emptyset$ Hz]
  - CAPACITORS: $z = \frac{1}{sC} = \infty \rightarrow$ OPEN CIRCUIT
  - INDUCTORS: $z = sL = 0 \rightarrow$ SHORT CIRCUIT
- **AC** [∞ Hz]
  - CAPACITORS: $z = \frac{1}{sC} = 0 \rightarrow$ SHORT CIRCUIT
  - INDUCTORS: $z = sL = \infty \rightarrow$ OPEN CIRCUIT

• Poles

To compute the poles of the circuits:

1. Switch OFF the generators ($v=0 \rightarrow$ short circuit / $i=0 \rightarrow$ open circuit)
2. compute the overall Resistance seen by the capacitor



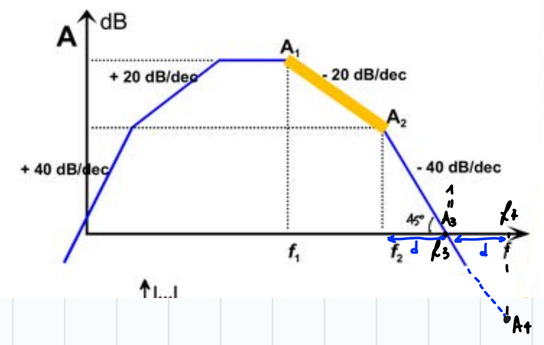
3. compute the pole as: $f_p = \frac{1}{2\pi \underbrace{R_{eq} C}_{\tau}}$
4. A pole on the Bode diagram causes a decrease of 20 dB/dec

-20dB/dec: $A_1 \cdot f_1 = A_2 \cdot f_2$ the so called constant Gain-Bandwidth Product GBWP

+20dB/dec: $\frac{f_1}{A_1} = \frac{f_2}{A_2}$

-40dB/dec: $\sqrt{A_1} \cdot f_1 = \sqrt{A_2} \cdot f_2$

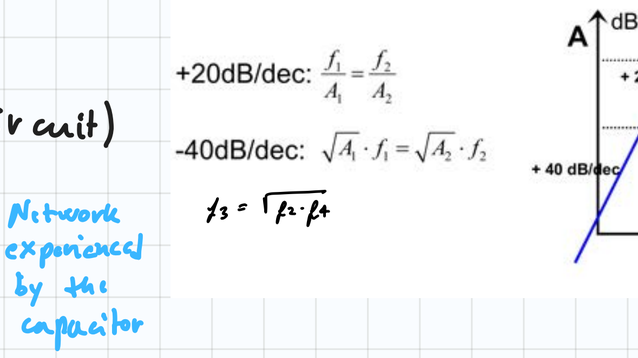
$f_3 = \sqrt{f_2 \cdot f_4}$



• Zeros

To compute the zeros of the circuits:

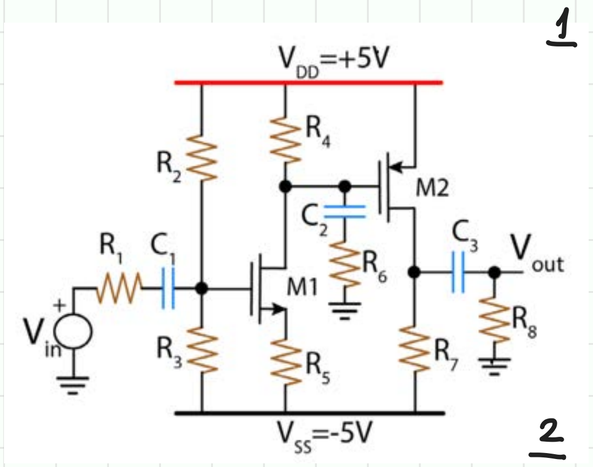
1. Switch OFF the generators ($v=0 \rightarrow$ short circuit / $i=0 \rightarrow$ open circuit) APART FROM THE INPUT
2. apply a generic (not null) signal at the input
3. apply a generic signal V(t) and i(t) on the C under test
4. Compute if output can be null (NIL)
5. Compute the R = resistors seen by C in these conditions
6. Compute the zero as: $f_z = \frac{1}{2\pi \underbrace{RC}_{\tau}}$



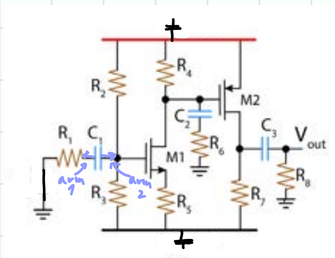
Note: All the previous arguments are still valid also for inductors instead of capacitor, just consider $\tau = \frac{R}{L}$

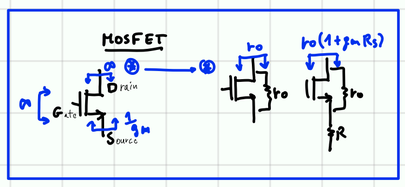
Ex.

Pole computation

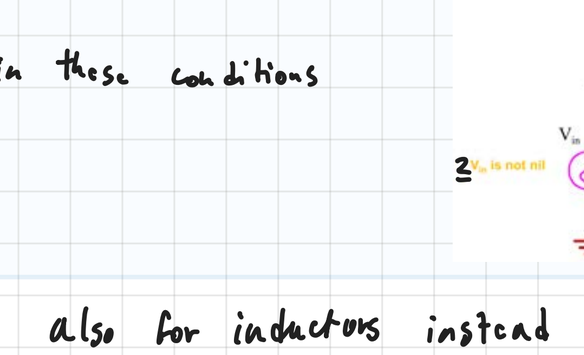


1. Switch OFF the generators





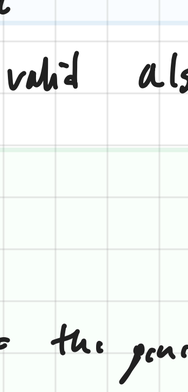
2. R_{eq} computation



the input impedance looking into a gate is ∞

$R^* \rightarrow \infty$

1. $R_{eq1} = \underbrace{R_1}_{\text{arm 1}} + \underbrace{R_2 \| R_3 \| R^*}_{\text{arm 2}}$



2. $R_{eq2} = R_4 + R_6$

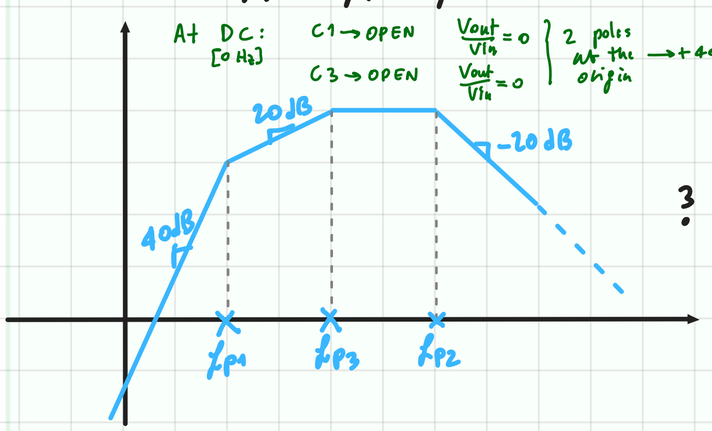
3. $R_{eq3} = R_7 + R_8$

3. $\tau_1 = C(R_1 + R_2 \| R_3 \| R^*)$

$\tau_2 = C(R_4 + R_6)$

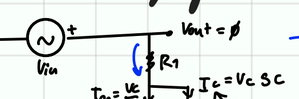
$\tau_3 = C(R_7 + R_8)$

4. for ex. $f_{p1} < f_{p3} < f_{p2}$



Zero computation → Common cases

1. Capacitor along the signal path ⇒ ZERO at the origin

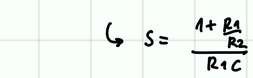


→ Ic must be ∅ :

$i_c(t) = C\frac{dv}{dt} \xrightarrow{\mathcal{L}} I_c(s) = sC V_c(s) = \emptyset$

$s=0 \rightarrow f_z = 0$

2. RC-shunt along the signal path ⇒ ZERO at finite freq.



→ current through R equal to Ic:

$I_c = sCV_c \quad \hookrightarrow sCV_c = -\frac{V_c}{R} \rightarrow s = -\frac{1}{RC}$

$\hookrightarrow f_z = \frac{1}{2\pi RC}$

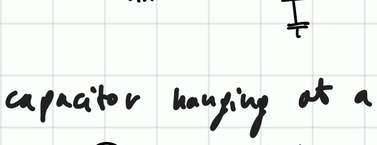
3. RC-series hanging at a node ⇒ a finite ZERO

→ Voltage on R equal to Vc:

$RsCV_c = -V_c \rightarrow s = -\frac{1}{RC} \rightarrow f_z = \frac{1}{2\pi RC}$

4. A capacitor hanging at a node ⇒ NO ZERO

5. A capacitor hanging at a node off the signal path ⇒ finite ZERO



→ the voltage on R_1 must be equal to V_c

$-\left(V_c sC + \frac{V_c}{R_2}\right)R_1 = V_c$

$-sCR_1 - \frac{R_1}{R_2} = 1$

$\hookrightarrow s = \frac{1 + \frac{R_1}{R_2}}{R_1 C} = \frac{R_1 + R_2}{R_1 R_2 C} = \frac{1}{(R_1 \| R_2)C} \rightarrow f_z = \frac{1}{2\pi (R_1 \| R_2) C}$

Amplifiers

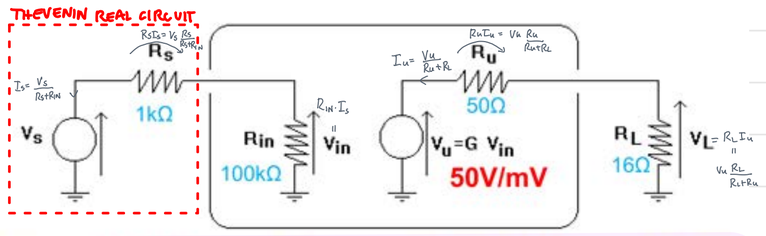
Intro from Zappa's Book p. 22

In analog electronics, it is important to extract signals from sensors and amplify them properly in order to make these signals strong enough to be treated by analog filtering and processing blocks or analog/digital conversion blocks placed downstream. Usually, the key feature of an amplifier is to have its **gain** G between its input and its output as constant and linear as possible. However, there are also other performance criteria that characterize an amplifier, and those criteria have to be considered in the analysis or design phase of an electronic system. For instance, the input **impedance** of an amplifier stage may cause a significant loss of amplitude of the signal to be amplified.

• Voltage amplifiers

Voltage to Voltage:

$$G = A_V = \frac{V_L}{V_S} \quad \text{V/V}$$



V_L/V_S = 0.99 x 50,000 x 0.24 = 12,000 V/V = 12,000

→ computations:

- $V_{IN} = V_S \frac{R_{in}}{R_{in}+R_S}$
- $V_u = G V_{IN} = G V_S \frac{R_{in}}{R_{in}+R_S}$ (voltage gain ≈ 50000)
- $V_L = V_u \frac{R_L}{R_L+R_u} = G V_S \frac{R_{in}}{R_{in}+R_S} \cdot \frac{R_L}{R_L+R_u}$

$\frac{V_L}{V_S} = \frac{R_{in}}{R_{in}+R_S} \cdot G \cdot \frac{R_L}{R_L+R_u}$ ≈ 0.99 · 50000 · ≈ 0.24

≈ 50000

→ GOOD VOLTAGE AMPLIFIER:

- $R_{in} >> R_S$ HIGH INPUT IMPEDANCE
- $R_u << R_L$ LOW OUTPUT IMPEDANCE

Unfortunately there are **voltage drops**:

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

↳ Voltage drops

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

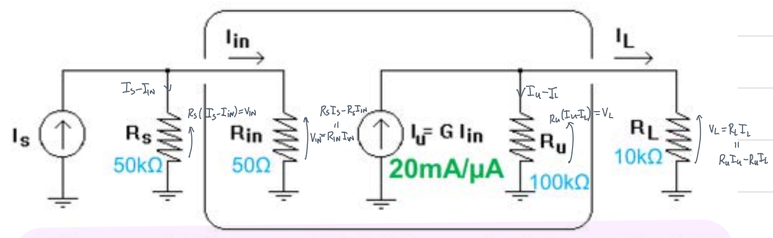
$R_{in} >> R_S$ very high input impedance $\qquad R_u << R_L$ very low output impedance

to improve in this case we have to choose $R_u << R_L$

• Current amplifier

Current to Current:

$$G = A_I = \frac{I_L}{I_S} \quad \text{A/A}$$



I_L/I_S = 0.994 x 20,000 x 0.91 = 18,000

→ computations:

- $I_{in} = I_S \frac{R_S}{R_S+R_{in}}$
- $I_u = G I_{in} = G I_S \frac{R_S}{R_S+R_{in}}$ (current gain ≈ 20000)
- $I_L = I_u \frac{R_u}{R_u+R_L} = G I_S \frac{R_S}{R_S+R_{in}} \cdot \frac{R_u}{R_u+R_L}$

$\frac{I_L}{I_S} = \frac{R_S}{R_S+R_{in}} \cdot G \cdot \frac{R_u}{R_u+R_L}$ ≈ 0.999 · 20000 · ≈ 0.91

≈ 20000

→ GOOD CURRENT AMPLIFIER:

- $R_{IN} << R_S$ LOW INPUT IMPEDANCE
- $R_u >> R_L$ HIGH OUTPUT IMPEDANCE

Unfortunately there are **current shunts**:

$$I_{in} = I_S \cdot \frac{R_S}{R_S+R_{in}} \qquad I_L = I_u \cdot \frac{R_u}{R_u+R_L}$$

↳ current shunts

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

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

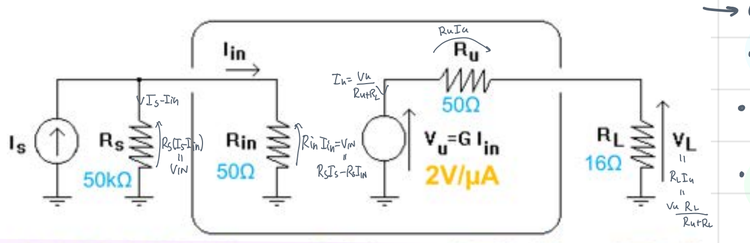
to improve it

• Transimpedance amplifier

I → V

Current to Voltage:

$$G = A_R = \frac{V_L}{I_S} \quad \text{V/A=}\Omega$$



V_L/I_S = 0.994 x 2,000,000 V/A x 0.24 = 477000 V/A = 477 kΩ

→ computations:

- $I_{in} = I_S \frac{R_S}{R_S+R_{in}}$
- $V_u = G I_{in} = G I_S \frac{R_S}{R_S+R_{in}}$
- $V_L = V_u \frac{R_L}{R_L+R_u} = G I_S \frac{R_S}{R_S+R_{in}} \cdot \frac{R_L}{R_L+R_u}$

$\frac{V_L}{I_S} = \frac{R_S}{R_S+R_{in}} \cdot G \cdot \frac{R_L}{R_L+R_u}$ ≈ 0.999 · 2·10⁶ [V/A] · 0.24

≈ 2 MΩ

→ GOOD TRANSIMPEDANCE AMPLIFIER:

- $R_{IN} << R_S$ LOW INPUT IMPEDANCE
- $R_u << R_L$ LOW OUTPUT IMPEDANCE

Unfortunately there are **losses**:

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

↳ Losses

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

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

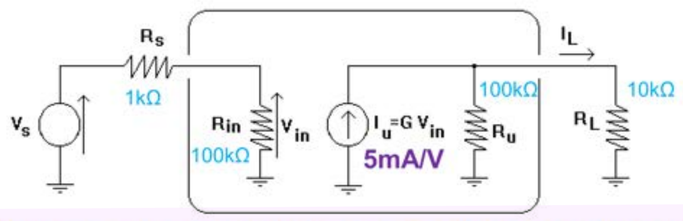
to improve it

• Transconductance amplifier

V → I

Voltage to Current:

$$G = A_G = \frac{I_L}{V_S} \quad \text{A/V=S}$$



I_L/V_S = 0.99 x 5 mA/V x 0.91 = 4.5 mA/V = 0.0045 Ω^{-1}

→ Computations

- $V_{IN} = V_S \frac{R_{IN}}{R_{IN}+R_S}$
- $I_u = G V_{IN} = G V_S \frac{R_{IN}}{R_{IN}+R_S}$
- $I_L = I_u \frac{R_u}{R_u+R_L} = G V_S \frac{R_{IN}}{R_{IN}+R_S} \cdot \frac{R_u}{R_u+R_L}$

$\frac{I_L}{V_S} = \frac{R_{IN}}{R_{IN}+R_S} \cdot \frac{G}{5\cdot10^{-3}\ [\frac{A}{V}]} \cdot \frac{R_u}{R_u+R_L}$ ≈ 0.99 · · 0.91

≈ 0.005 Ω^{-1}

→ GOOD TRANSCONDUCTANCE AMPLIFIER

- $R_{IN} >> R_S$ HIGH INPUT IMPEDANCE
- $R_u >> R_L$ HIGH OUTPUT IMPEDANCE

Unfortunately there are **losses**:

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

↳ Losses

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

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

to improve it

Obs.

The stages that must drive some actuators (e.g. speakers) deserve a separate discussion. In fact, their action is proportional to neither the voltage delivered to their terminals nor the current flowing in them, but the total electrical power delivered to them. In this case, the maximum power transfer efficiency is obtained when the output impedance of the amplifier stage is exactly equal to the impedance R_L.

Note: so far we have used the terms resistance and impedance without caring about the fact that impedance generally includes reactive components and that it changes as the frequency of the input signal changes.

Feedback

Intro: Book p. 28

1.3.1 Invention of the Feedback

The idea of **negative feedback** belongs to the American Harold S. Black. He came up with that idea in a Tuesday morning, on August 2, 1927, while crossing the Hudson River on the Lackawanna ferry to get to work in Manhattan. He was 29 years old and had been working as an engineer at the laboratories of the American Telephone Company (today's Bell Telephone Laboratories) for six years. The object of his research was to design a system for long-distance telephone communication which requires the creation of equipment that would enable an efficient linkage between the two coasts of the United States and between the United States and Europe. The difficulties of that research were related to not only the quality of the components used, but mainly the fact that no one knew how to design amplifiers, which are sufficiently stable, which are linear, and which do not lead to excessive distortion of signals applied to them. In fact, non-linearity of the elements that made up the amplifiers, in the first place, the electronic tubes, resulted in the generation of unwanted harmonics in the output signal while the variations of the characteristics of the same elements, due to temperature or aging effects, determined a change in the performance of the amplifiers, in particular in their gain.

The research objective of H. S. Black was the improvement of the performance of the amplifiers placed as repeaters along the telephone lines so that it could be possible to simultaneously transmit multiple channels on the same line for a long distance. He soon realized that the characteristics required for an amplifier to accomplish that performance were so strict that it could not be considered to obtain them simply by making improvements to existing topologies. They needed a completely new idea.

After H. S. Black came up with the idea, he sketched, on a page of The New York Times, the diagram of a negative feedback loop and derived its basic properties. He signed his notes at the bottom of the newspaper, then arrived at the laboratory, and showed them to his manager, E. C. Blessing, shortly thereafter. The manager, convinced of the importance of that invention, also signed as a witness in footnotes. Those notes summarized the idea that the amplifier gain could be well controlled and that the distortion of the amplified signal could be extremely lowered if the output signal of the circuit were brought back to the input and added in phase with the applied signal.

NEGATIVE FEEDBACK

Four days later, on August 6, he made the effects of feedback on input and output impedance of a circuit clear, thereby achieving another important objective: to establish and stabilize the impedance of various stages of an amplifier, this impedance had to match that of the signal transmission cables perfectly.

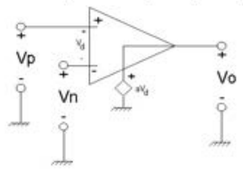
On December 29 of the same year, H. S. Black experimentally verified, for the first time, the characteristics of negative feedback systems through measuring a distortion reduction by a factor of 100,000 on input signals between 4 and 45kHz, using the first negative feedback amplifier in the history.

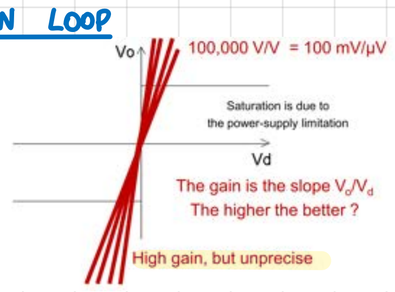
Practical idea: We don't want to buy an OpAmp for every possible gain, we can use the negative feedback of the output brought back to the OpAmp's input to adjust, through a controller, the GAIN

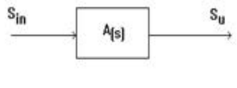
OpAmp differences:

OPEN LOOP

Here is the "open-loop" OpAmp:

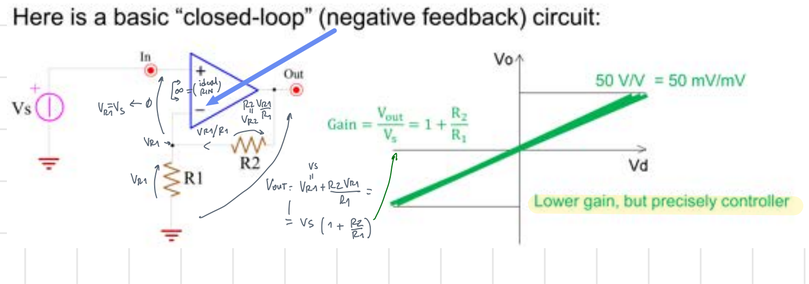


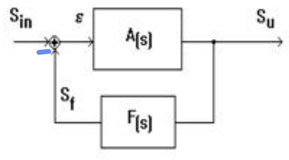




CLOSED LOOP

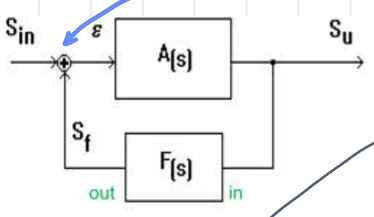
Here is a basic "closed-loop" (negative feedback) circuit:





Gain computation:

Closed-loop OpAmp with feedback:



with ⊕ ⊛

- $S_u = A(s) \cdot \varepsilon$
- $\varepsilon = S_{in} + S_f \rightarrow S_{in} = \varepsilon - S_f = \varepsilon - \varepsilon A(s) F(s)$
- $S_f = S_u \cdot F(s) = \varepsilon \cdot A(s) \cdot F(s)$
- $S_{in} = \varepsilon (1 - A(s)F(s))$

Real (closed-loop) gain: $G(s) = \frac{S_u}{S_{in}} = \frac{A(s)}{1 - A(s) \cdot F(s)} = \frac{A(s)}{1 - G_{loop}} = \frac{A(s)}{G_{loop}} \cdot \frac{G_{loop}}{1 - G_{loop}} = \frac{1}{F(s)} \cdot \frac{G_{loop}}{G_{loop}\left(\frac{1}{G_{loop}} - 1\right)} = -\frac{1}{F(s)} \cdot \frac{1}{1 - \frac{1}{G_{loop}}} \rightarrow -\frac{1}{F(s)}$

Real gain · G_{loop} · Correction factor · Ideal gain

with ⊕ here − (with ⊖ here +)

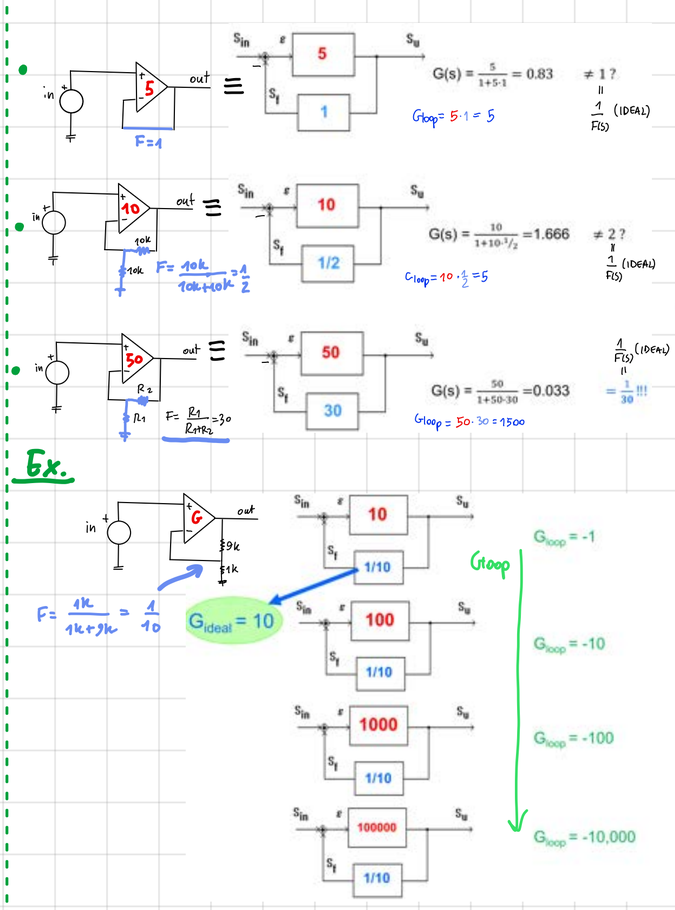
Loop gain G_{loop}: measures the **feedback effectiveness**, the higher the better !

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!

Ex.



↳ No practical use if |Gloop| is NOT >>1 !

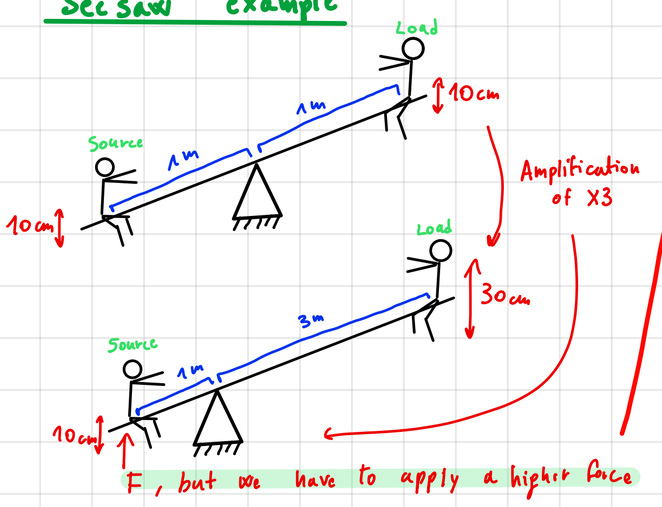
Ex.

↳ If |Gloop|>>1, gain is set just by the feedback, whatever the OpAmp is!

Virtual ground:

- **Idea:** If we have a high load wrt a small source we have to put an amplifier stage between the source and the load

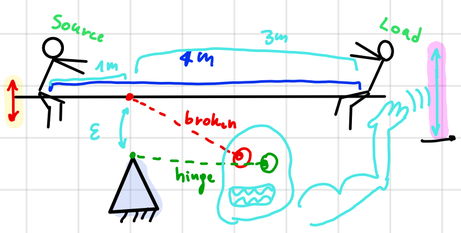
"Seesaw" example



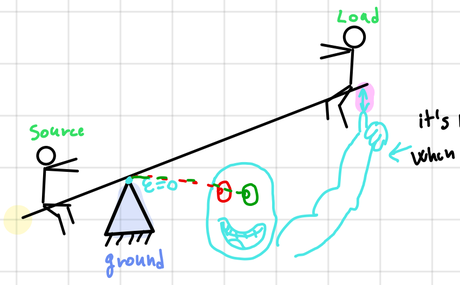
Suppose now to use a "broken" seesaw (without the hinge to the ground)

→ If we're above the "load" end doesn't move

→ Now, suppose there's a strong monster (amplifier) that knows the seesaw it's broken and where the hinge should be

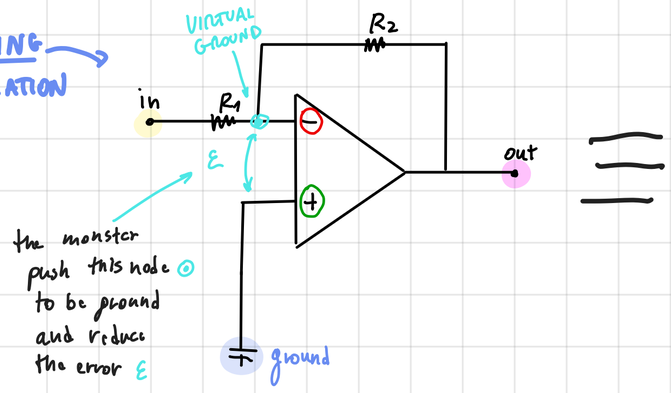


The monster sees that it's broken and there should be an hinge to the ground in that point, so it sees there's an error ε. Therefore the monster start to move itself the other end of the seesaw making big movements to reduce the error.



it's happy when ε=0

INVERTING CONFIGURATION



the monster push this node ⊙ to be ground and reduce the error ε

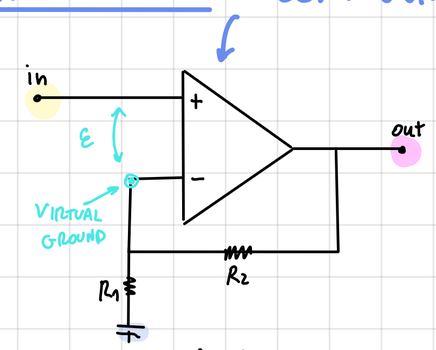
↳ if ε>0 → the monster raises the output ↑

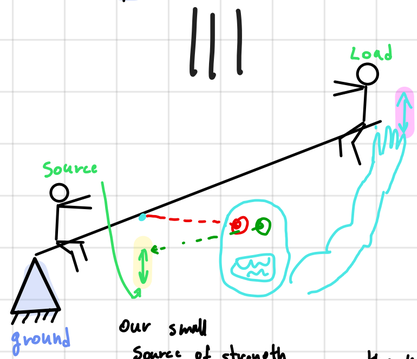
⇓

Since we have a negative feedback the point where the hinge is broken will go down ↓ till it touches the hinge (ground) → ε=0

⇐ this means that however small the input source, the monster (amplifier) will be able to raise the output to satisfy the feedback error to be 0

NON-INVERTING CONFIGURATION





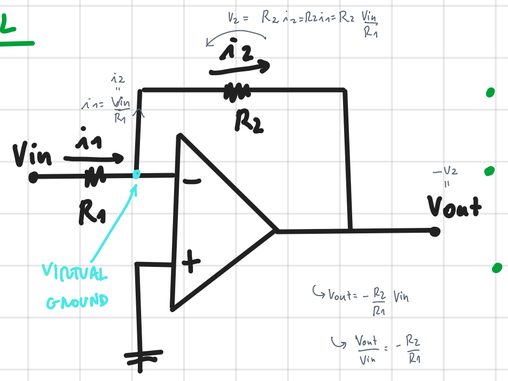
Our small source of strength tries to apply a force here (but can't) → the monster sees that the input source want to move and sees where the 1m/3m (R1/R2) partition should be.

then the monster will move ↕ the output in order to have the desired input movement

Configurations:

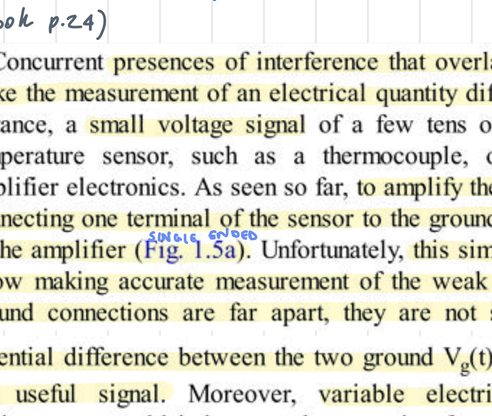
INVERTING CONFIGURATION

IDEAL



- Voltage GAIN: $A_v = -\frac{R_2}{R_1}$
- Input Impedance: $R_{in} = R_1$
- Output Impedance: $R_{out} \simeq 0$

REAL



- Voltage GAIN: $A_v = \frac{A_{v\,ideal}}{1 - G_{loop}} = -\frac{R_2}{R_1}\frac{1}{1 - G_{loop}}$
- Input Impedance: $R_{in} = R_1 + R_A \simeq R_1$ (*)
- Output Impedance: $R_{out} = \frac{R_{out} \,||\, (R_2 + R_1 || R_{in})}{1 - G_{loop}} \simeq 0$ (*)

(*) A won't actually be the ground (virtual ground)
↳ R_A is the equiv. resistance seen at node A

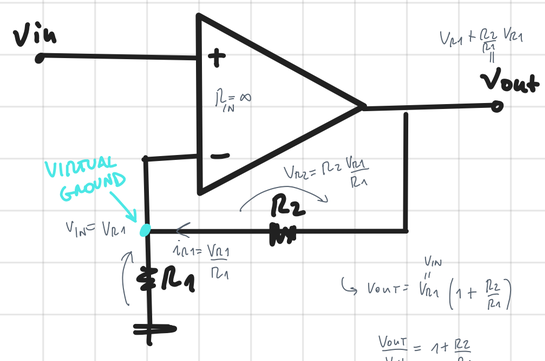
↳ $R_A = \frac{\mathcal{E}}{i_1} = \frac{\mathcal{E}}{i_{IN} + i_2} = \frac{\mathcal{E}}{\frac{\mathcal{E}}{R_{IN}} + \frac{\mathcal{E} - A(s)\mathcal{E}}{R_2}} = \frac{R_{IN} \cdot R_2}{R_2 + R_{IN} - A(s) R_{IN}} = \frac{R_{IN} \cdot R_2}{R_{IN} + R_2} \cdot \frac{1}{1 - A(s)\frac{R_{IN}}{R_{IN} + R_2}} = \frac{R_{IN} || R_2}{1 - A(s)\frac{R_{IN}}{R_{IN} + R_2}}$

($\mathcal{E}$: voltage; i_1: current; $\frac{V_{R2}}{R_2}$; correcting factor; G_{loop})

$\Rightarrow R_A = \frac{R_{IN} || R_2}{1 - G_{loop}} \simeq 0$

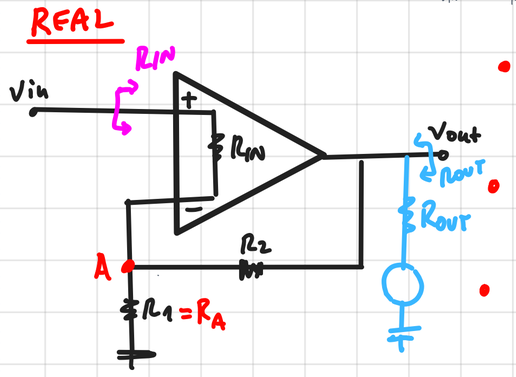
NON-INVERTING CONFIGURATION

IDEAL



- Voltage GAIN: $A_v = 1 + \frac{R_2}{R_1}$
- Input Impedance: $R_{IN} \simeq \infty$
- Output Impedance: $R_{OUT} \simeq 0$

REAL



- Voltage GAIN: $A_v = \frac{A_{v\,ideal}}{1 - G_{loop}} = \left(1 + \frac{R_2}{R_1}\right) \cdot \frac{1}{1 - G_{loop}}$
- Input Impedance: $R_{IN} = (R_{IN} + R_1 || R_2)(1 - G_{loop}) \to \infty$ (*)
- Output Impedance: $R_{OUT} = \frac{R_{OUT} \,||\, (R_2 + R_1 || R_{IN})}{1 - G_{loop}} \simeq 0$ (*)

Differential Amplifiers:

(Book p.24)

Concurrent presences of interference that overlaps the useful signal often make the measurement of an electrical quantity difficult. Let us consider, for instance, a small voltage signal of a few tens of μV, developed across a temperature sensor, such as a thermocouple, occurring away from the amplifier electronics. As seen so far, to amplify the signal, you might think of connecting one terminal of the sensor to the ground and the other to the input of the amplifier (Fig. 1.5a). Unfortunately, this simple configuration does not allow making accurate measurement of the weak signal. In fact, if the two ground connections are far apart, they are not strictly equipotential. The potential difference between the two ground $V_g(t)$ would hence be added to the useful signal. Moreover, variable electric fields present in the environment would induce an electromotive force in series with the sensor signal, which is proportional to the area of the loop linked with the electromagnetic field.

To ensure the amplification of the useful signal alone, connecting both terminals of the sensor to the amplifier with a matched pair of wires can be considered (Fig. 1.5b). In this new configuration, the potential difference between the terminal of the amplifier and the ground, named the differential signal, is equal to the sum of the useful signal $V_d(t)$ and the only interference induced in the loop (identified by the two connecting wires) by the electromagnetic fields. Since the area of this loop is much less extensive than the previous one, the electromagnetic interference is proportionally reduced while the disadvantage of the non equipotentiality of the two ground connections is entirely eradicated.

Nevertheless, it can be useful, and sometimes essential, that the sensor is separated from the ground. In these cases, it is necessary to use a voltage amplifier with two input terminals that is able to amplify only the differential signal present at its terminals (V'-V''). Amplifiers designed specifically for these applications are called differential amplifiers (Fig. 1.5c). However, note that a potential difference between the sensor and the amplifier ground $V_{cm}(t)$ still remains. This signal is called the common mode signal and equal to the average of the potentials of the two wires. If the differential amplifier were ideal, this common mode signal would have no effect on the output value, and the signal $V_o(t)$ would simply be proportional to the differential signal.

As shown in Fig. 1.5c, twisting the two wires that connect the sensor and the amplifier allows reducing even the small differential interference still present due to the loop between the two wires. In this way, on each conductor, the directions of the induced electromotive forces alternate from a lobe to the next, thus canceling out.

Like all voltage amplifiers, the differential amplifier must show, between the two inputs V' and V'', high impedance, which is ideally infinite, while the output has to have low impedance, which is ideally zero. The circuit is characterized by a differential gain G_d that defines its ability to amplify the differential signal $V_d=V'-V''$ and by a common mode gain G_{cm} that accounts for the residual amplification, which is undesired, of the common mode signal $V_{cm}=(V'+V'')/2$.

The simultaneous presence of the two types of signal at the input means, by and large, that the transfer of a real amplifier is given by:

$$V_o = G_d \cdot V_d + G_{cm} \cdot V_{cm}$$

Obviously, a good differential amplifier will have a high differential gain G_d and a low common mode transfer G_{cm}. For instance, if we assume that $V_d=50\mu V$ while $V_{cm}=500mV$, to prevent the common mode from completely hiding the amplified differential signal at the output, the ratio G_d/G_{cm} must be much greater than $(500mV/50\mu V)=10^4$.

The Common Mode Rejection Ratio (CMRR) is an important figure of merit of a differential amplifier and defined as $CMRR=G_d/G_{cm}$. Usually, the two amplifications differ by several orders of magnitude, so the CMRR, expressed in dB, is between 80 and 100dB.

For specific applications, differential amplifiers can be designed and implemented with discrete components, i.e. by assembling individual components on a printed circuit board. Nonetheless, nowadays, in most cases, integrated differential amplifiers are used, in which the whole circuit is built on the same silicon substrate. Operational amplifiers are integrated differential amplifiers characterized by differential amplifications of $10^5 \div 10^6$ and a CMRR of 100÷120dB with input resistance up to some GΩ and output resistance lower than 100Ω. These components are among the most widely used in the production of electronic circuits and are called operational because, if properly connected, they allow performing many operations on the input variables (sums, differences, derivations, integrations, etc.).

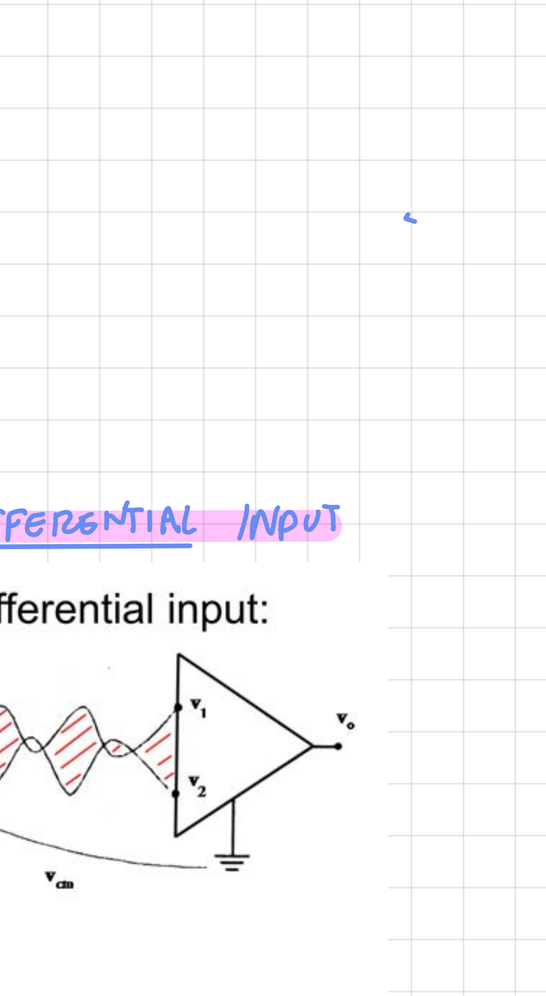
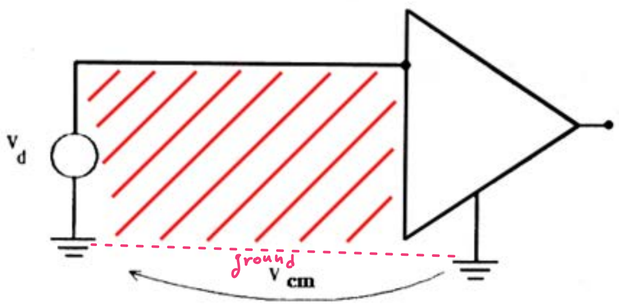


Fig. 1.5: Connection schemes of a sensor to an amplifier: a) single-ended through ground; b) single ended through two wires; c) differential sensing.

↳ Recap.:

SINGLE-ENDED INPUT

Single-ended input:



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

DIFFERENTIAL INPUT

Differential input:

$$V_o = A_d \cdot V_d + A_{cm} \cdot V_{cm}$$

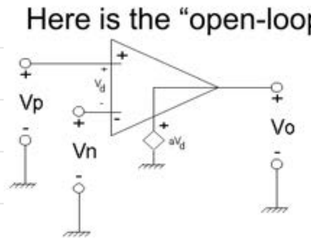
- let's amplify just Differential signal
- let's reject Common Mode disturbance

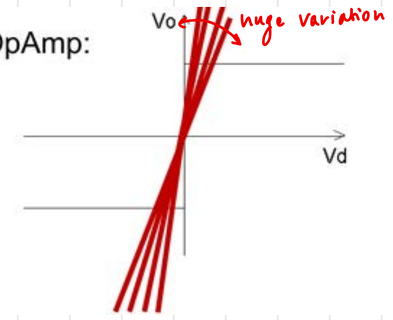
$$CMRR = \frac{A_d}{A_{cm}} \approx 80\text{–}100dB$$

Feedback effect on amplifier's mismatches and drifts:

OPEN-LOOP

Here is the "open-loop" OpAmp:





$$G = \frac{V_o}{V_d} = A$$

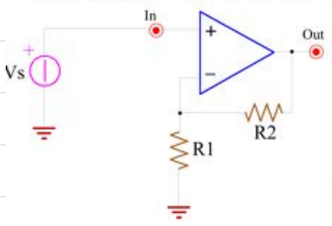
$$\frac{dG}{G} = \frac{dA}{A}$$

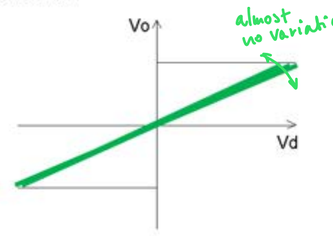
e.g.=±50%

→ if we buy a OpAmp and for ex. for aging it changes by 50% ($\frac{dA}{A}$) our gain will change by 50% ($\frac{dG}{G}$)

CLOSED-LOOP

Here is the "closed-loop" circuit:





$$G = \frac{V_o}{V_d} = \frac{A}{1 - A \cdot F} = \frac{A}{1 + G_{loop}}$$

$$\frac{dG}{G} = \frac{dA}{A} \cdot \frac{1}{1 - G_{loop}}$$ (correction factor; $\frac{dA}{A}$: 50% variation)

e.g.=±50% · $\frac{1}{1+100}$ = ± 5‰

But what about $\frac{dG}{G} = \frac{dF}{F}$ ...??

It happens that $\frac{dG}{G} = \frac{dF}{F}$ ← our feedback F (depends on R1, R2)

(Book p. 33)

It is important to note how negative feedback makes the transfer less sensitive to changes in the parameters of the forward block, but that it has no effect with respect to the variation in the transfer function F(s) of the feedback block. In fact, according to Eq. 1.3, it is just the transfer function F(s) that determines the transfer function of the system. Therefore, in order to have a transfer of the feedback amplifier that is reproducible, we must care about the reproducibility of the parameters of the transfer function F(s). In practice, this condition is easily satisfied with building the feedback branch only with passive components (usually resistors) that have a sufficient margin of tolerance. Thus, the amplification A(s) of the forward block has to only insure a high loop gain.

$\otimes$ $G(s) \cong -\frac{1}{F(s)}$

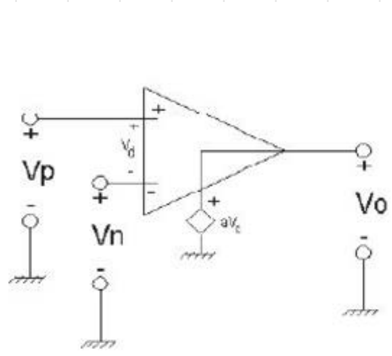
$dG = \frac{dF}{F^2}$

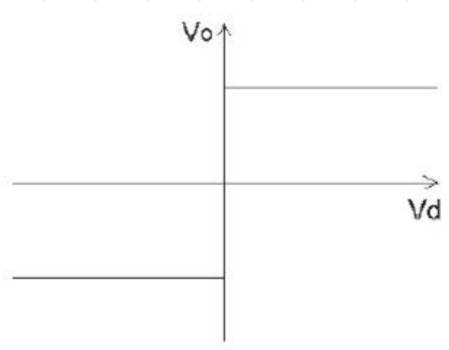
$G = -\frac{1}{F}$

$\frac{dG}{G} = \frac{dF}{F^2}(-F) = (-)\frac{dF}{F}$

Parameters of OpAmp:

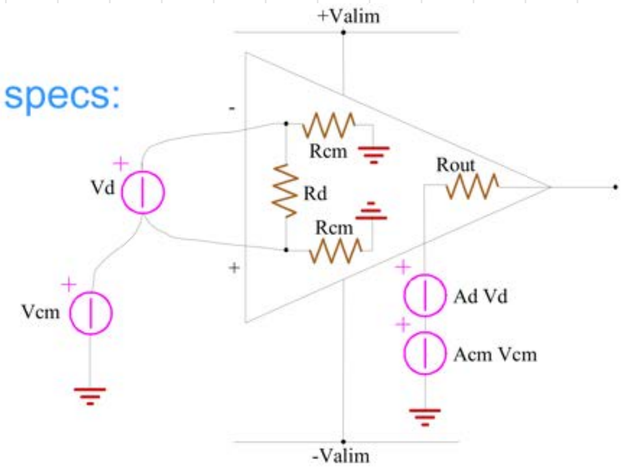
IDEAL OpAmp





- Infinite differential gain (A_d);
- Common mode gain (A_{cm}) equal to zero;
- Infinite bandwidth (BW);
- Infinite input differential impedance (Z_d);
- Infinite input common mode impedance (Z_{cm});
- Output impedance (Z_o) equal to zero;

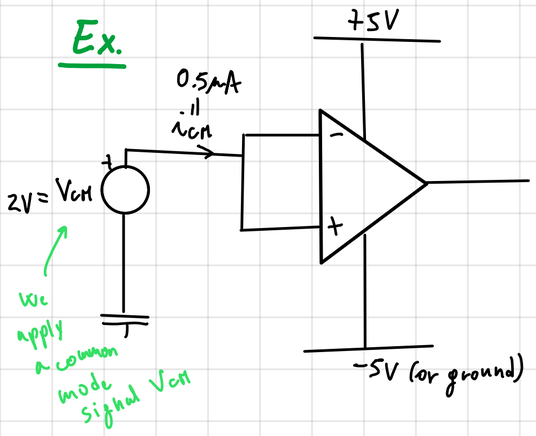
NON-IDEAL OpAmp



Non-ideal **small-signal** specs:

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	

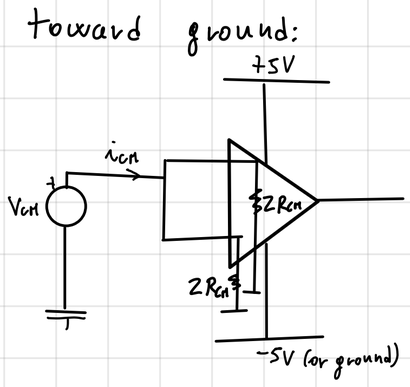
Ex.



→ If we measure a current it means there's an input impedance

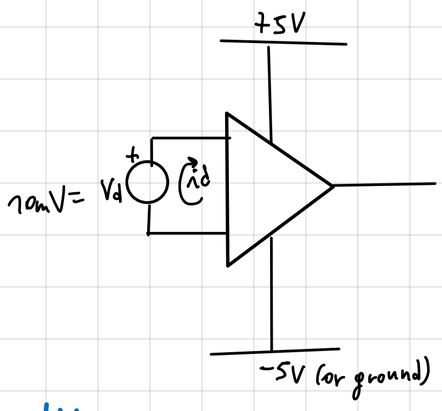
$$R_{IN} = \frac{2V}{0.5\mu A} = 4\,M\Omega$$

⟹ In the case of common mode we can model R_{IN} as 2 resistances toward ground:



with $R_{CM} = \frac{2V}{0.5\mu A} = 4\,M\Omega$

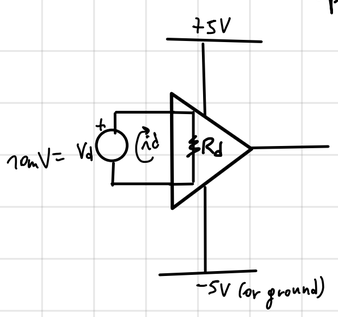
→ If instead we apply a differential voltage:



→ If we measure a differential current we find an input impedance:

$$R_{IN} = \frac{V_d}{i_d} = \frac{10mV}{1\mu A} = 10k\Omega$$

⟹ In the case of differential mode we can model R_{IN} as a differential impedance R_d:

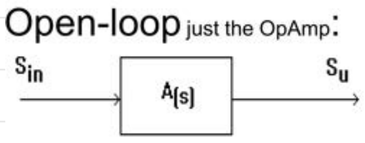


with $R_d = \frac{V_d}{i_d} = \frac{10mV}{1\mu A} = 10k\Omega$

Bandwidth

OPEN-LOOP

Open-loop just the OpAmp:



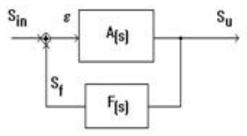
SLOPE: x/x 10/10 -20dB/dec -45°

$$A = \frac{G_0}{1 + s\tau}$$

CLOSED-LOOP

Closed-loop

OpAmp with feedback:

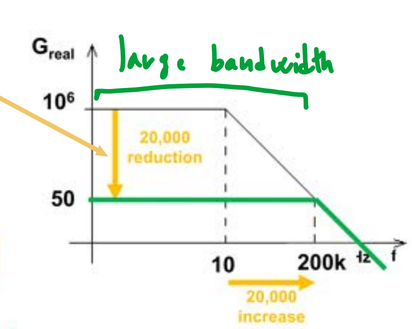


$$\frac{A}{1 + A\beta} = \frac{\frac{G_0}{1+s\tau}}{1 + \beta \cdot \frac{G_0}{1+s\tau}} = \frac{G_0}{(1 - Gloop) \cdot \left(1 + \frac{s\tau}{1 - Gloop}\right)}$$

$G_{loop} = |A(0)||F(0)| = A_0 F_0 = G_0 \beta$

Gain gets lower $\frac{G_0}{1+Gloop}$

Bandwidth gets wider polo=(1/τ)·(1+Gloop)

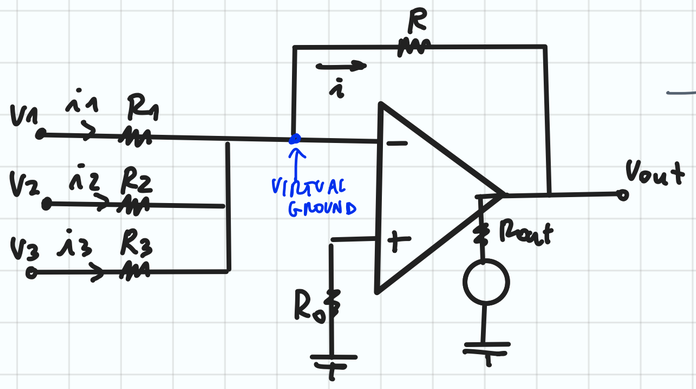


For example: F=1/50 G_{loop} = -20,000 $G_{real} \approx 50$ pole = 200,000 Hz !

We can generally say that the negative feedback, together with other benefits in terms of stabilization of the transfer and of impedentialmatching, also has a beneficial effect on the frequency response of the circuit since it tends to widen the bandwidth.

Voltage and current adder

Because of mutual ground, all currents sum up to - pin to create i, the output is the weighted sum of the input voltages. This type of scheme allows to avoid crosstalks between all voltage sources by creating a fake (virtual) ground. The circuit adds the voltages of the input terminals with a gain equal to R/R1 (possibly an attenuation if the ratio is lower than one), and the output gives the inverted voltage (due to the sign "-").



→ Computations:

$i_1+i_2+i_3 = i$

$\frac{V_1}{R_1}+\frac{V_2}{R_2}+\frac{V_3}{R_3}=-\frac{V_0}{R}$

- Voltage GAIN: $V_{out} = -\left(\frac{R}{R_1}V_1+\frac{R}{R_2}V_2+\dots+\frac{R}{R_n}V_n\right) = -R\sum_{i=0}^{n}\frac{V_i}{R_i}$

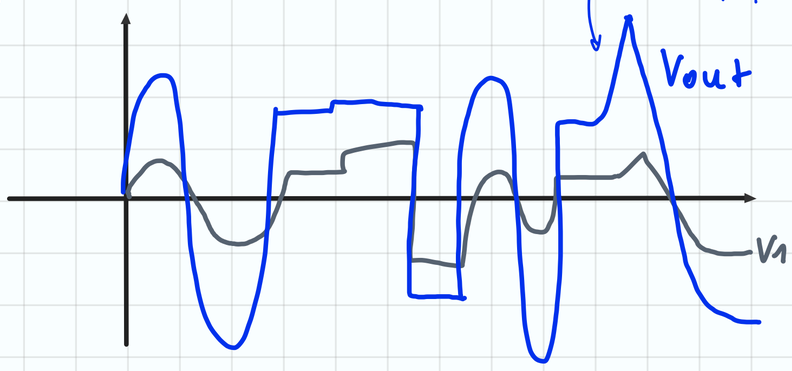
  ↓ if $R=R_i$

  $V_{out} = -(V_1+V_2+\dots+V_n)$

  ADDER

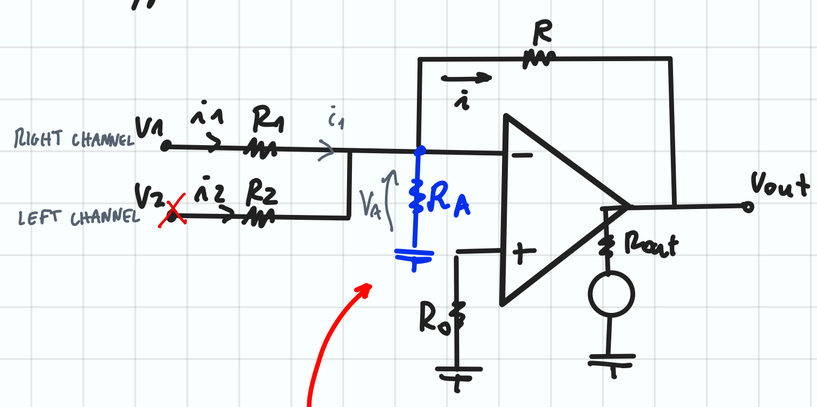
- Input Impedance: $Z_{in_i} = R_i$ (for every channel)
- Output Impedance: $Z_{out} \simeq 0$

Voltage characteristic:



Ex. • AUDIO MIXER with resistor instead of virtual ground

Suppose we have a resistor instead of the virtual ground



→ Imagine V_2 is not operating ✗ while V_1 is working

↓

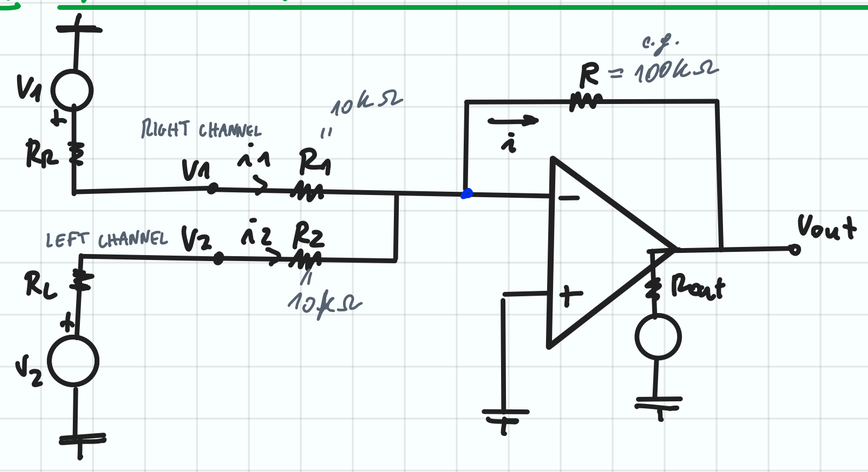
i_1 creates a voltage on R_A and this voltage is on the left microphone (V_2)

↓

PROBLEM ← CROSS-TALK: one microphone talks (V_1) and the other one (V_2) listen to V_1

We have to put the ground

Ex. • AUDIO MIXER with different channel internal resistors on real source



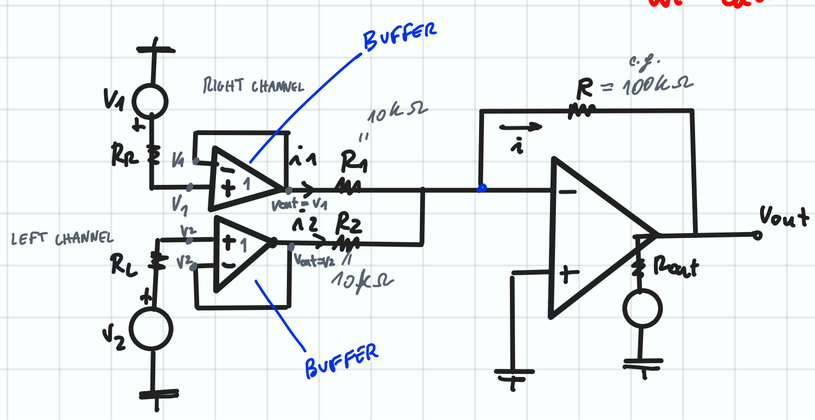
→ If we have $R_1=R_2=10k\Omega$ → to have $G_L = G_R = -\frac{R}{R_i} = -10$

but $R_L \neq R_R$ it means

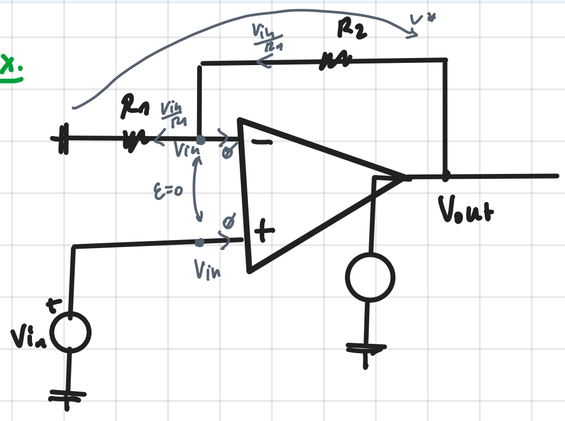
that $G_L \neq G_R$ → PROBLEM

↓ We can use a BUFFER configuration

↙ ↓ to keep the stages separed



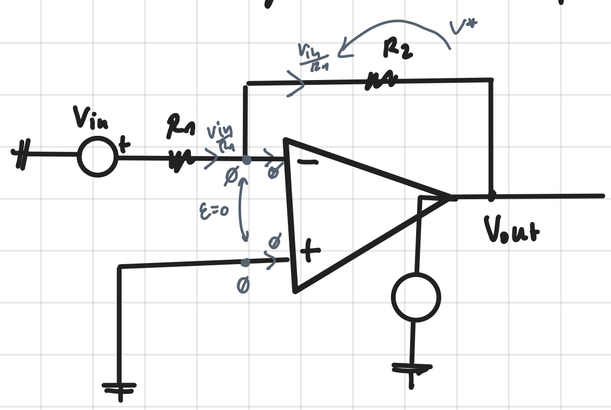
Ex.



→ $V_{out} = V^* = \frac{V_{in}}{R_1}(R_1+R_2)$

↳ $\frac{V_{out}}{V_{in}} = 1+\frac{R_2}{R_1} \geq 1$ ← AMPLIFICATION (NEVER ATTENUATED)

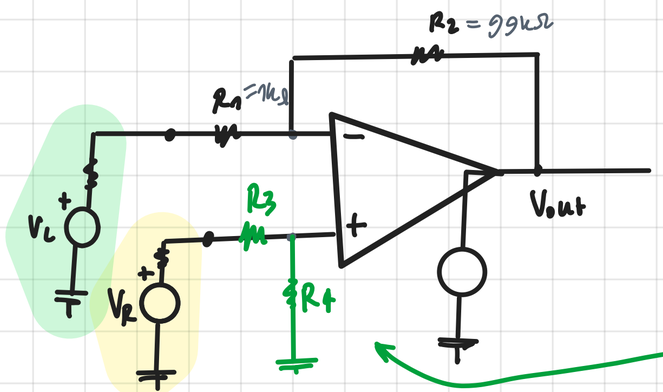
- If we change a bit the previous circuit



$V_{out} = -V^* = -\frac{V_{in}}{R_1}R_2$

↳ $\frac{V_{out}}{V_{in}} = -\frac{R_2}{R_1}$ (negative) $\gtrless |1|$

- Let's suppose to implement an amplifier with both the previous stages



→ $G_L = -\frac{R_2}{R_1} = -99$

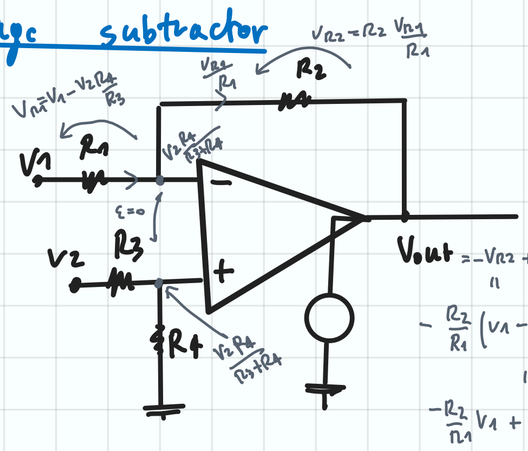
→ $G_R = 1+\frac{R_2}{R_1} = +100$

≠ → Problem → we have to introduce an attenuation

↳
- $G_L = -\frac{R_2}{R_1} = -99$
- $G_R = \frac{R_4}{R_3+R_4}\cdot\left(1+\frac{R_2}{R_1}\right) = 99$

R3, R4 chosen in order to have the same gain

Voltage subtractor



$$\rightarrow V_{out_1}=-\frac{R_2}{R_1}V_1$$

$$V_{out_2}=\frac{R_4}{R_3+R_4}\left(1+\frac{R_2}{R_1}\right)V_2$$

$$V_{out}=-V_{R2}+V_2\frac{R_4}{R_3+R_4}$$

$$-\frac{R_2}{R_1}\left(V_1-V_2\frac{R_4}{R_3+R_4}\right)+V_2\frac{R_4}{R_3+R_4}$$

$$-\frac{R_2}{R_1}V_1+V_2\frac{R_4}{R_3+R_4}\left(1+\frac{R_2}{R_1}\right)=V_{out1}+V_{out2}$$

$$-\frac{R_2}{R_1}V_1+V_2\left(\frac{R_4}{R_3+R_4}\right)\left(\frac{R_3+R_4}{R_3}\right)=-\frac{R_2}{R_1}V_1+\frac{R_4}{R_3}V_2$$

- Differential GAIN: $\left[\frac{R_2}{R_1}=\frac{R_4}{R_3}\right]$

$$V_{out}=-\frac{R_2}{R_1}V_1+\frac{R_4}{R_3}V_2=-\frac{R_2}{R_1}(V_1-V_2)$$

↓ if $R_1=R_2=R$

$V_{out}=V_2-V_1$ SUBTRACTOR

- common mode GAIN: $V_{out}=2\cdot\frac{R_2}{R_1}V_{control}$

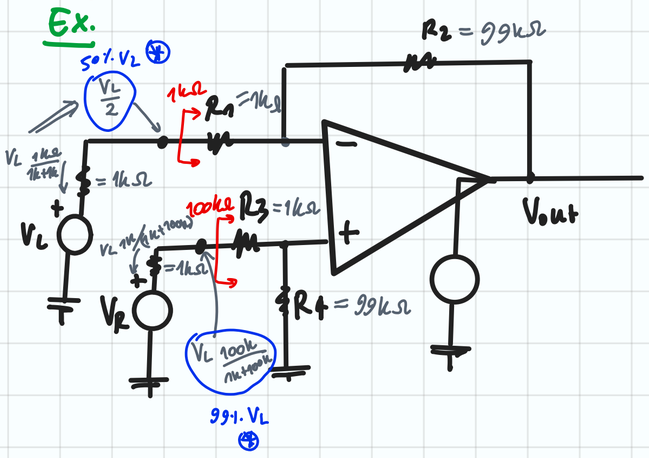
- Input Impedance: $R_{pos}=R_3+R_4$

PROBLEM

$R_{neg}=R_1$

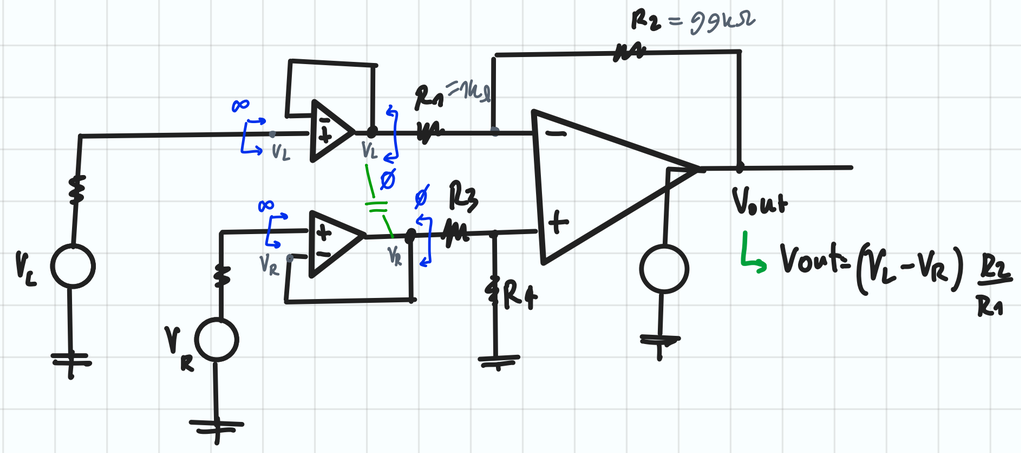
- Output Impedance: $R_{out}\approx 0$

Ex.



→ If the $V_L=V_R$ will have a voltage mismatch at the input

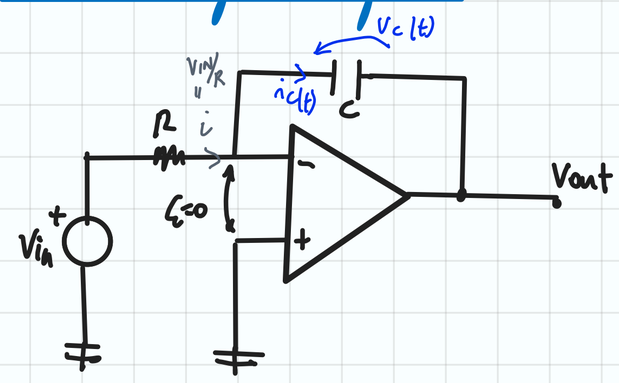
↳ we can solve it with this configuration →



↳ $V_{out}=(V_L-V_R)\frac{R_2}{R_1}$

FILTERS

IDEAL voltage integrator



- capacitor $\quad i(t)=i_C(t)=C\frac{dV_C(t)}{dt}$

$i(t)=\frac{V_{IN}}{R}$

↳ $V_{out}=-V_C(t)=-\frac{1}{C}\int i(t)dt=-\frac{1}{RC}\int V_{IN}(t)dt$ TIME DOMAIN

↓ LAPLACE TRANSFORM

$V_{out}=-\frac{1}{RCs}V_{IN}=-\frac{1}{s\tau}V_{IN}$ FREQUENCY DOMAIN

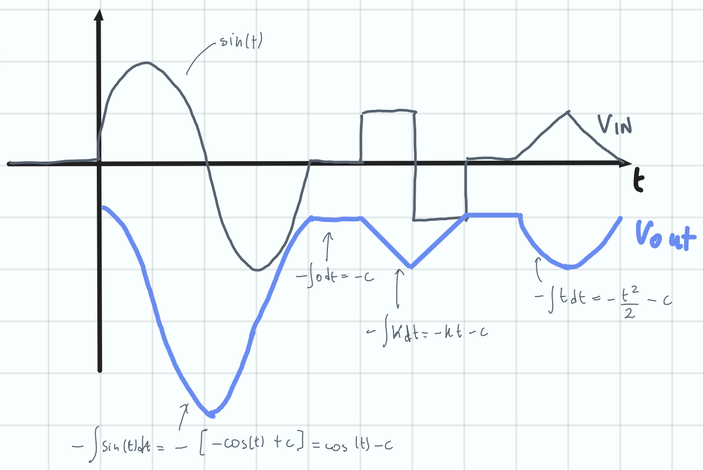
- Characteristic in time

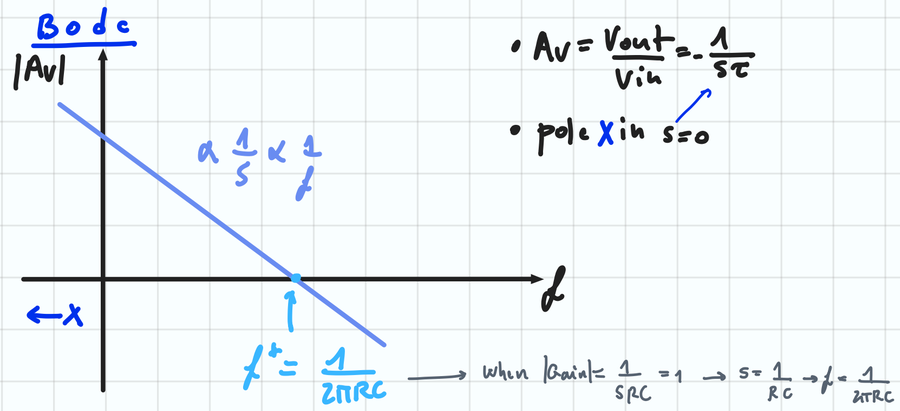
Issues

- DC gain $\rightarrow\infty$ (when $f\rightarrow 0$ ($-\infty$ in log-log))

⇓

OpAmp will eventually saturates

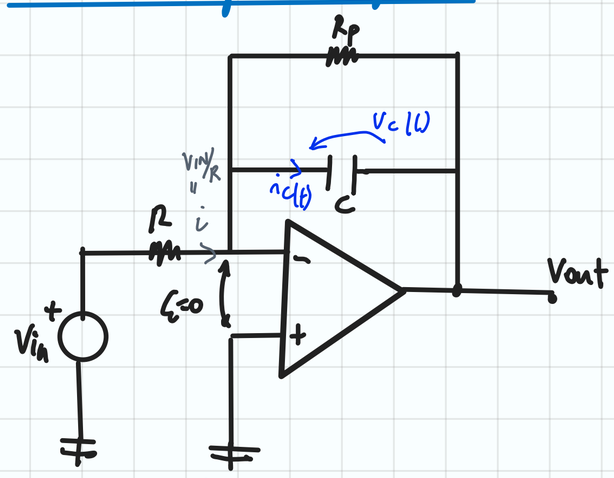




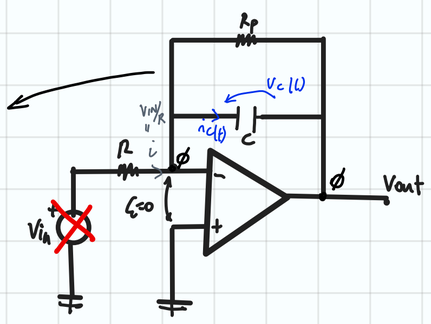
- $A_V=\frac{V_{out}}{V_{in}}=-\frac{1}{s\tau}$
- pole x in s=0

→ when $|Gain|=\frac{1}{sRC}=1\rightarrow s=\frac{1}{RC}\rightarrow f=\frac{1}{2\pi RC}$

REAL Voltage integrator

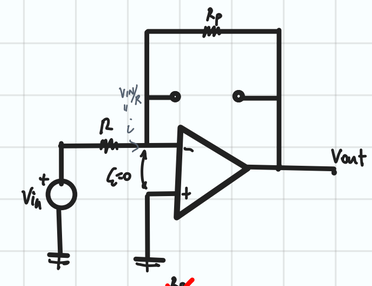


- $R_{eq}=R_p$



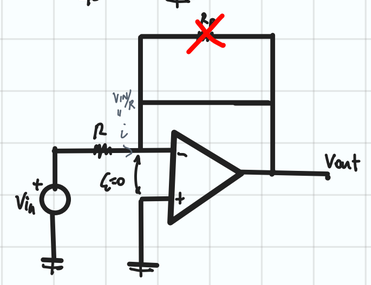
- pole at: $f_p=\frac{1}{2\pi R_{eq}C}=\frac{1}{2\pi R_pC}$

- at DC ($f\rightarrow 0$)

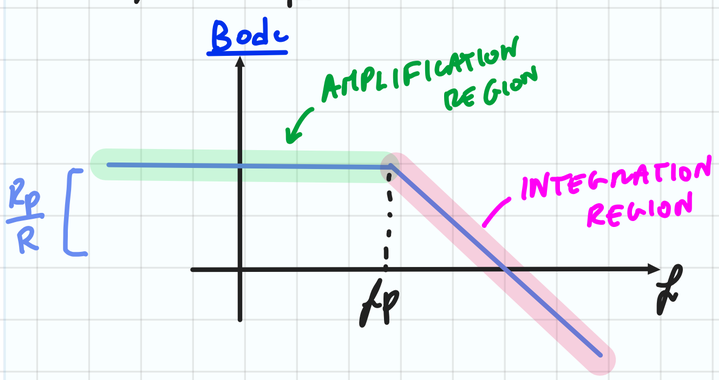


$A_V(0)=\frac{V_{out}}{V_{in}}=-\frac{R_p}{R}$

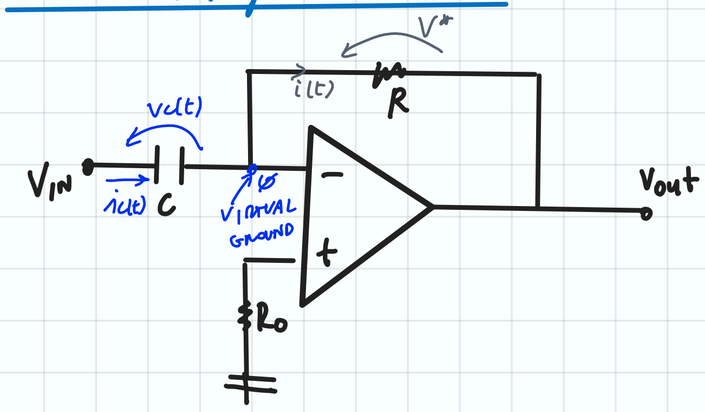
- at HF (High Frequency) ($f\rightarrow\infty$)



$A_V(\infty)=0$



IDEAL Voltage Derivator

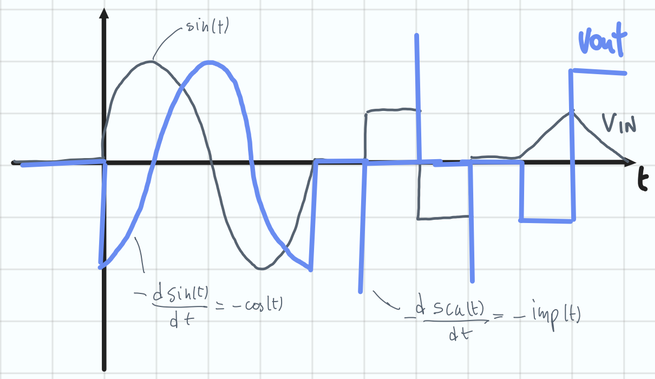


- **Capacitor** $i(t) = i_C(t) = C\frac{dV_C(t)}{dt} = C\frac{dV_{IN}(t)}{dt}$

$$V_{out} = -V^* = -R\,i(t) = -RC\frac{dV_{IN}(t)}{dt}$$ TIME DOMAIN

$$V_{out}(s) = -\frac{RC}{\tau}\,s\,V_{IN}(s) = -s\tau\,V_{IN}(s)$$ FREQUENCY DOMAIN

- **Characteristic in time**



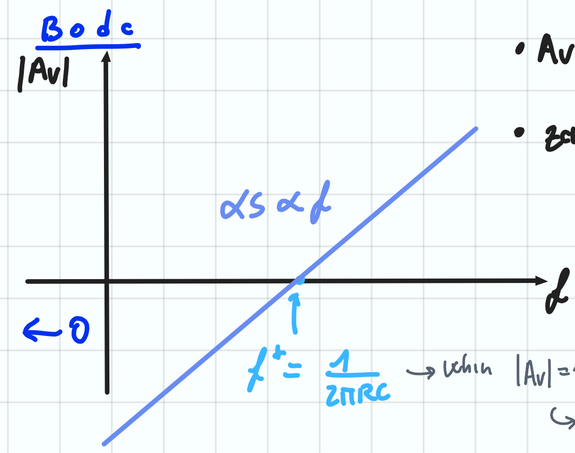
Issues

- HF gain $\to \infty$ (when $f \to \infty$)

⇓

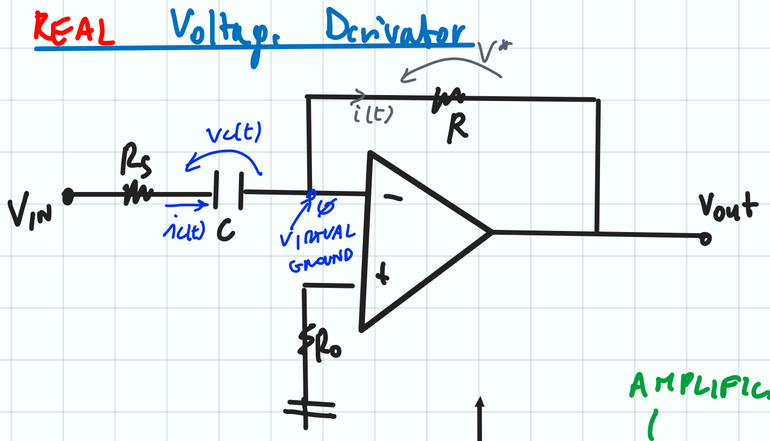
OpAmp will be to sensitive to HF noise

Bode

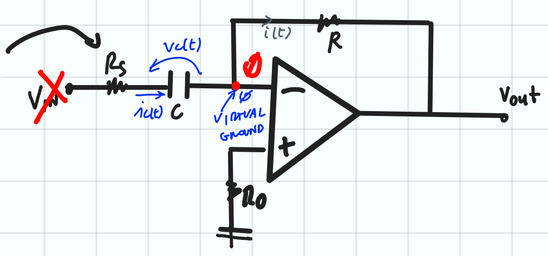


- $A_V = \frac{V_{out}}{V_{in}} = -s\tau$
- zero Oin $s=0$

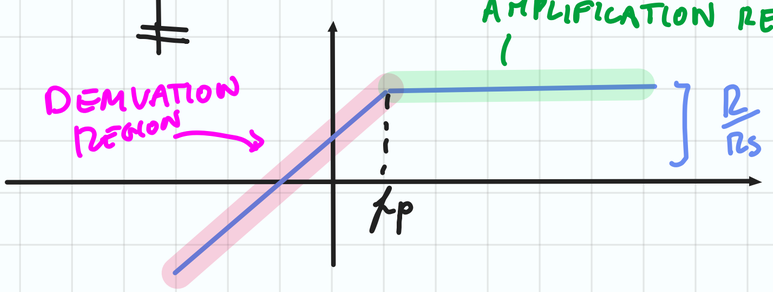
REAL Voltage Derivator



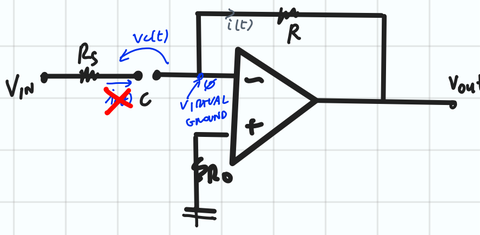
- $R_{eq} = R_S$



- pole at: $f_P = \frac{1}{2\pi R_{eq} C} = \frac{1}{2\pi R_S C}$

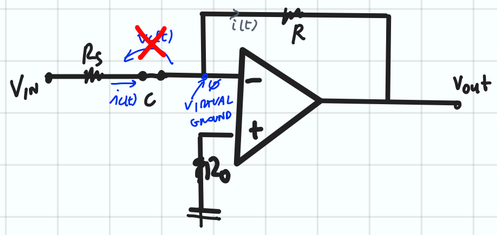


- **at DC** $[f \to 0]$



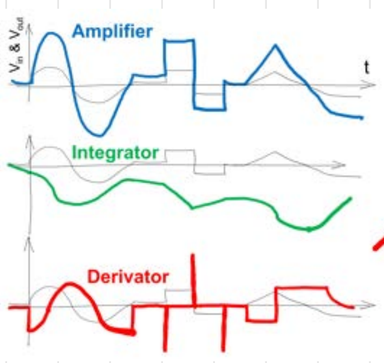
$A_V(0) = 0$

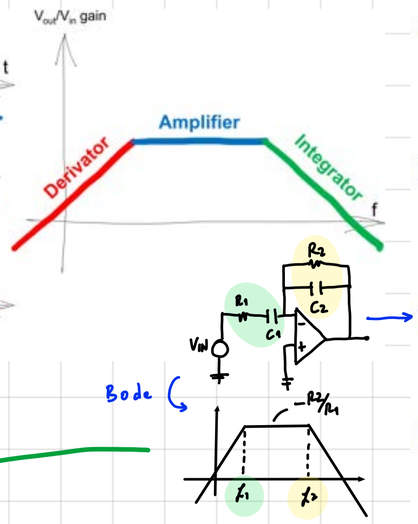
- **at HF** $[f \to \infty]$

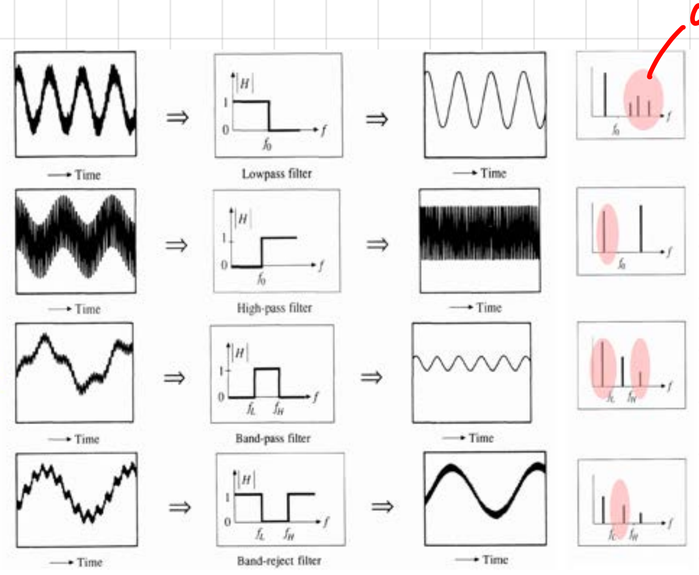


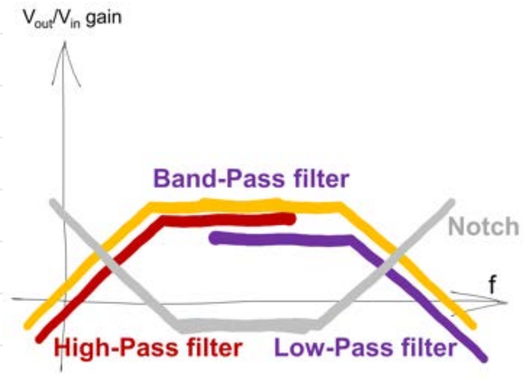
$A_V(\infty) = -\frac{R}{R_S}$

RECAP

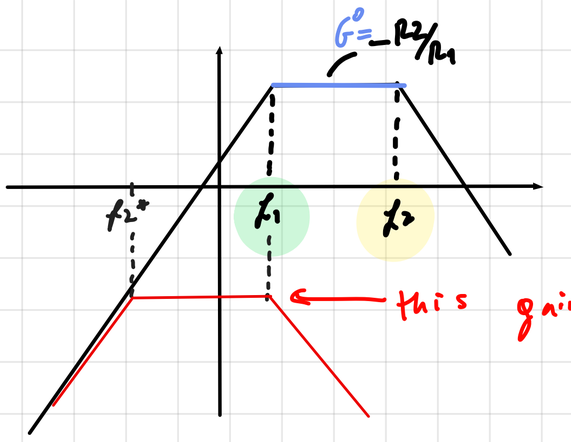








Ex. Band pass with WRONG-SIZING



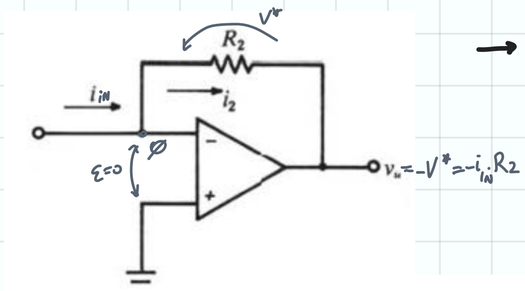
→ if $f_1 > f_2^*$

it means that the circuit with R_2, C_2 switches to short circuit $f > f_2\ (< f_1)$ before the circuit R_1, C_1 it will kill the signal before that circuit

this gain is $G = \frac{\left|-\frac{R_2}{R_1}\right|}{\frac{1}{2\pi C_1 R_1}} = \frac{\frac{R_2}{R_1} C_2 R_2}{C_1 R_1} = \left(\frac{R_2}{R_1}\right)^2 \frac{C_2}{C_1} < G^0$

CONVERTERS

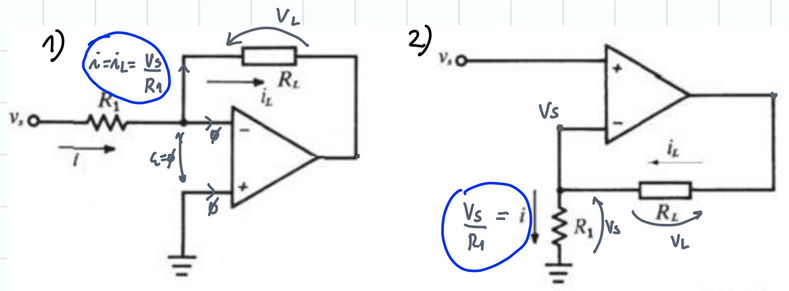
I → V (transimpendance) converter



→ $\frac{V_{OUT}}{i_{IN}} = -R_2\ [\Omega]$

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

V → I (transconductance) converter



→ from 1) $i_{OUT} = i_L = \frac{V_S}{R_1}$

→ from 2) $i_{OUT} = \frac{V_S}{R_1}$

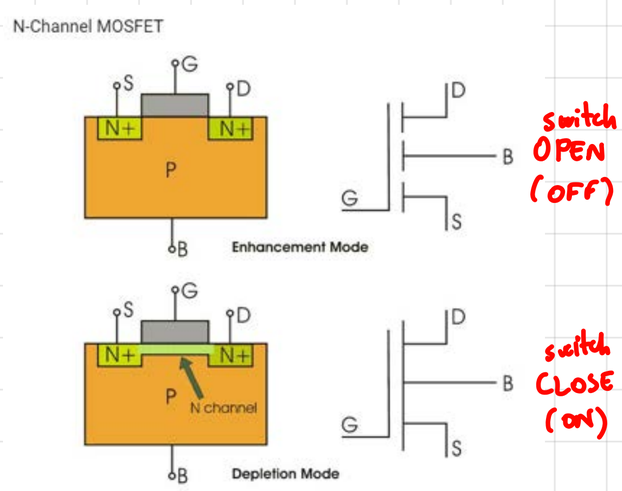
Issues:

- If $R_1 \uparrow$ Gain: $\frac{i_{out}}{V_{IN}} = \frac{1}{R_1} \downarrow$
- the current is limited by the current capability of the OpAmp that also limits the voltage capability

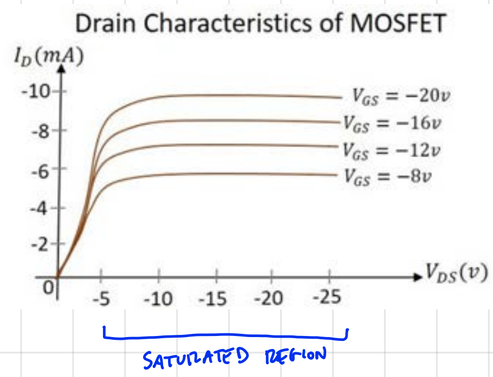
↳ If the current delivered by the OpAmp is too small → use MOSFETs (transistors) next.

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

TRANSISTORS RECAP



The drain and source are heavily doped N+ region and the substrate is p-type. The current flows due to the flow of negatively charged electrons, also known as n-channel MOSFET. When we apply the positive gate voltage the holes present beneath the oxide layer experience repulsive force and the holes are pushed downwards into the bound negative charges which are associated with the acceptor atoms. The positive gate voltage also attracts electrons from the N+ source and drain region into the channel thus an electron reach channel is formed.



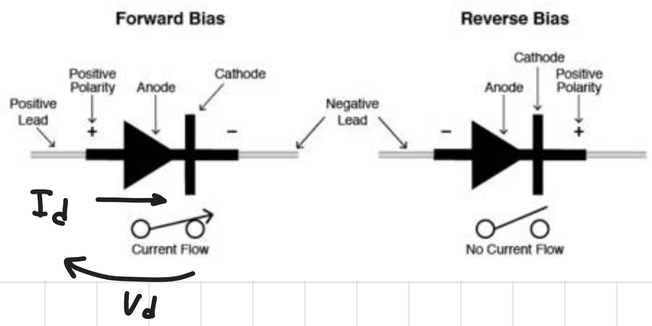
to be in the saturated region (threshold $(V_{GS} > V_T)$)

Formulas

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

coefficients; physical property; V_{OD} overdrive

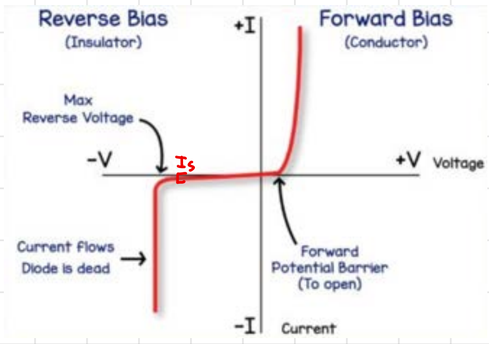
DIODE RECAP



A diode is a semiconductor device that essentially acts as a one-way switch for current. It allows current to flow easily in one direction, but severely restricts current from flowing in the opposite direction.

Diodes are also known as rectifiers because they change alternating current (ac) into pulsating direct current (dc). Diodes are rated according to their type, voltage, and current capacity.

Diodes have polarity, determined by an anode (positive lead) and cathode (negative lead). Most diodes allow current to flow only when positive voltage is applied to the anode. When a diode allows current flow, it is forward-biased. When a diode is reverse-biased, it acts as an insulator and does not permit current to flow.

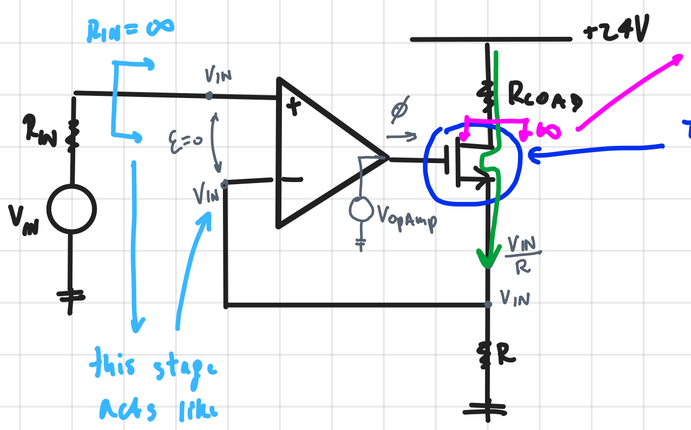


Formulas

- FORWARD: $I_d = I_s e^{\frac{V_D q}{KT}}$ (T: temperature)
- REVERSE: $I_d = I_s\left[e^{\frac{V_d q}{KT}} - 1\right]$

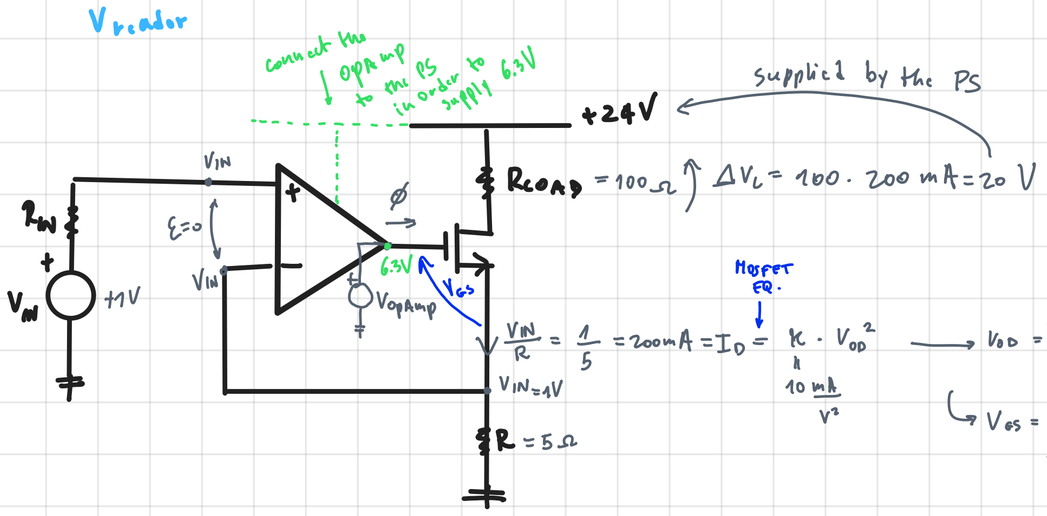
Non-linear circuits with OpAmp

In the previous analysis of the V→I converter we saw that one of the issue was the capability of the OpAmp to supply a possibly big current that must go to the load. We can use the following configuration:



→ the OpAmp can provide no current but just a voltage, we'll have the current $\frac{V_{IN}}{R}$ flowing through the transistor instead of directly the OpAmp

Ex.

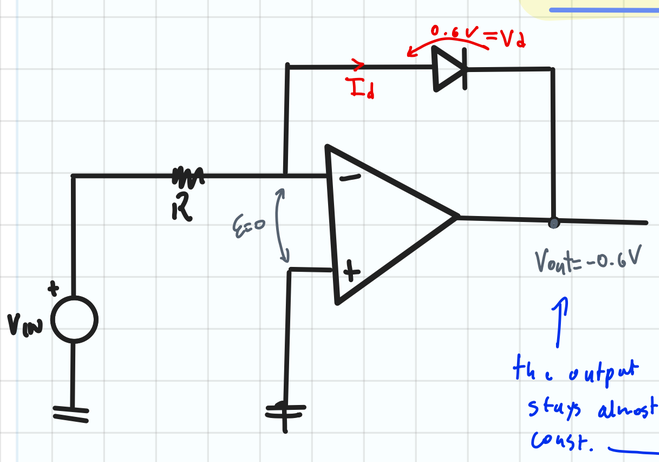


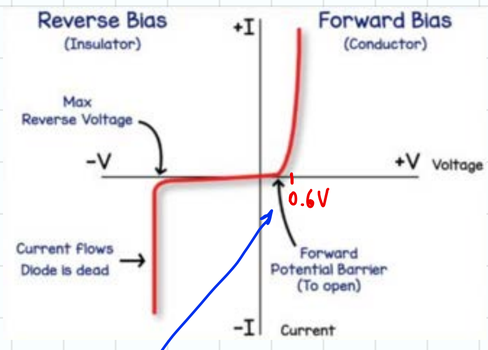
$$V_{OD} = \sqrt{\frac{200\,mA}{\frac{10\,mA}{V^2}}} = \sqrt{20} = 4.2V$$

$$V_{GS} = \underset{1.1V}{V_T} + V_{OD} = 5.3V \Longrightarrow V_{OpAmp} = V_{GS} + \underset{\text{on R}}{V_{IN}} = 6.3V$$

Exponential and Logarithmic CONVERTERS

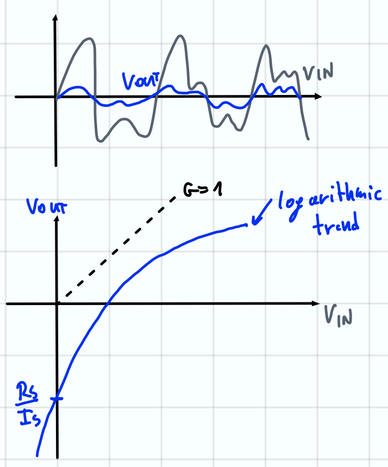
LOG



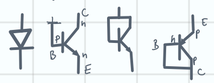


↳ If we analyze the diode equation we know that in the forward region:

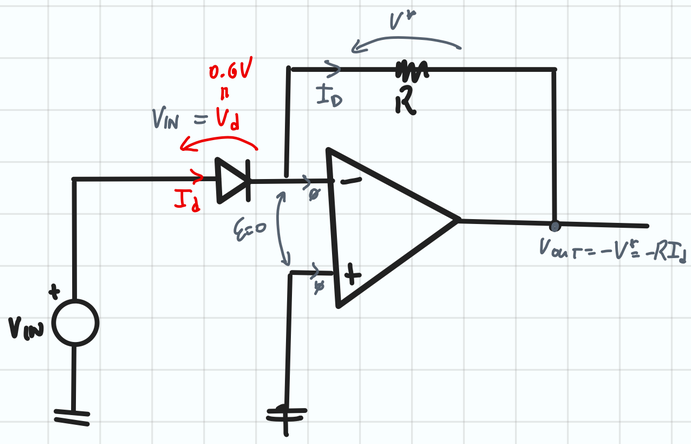
- $I_D = I_s e^{\frac{V_d q}{kT}}$
- $V_{out} = -V_D = -\underbrace{\frac{kT}{q}}_{V_{TH}} \ln\frac{V_{IN}}{R \cdot I_S} = -V_{TH} \cdot \ln\frac{V_{IN}(t)}{R \cdot I_S}$

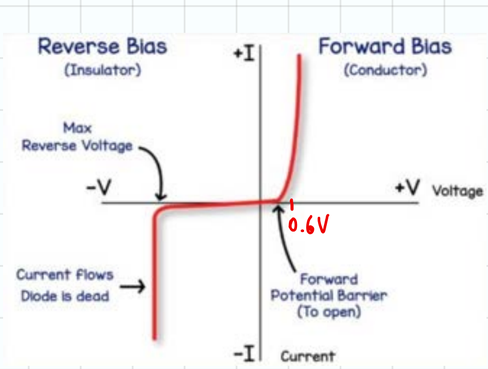


Note We could also have used transistor:



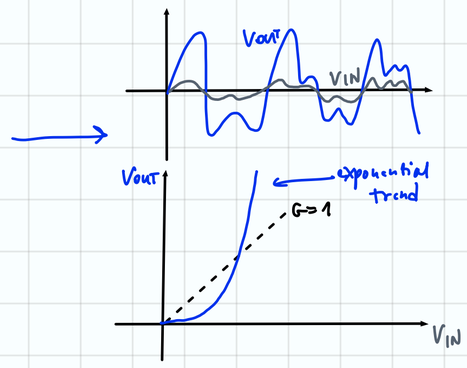
EXP



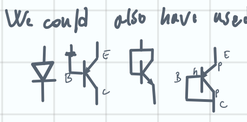


↳ If we analyze the diode equation we know that in the forward region:

- $I_D = I_s e^{\frac{V_d q}{kT}}$ ($V_d = V_{IN}$)
- $V_{out} = -V_r = -RI_D = -RI_s e^{\frac{V_{IN} q}{kT}}$

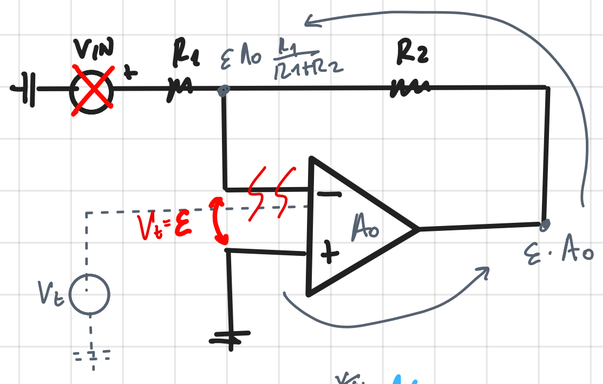


Note We could also have used transistor:



Ex. Gloop computation

1.



→ Turn off the source ✗

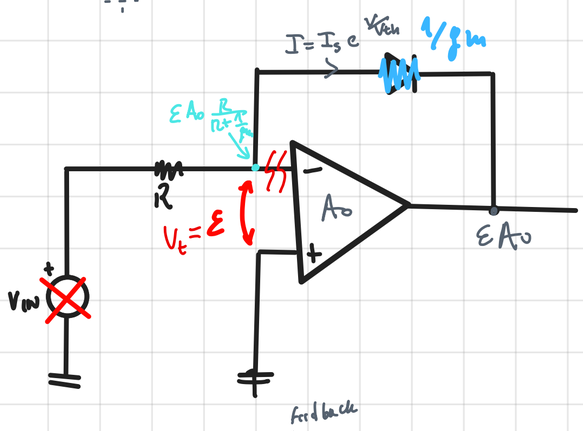
→ Apply a test voltage ε ($\not\!\!\not$ cut the loop)

V_t

↳ $V_f = \varepsilon A_0 \frac{R_1}{R_1+R_2}$

$$G_{loop} = \frac{V_f}{V_t} = A_0 \frac{R_1}{R_1+R_2}$$

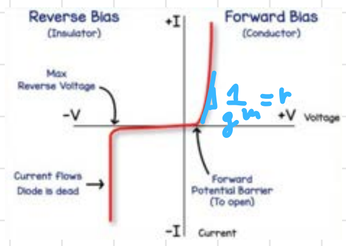
2.



→ Turn off the source ✗

→ Apply a test voltage ε ($\not\!\!\not$ cut the loop)

V_t



Diode Eq.: $I_D = I_S e^{V_D/V_{TH}}$

$$\frac{dI}{dV} = \frac{I_S}{V_{TH}} e^{V_D/V_{TH}} = \frac{I_D}{V_{TH}}$$

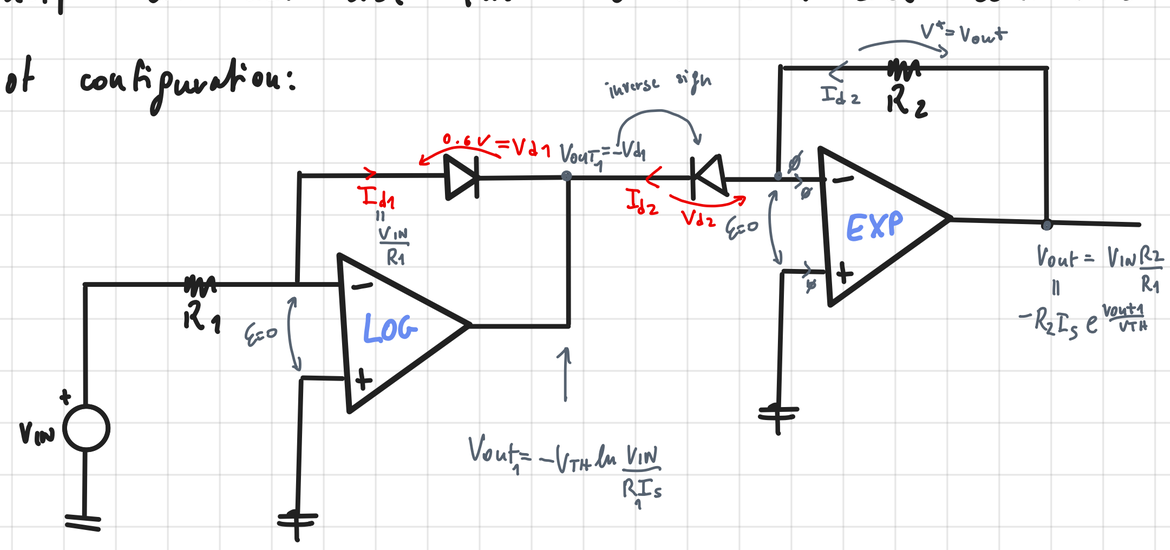
$\frac{dI}{dV} = g_m$ (TRANSCONDUCTANCE)

We can consider the diode with its small signal equivalent $\frac{1}{g_m}$

$\Longrightarrow V_f = \varepsilon A_0 \frac{R}{R+\frac{1}{g_m}}$ → $G_{loop} = A_0 \frac{R}{R+\frac{1}{g_m}}$

Note: If the diode it's in reverse mode the feedback is OFF.

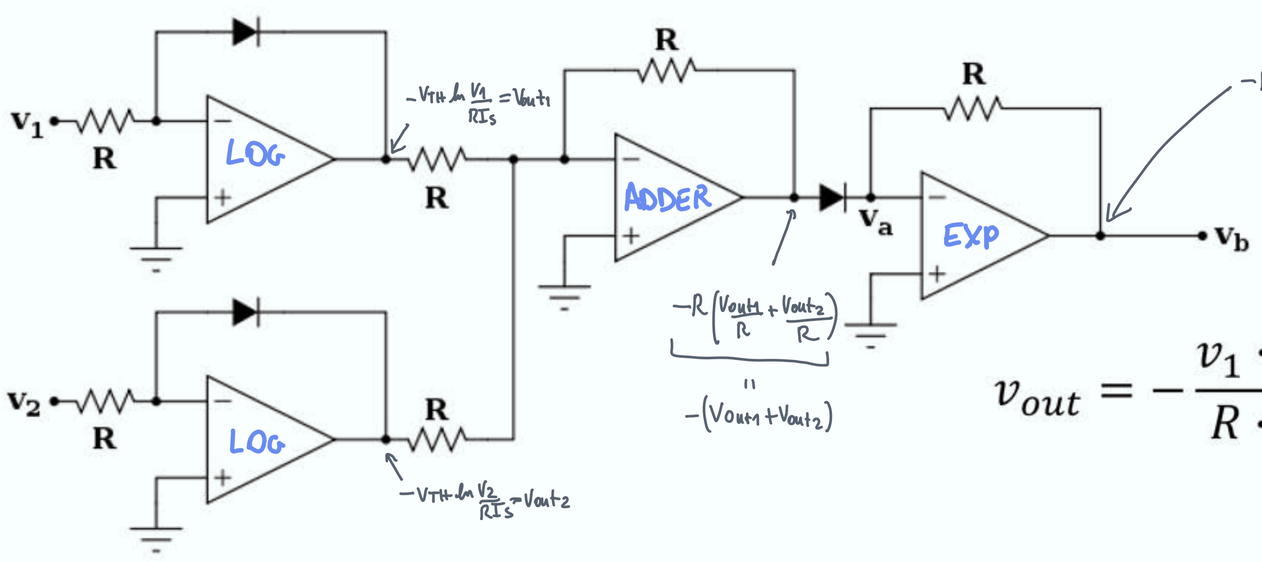
Ex. In principle we can use the EXP and LOG converters to obtain the same input ($\ln(e^{V_{IN}}) = V_{IN}$) with this type of configuration:



$\Longrightarrow V_{out} = V_{IN}\frac{R_2}{R_1}$

But in practice it wouldn't be that useful, we can use instead these configuration in a voltage multiplier configuration....

Voltage Multiplier



$$v_{out} = -\frac{v_1 \cdot v_2}{R \cdot I_S}$$

$-RI_S e^{-\frac{(V_{out1}+V_{out2})}{V_{TH}}} = -RI_S e^{\frac{+V_{TH}\left(\ln\left(\frac{V_1}{RI_S}\cdot\frac{V_2}{RI_S}\right)\right)}{V_{TH}}} = -RI_S \frac{V_1 \cdot V_2}{R^2 I_S^2} = -\frac{V_1 V_2}{RI_S}$

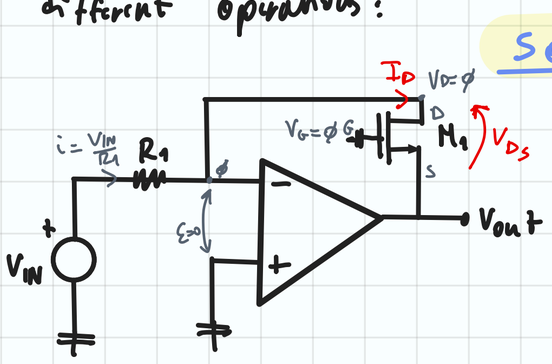
We can compute the product between two input voltages!

Historical Note The Op Amp is called OPERATIONAL because indeed it can perform, through different amplifier configuration, multiple operations ($+, -, \times, \frac{d}{dt}, \int dt \dots$)

Square Root And Power 2 Converters

In the previous stages we've used diodes or simple transistors (BJT), while if we use MOSFET we can obtain different operations:

SQRT



MOSFET Equation (to be in the saturated region)

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

threshold ($V_{GS} > V_T$)

coefficients; physical property; V_{OD} overdrive

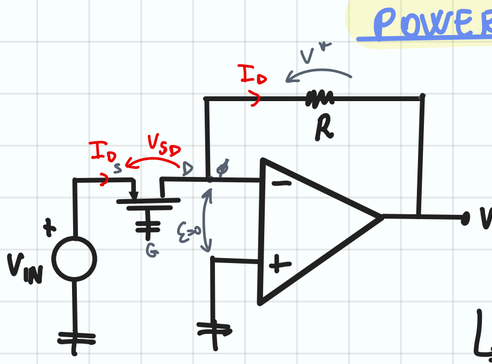
$i = \frac{V_{IN}}{R_1}$

↳ $V_{OD}^2 = (V_{GS}-V_T)^2 = \frac{I_D}{K} = \frac{V_{IN}}{R_1 K}$

↳ $V_{GS} = V_T + \sqrt{\frac{V_{IN}}{R_1 K}}$

$V_{OUT} = -V_{DS} = -(V_G - V_S) = V_S = V_G - V_{GS} = -V_T - \sqrt{\frac{V_{IN}}{R_1 K}}$ ← Square root

POWER2



MOSFET Equation

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

↳ $V_{out} = -RI_D = -RkV_{OD}^2 = -Rk(V_{IN}-V_T)^2$ ← Power 2

$V_D = 0$

$V_S = -V_{IN}$, $V_G = \phi$ → $V_{GS} = V_{IN}$

→ $V_{OD} = V_{GS} - V_T = V_{IN} - V_T$

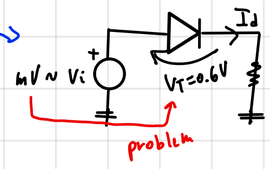
PRECISION RECTIFIER: Super Diode

(Book p. 159)

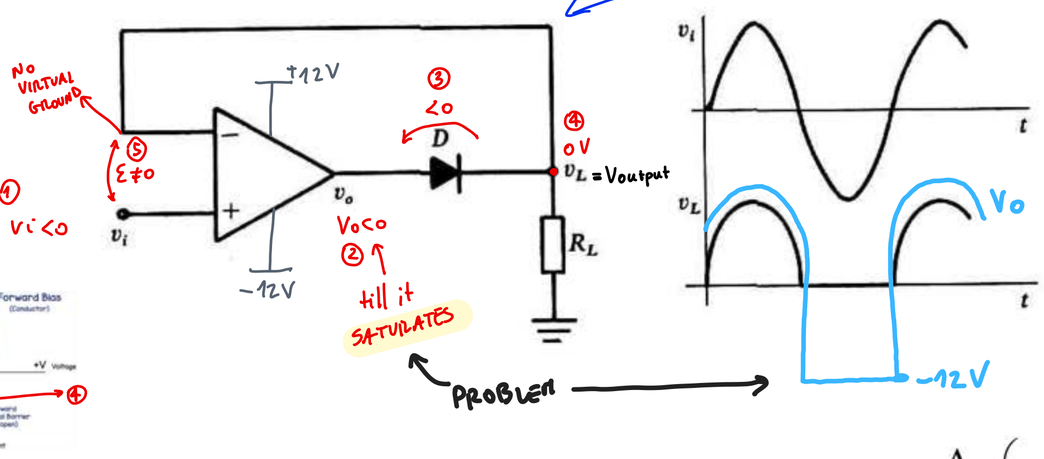
There are many circuits with OpAmps that realize non-linear functions (for which the superimposition principle does not apply). We can start considering the precision rectifiers.

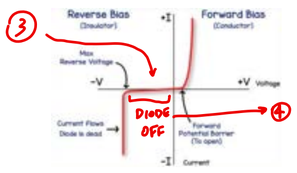
In the case of non-DC signals (AC) with amplitude of some millivolts, they can be rectified and become DC signals if the threshold of the rectifier is properly infinitesimal. For this reason, simple diodes are not sufficient because they have thresholds of about V_γ=0.6-0.7V. A combined use of diodes and OpAmps can reduce the rectifier threshold.

→ The amplifier is useful in the case where we want to rectify a small signal (input v_i), indeed with just a diode we wouldn't be able to work properly given that the small signal's voltage << V_T (0.6-0.7 V), so we amplify this signal in order to correctly process (rectify) it through a diode. V_γ = threshold of the diode



improvement wrt this circuit





Computations:

no virtual ground

- for a differential voltage $\varepsilon = V_i - V_L \neq 0$

→ $V_o = A \cdot \varepsilon = V_L + V_\gamma$ → $V_L = A(V_i - V_L) - V_\gamma$

(V_o: open loop gain (no feedback))

↳ $(1+A)V_L = AV_i - V_\gamma$

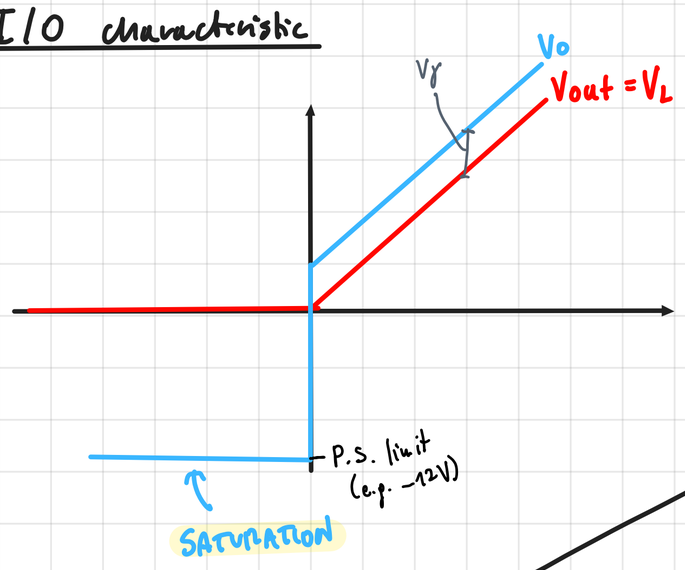
$$\Rightarrow V_L = \frac{A}{1+A}\left(v_i - \frac{V_\gamma}{A}\right) \cong v_i - \frac{V_\gamma}{A}$$

↳ for $\varepsilon \to 0$ $V_L \simeq v_i$

- positive V_i …. current through R_L… diode goes ON: $v_L = \frac{A}{1+A}\left(v_i - \frac{V_\gamma}{A}\right) \cong v_i - \frac{V_\gamma}{A}$

- negative V_i …. no current allowed through R_L… since diode is OFF: $v_L = 0$ → on the circuit

I/O characteristic

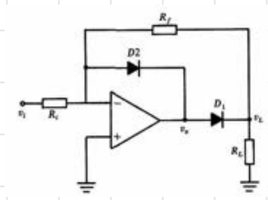


→ Issue: SATURATION → to improve this problem we can choose a configuration with 2 diodes.

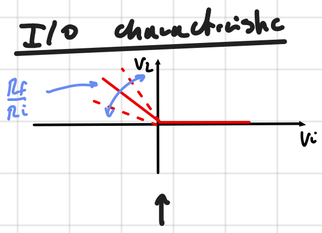
Super-diode applications: For NO SATURATION

INVERTING

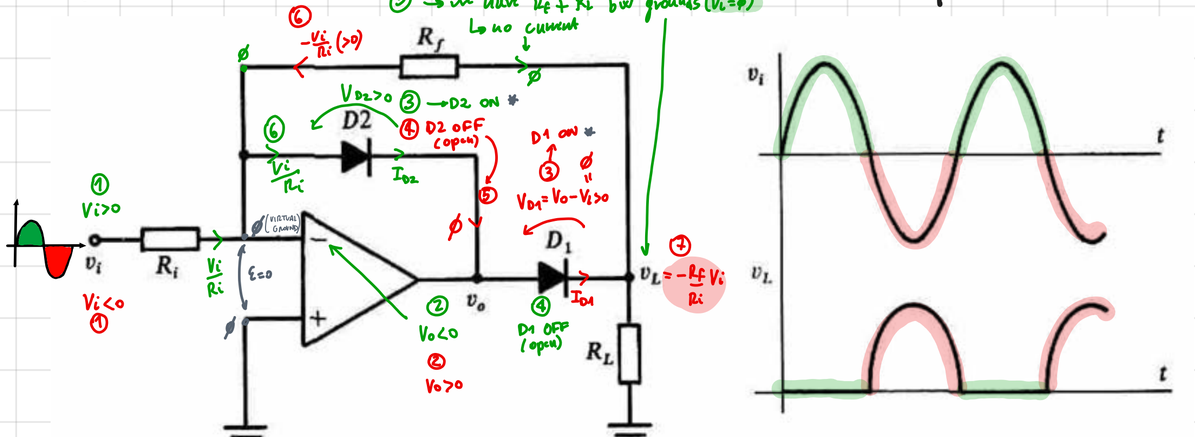
circuit:



I/O characteristic



Computations:

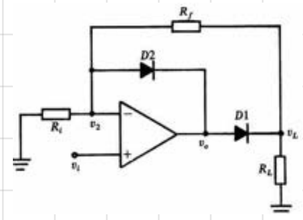


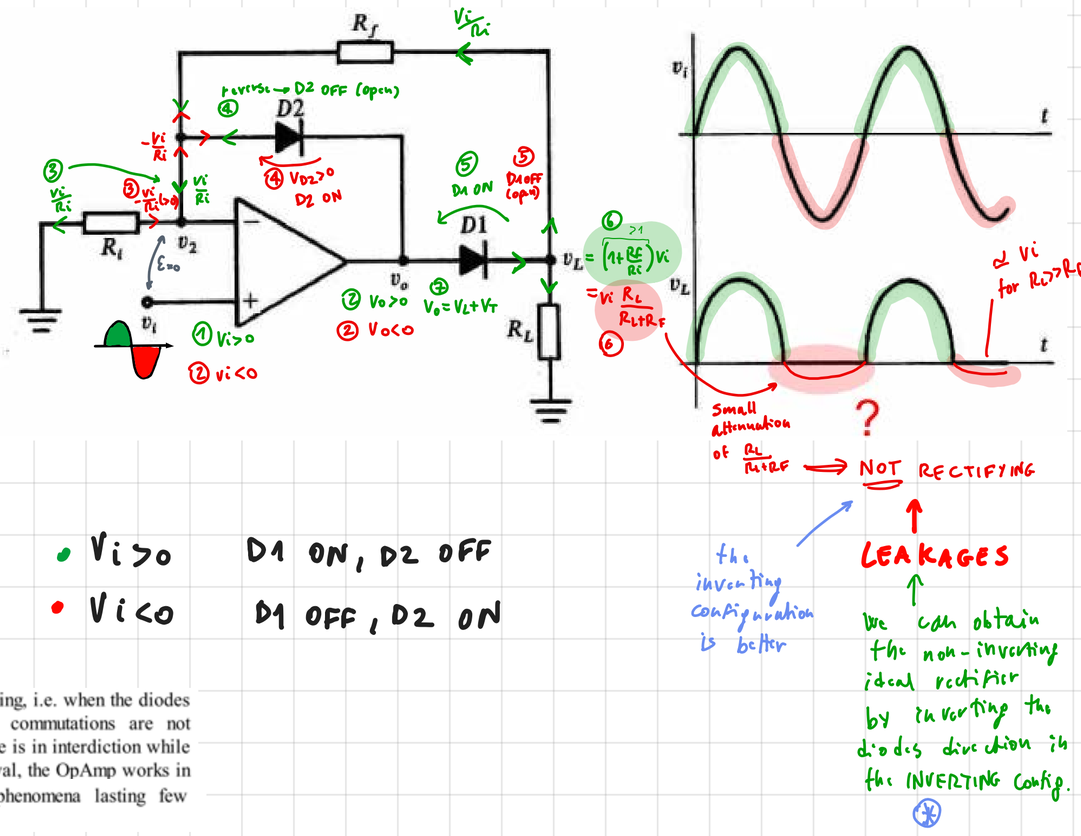
* Remember that for DIODE to be ON at least V_γ = 0.6 V is required across the diode

- V_i>0 D2 ON (forward), D1 OFF (reverse) $I_{D_2} = \frac{V_i}{R_i}$ (>0)
- V_i<0 D1 ON (forward), D2 OFF (reverse) $I_{D1} = -\frac{V_i}{R_i}$ (>0)

NON-INVERTING

circuit:





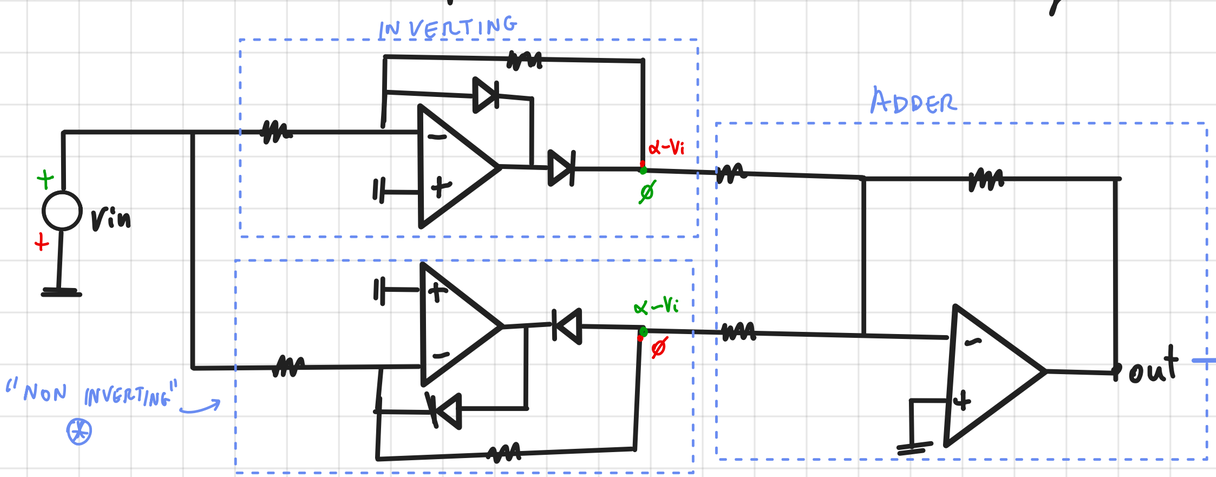
- V_i>0 D1 ON, D2 OFF
- V_i<0 D1 OFF, D2 ON

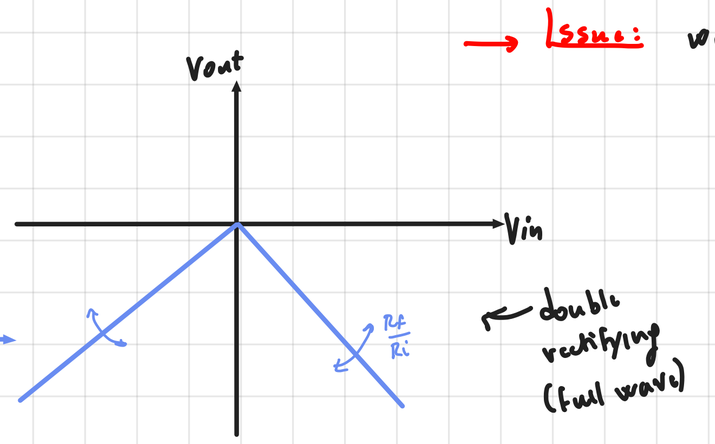
(Book p. 162)

In both cases, there is the problem of phase switching, i.e. when the diodes are switched, conductive, or in interdiction, the commutations are not immediate, and there is a period during which a diode is in interdiction while the other is not yet conductive. During this time interval, the OpAmp works in an open loop. However, these are very short phenomena lasting few microseconds.

Double-Rectifier

We can take the 2 previous half-wave rectifier configurations and sum them together:





← double rectifying (full wave)

→ Issue: we need too many components

↳ 3 OpAmp

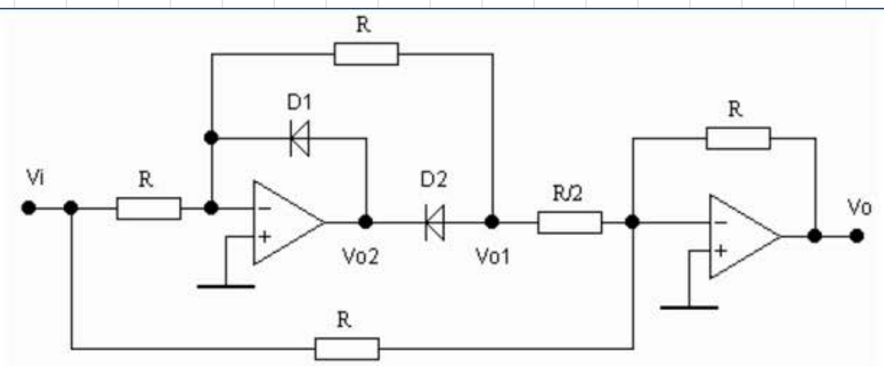
↳ 4 diodes

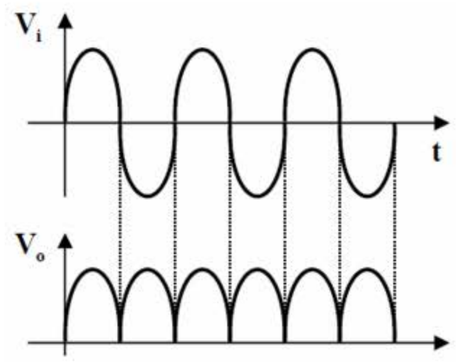
↳ Solution: Other configurations

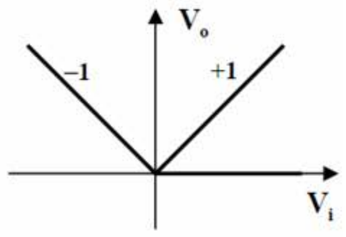
Super Double-Rectifier

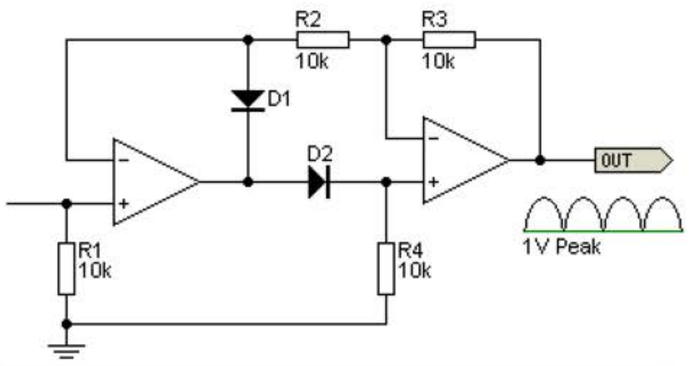
Configurations with just 2 OpAmp and 2 diodes

(See appendix for further analysis)



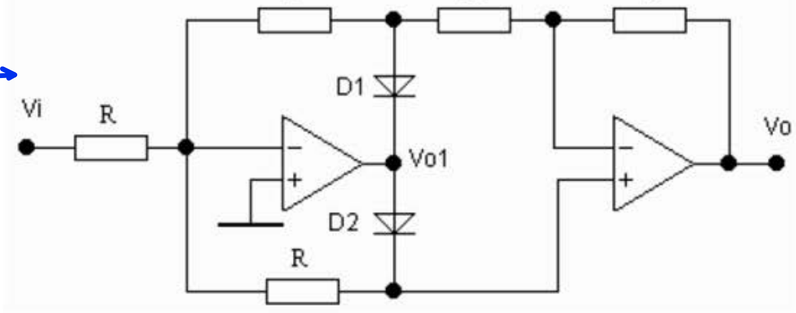






(Book p. 163)

Consider now the circuit shown in Fig. 2.39; it is a full-wave rectifier. When v_i is positive, and the output v_o is negative, the diode D_1 conducts (the v_2' terminal is a virtual ground) because D_2 is in interdiction (the v_1'' terminal is virtually grounded because there is no current flowing through the resistance R).



Comparator

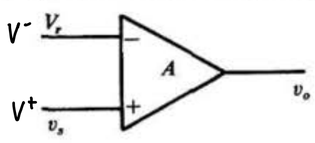
(Book p. 165)

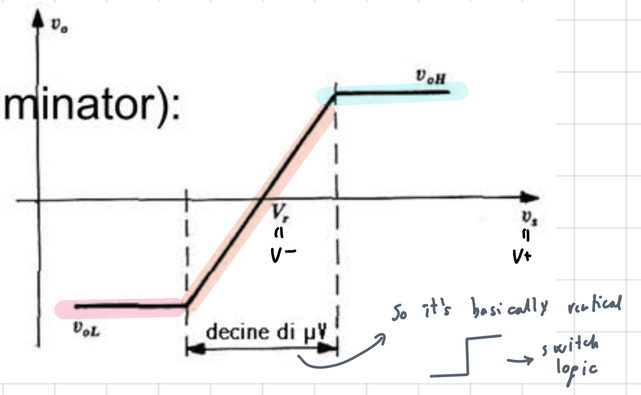
The comparator presents two inputs connected to a constant reference voltage V_r and to the signal v_s. The output can assume only two distinct values (Fig. 2.42) to detect if the signal v_s is above or below the reference voltage.

Using a differential amplifier without a feedback network (i.e. with a very high open loop gain A), it is possible to obtain a very sharp input linear region. In fact, for a standard OpAmp, only few tens of μV are enough to saturate the output, as shown in Fig. 2.43. Often, this switch interval is defined as the "resolution". There are OpAmps with the intended purpose of comparing voltages with very high resolution, for example for the μA311, it is below 15 μV.

Generally the voltages v_{oL} and v_{oH} at which the OpAmp or the comparator saturates are close to the supply voltages of the device, for example 0-5V, ±5V, ±12V. To make the output signal compatible with digital circuits (or, more generally, to make it independent of supply voltages), dedicated component are required. Another method to limit the voltage is the use of Zener diodes, as

Open-loop OpAmp (one-threshold discriminator):





↳ • $V^+ > V^- \Rightarrow V_o \propto (V^+ - V^-) > 0$

• $V^- > V^+ \Rightarrow V_o \propto (V^+ - V^-) < 0$

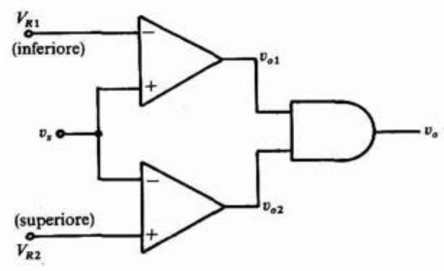
• $V^+ \simeq V^- \Rightarrow V_o \simeq 0$

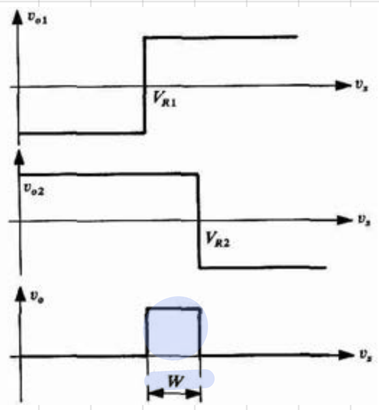
Window comparator

(Book p. 168)

The combination of an inverting and a non-inverting comparator, with two different reference voltages, allows the implementation of a circuit suitable for detection if the input voltage v_s is in a certain range. The comparator output is connected to an AND gate which provides a high output value only if both inputs are high. This is verified only in the window $W=V_{R2}-V_{R1}$ (Fig. 2.48).

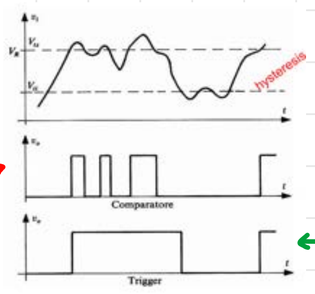
"Window" (two-thresholds) discriminator:





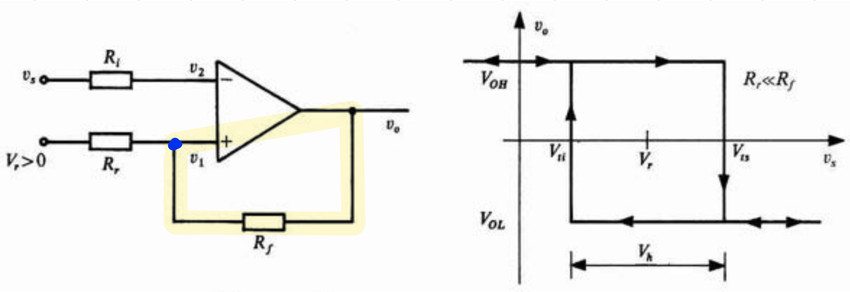
Schmitt Trigger

The comparators presented in the preceding chapters have a single reference voltage for threshold switching: it is evident that normal noise on this reference can cause many undesired commutations. The comparator with hysteresis (also known as Schmitt Trigger) does not suffer from this problem as it has two different input thresholds, V_{ts} for the rising signal and V_{ti} for the decreasing signal. The difference between them is the *hysteresis* or the dead zone V_h.



← TRIGGER SCHMITT

INVERTING Schmitt Trigger



Usually let's choose $R_f >> R_r$

so ...

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

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

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

To introduce the hysteresis in a comparator stage, it is sufficient to add slightly positive feedback to the OpAmp through the resistors R_f and R_r, as shown in Fig. 2.50. The signal v_s is applied to the non-inverting input terminal through the R_i resistance (irrelevant to the operation of the system) while the external reference voltage V_r is applied to the resistance R_r. The voltage applied to the non-inverting terminal is (with the superposition principle):

$$V_1 = \frac{R_r}{R_f + R_r}\cdot V_o + \frac{\cdot R_f}{R_f + R_r}\cdot V_r$$

Assuming that $v_s < v_1$ is the initial state, the output voltage v_o is equal to V_{OH}. If the signal v_s increases, the output commutation will be for $v_s = v_1$, i.e. corresponding to the upper threshold equal to:

$$V_{th} = v_{1(th)} = \frac{R_r}{R_f + R_r}\cdot V_{OH} + \frac{R_f}{R_f + R_r}\cdot V_r = \frac{R_r V_{OH} + R_f V_r}{R_f + R_r}$$

In this case, the output v_o switches to low, becoming equal to V_{OL}. The output continues to remain low even if the signal v_s starts to grow.

When the signal v_s decreases, the output switches to high for $v_s = v_1$, but the value v_1 has changed compared to the previous case because of the different value of the output voltage v_o. This new value is the lower threshold equal to:

$$V_{tl} = v_{1(tl)} = \frac{R_r}{R_f + R_r}V_{OL} + \frac{R_f}{R_f + R_r}V_r = \frac{R_r V_{OL} + R_f V_r}{R_f + R_r}$$

The hysteresis voltage is equal to the difference between the two different commutation thresholds:

$$V_h = V_{th} - V_{tl} = (V_{OH} - V_{OL})\frac{R_r}{R_f + R_r}$$

and it depends on both the ratio R_f/R_r and the output voltage, not the reference voltage V_r. Normally, the hysteresis is less than the output voltage, and this can be obtained with $R_r \ll R_f$. In this case, we have:

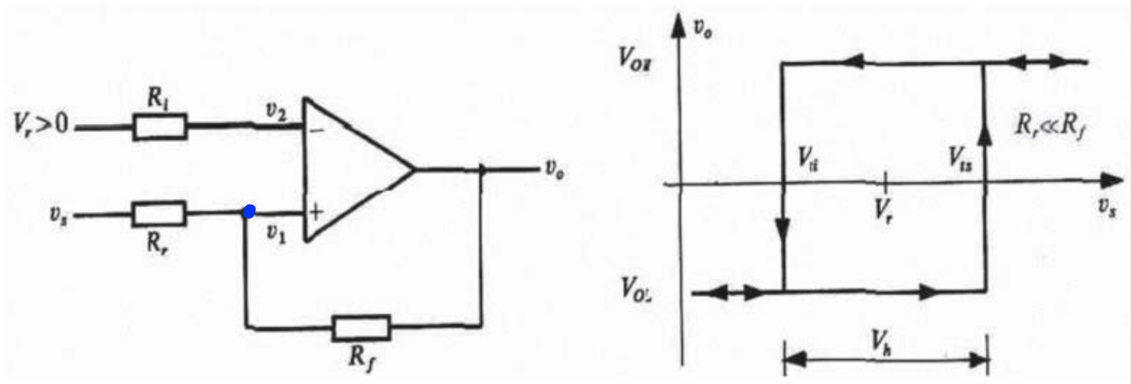
$$V_h \cong (V_{OH} - V_{OL})\frac{R_r}{R_f} \qquad V_{th} \cong V_r + \frac{V_h}{2} \qquad V_{tl} \cong V_r - \frac{V_h}{2}$$

Notice that V_r defines the "distance" between the center of the hysteresis cycle and the origin of the characteristic if $V_{OH} = -V_{OL}$. In fact, we generally have:

$$\frac{V_{th} + V_{tl}}{2} = \frac{R_r}{R_r + R_f}\frac{V_{OH} + V_{OL}}{2} + V_r$$

And particularly if $(V_{th}+V_{tl})/2=0$, we have a detector with no hysteresis.

NON- INVERTING Schmitt Trigger



The non-inverting Schmitt Trigger is similar to the inverting configuration, but the input signal v_s and the reference V_r are reversed, as shown in Fig. 2.51. With an analogous reasoning, as for the inverting configuration, we have:

$$v_1 = \frac{V_o}{R_f + R_r} R_r + \frac{V_S}{R_f + R_r} R_f$$

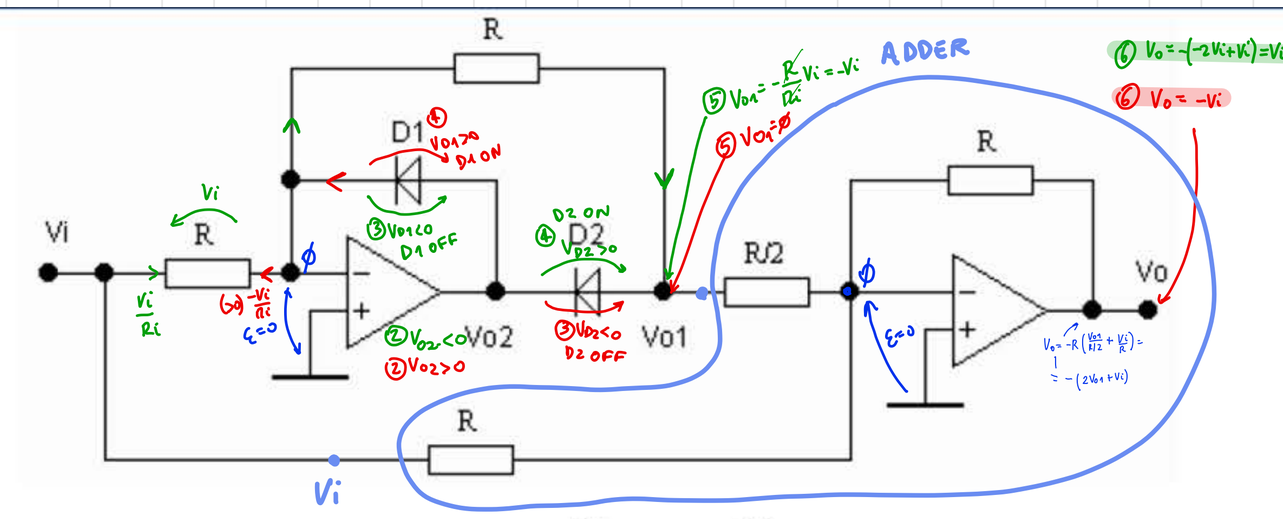
Until the voltage v_s is not enough to satisfy $v_1 \geq v_2 = V_r$, the output voltage v_o is equal to V_{OL}. The output switching is for:

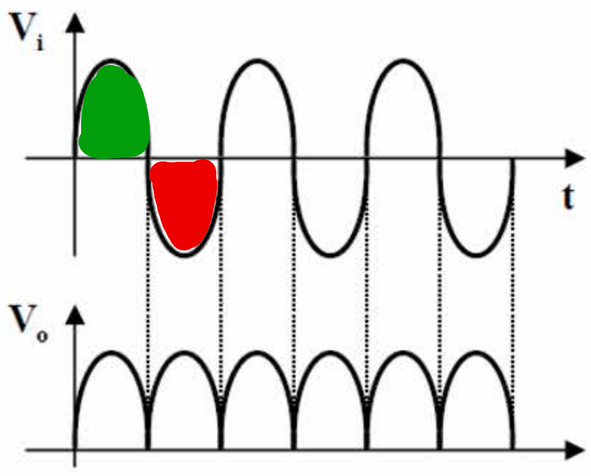
output is low

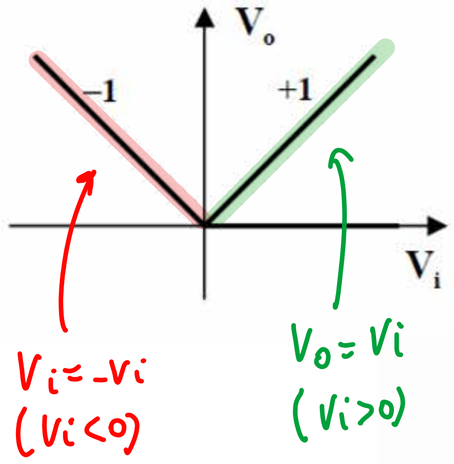
$$v_1 = V_r = \frac{V_{OL}}{R_f + R_r} R_r + \frac{v_s}{R_f + R_r} R_f \rightarrow \varepsilon = 0$$

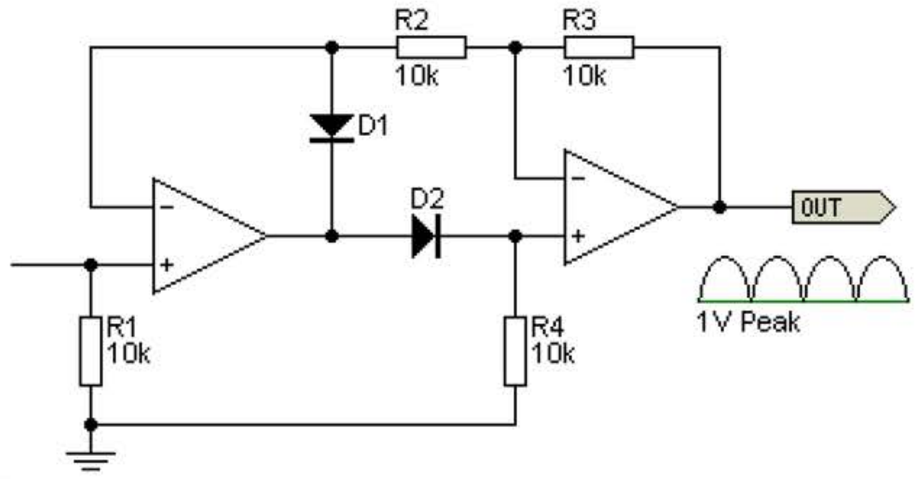
and thus $v_s = V_{th} = \frac{R_f + R_r}{R_f} V_r - \frac{R_r}{R_f} V_{OL}$. The output voltage v_o changes its value to V_{OH} and remains stable for every further increase of the signal v_s. When the signal decreases, we will have the commutation for $v_1 = V_r = \frac{V_{OH}}{R_f + R_r} R_r + \frac{v_s}{R_f + R_r} R_f$ and thus with $v_s = V_{tl} = \frac{R_f + R_r}{R_f} V_r - \frac{R_r}{R_f} V_{OH}$.

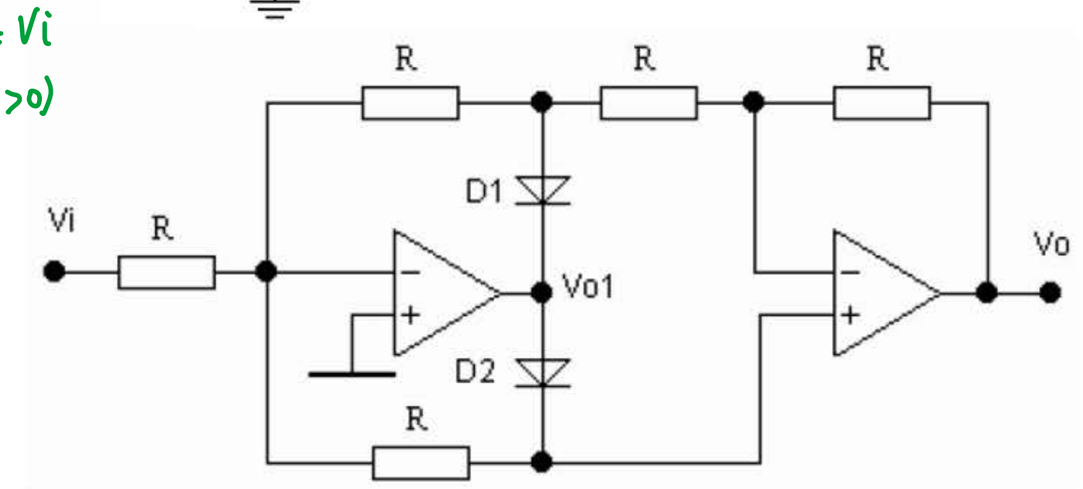
The hysteresis is: $V_h = V_{th} - V_{tl} = (V_{OH} - V_{OL}) \frac{R_r}{R_f}$


R
D1
Vi
R
Vi
Vi
R
ϕ
D2 ON
D2
R/2
R
ϕ
Vo
Vo1
R
Vi
ADDER
① Vi > 0
② Vi < 0
④ Vo1 > 0 D1 ON
③ Vo1 < 0 D1 OFF
② Vo2 < 0 Vo2
② Vo2 > 0
④ VD2 > 0
③ VD2 < 0 D2 OFF
⑤ Vo1 = 0
⑥ Vo = -Vi


Vi
t
Vo


Vo
−1
+1
Vi
Vi = -Vi
(Vi < 0)
Vo = Vi
(Vi > 0)


R2
10k
R3
10k
D1
D2
OUT
R1
10k
R4
10k
1V Peak


R
R
R
D1
Vi
R
Vo1
Vo
D2
R

ES04_FREQUENCY COMPENSATION

Electronic Systems

Intro

(Book p. 199)

The first thing to clarify in order to avoid any misunderstanding is that the frequency response study is a small-signal study; that is, a study that analyzes components or circuitry when they deviate from their operating point a little and do not come out of their linear operation range, reaching the saturation or even the cut-off.

Let us consider stages employing operational amplifiers (OpAmps) which have their own intrinsic frequency response A(s), the so-called "open loop", which represents the gain between the differential signal applied to the two input terminals and the corresponding output at different frequencies. Since typical values of A_0=A(0) are well above 100000, OpAmps cannot operate with an open loop, but have to be used in a negative feedback configuration with the output signal coming back (in full or reduced) to the inverting OpAmp pin. The "loop closure" leads to a reduction of the OpAmp gain, but, at the same time, makes it more precise, widens the stage bandwidth, reduces the output impedance, raises the input impedance (seen from the positive terminal), almost completely zeroes the impedance of the virtual ground (seen from the inverting terminal), and more. In other words, negative feedback redresses all the "defects" of the real OpAmp, enhancing the performance of the stage.

A noticeable disadvantage of feedback is that it may cause the instability of the stage. The aim of this chapter is to understand whether feedback applied to an OpAmp is still able to maintain stability or not.

Commonly, we say that an OpAmp is compensated if the second pole f_{p2} of its open-loop gain A(s) is at a frequency higher than the intersection of $G_{loop}(s)$ (in the case of the buffer configuration, $G_{loop}(s)$ overlaps A(s)) and the 0dB axis, as shown in Fig. 3.1. By connecting such an OpAmp in buffer configuration (see Fig. 3.1), the ideal closed-loop gain becomes unity up to high frequencies. In the real gain, even new poles will arise. This chapter will show how to compute these poles, how to verify whether the stage remains stable, and, in case of instability, how to make it stable, i.e. compensated. In the case of the buffer in Fig. 3.1, we will find out that the poles are placed at f_{plow} (corresponding to $G_{loop}(s)$=1) and at f_{phigh} (coinciding with the second OpAmp pole f_{p2}). The phase margin PM is the difference between the phase of G_{loop} (negative) and -180° (see Fig.3.1). In fact, an additional phase shift of G_{loop} of -180° makes the loop have positive feedback (with a net phase shift of -360°). The PM is a function of frequency, but we consider it only at the intersection of G_{loop} and the 0dB axis if we are interested in the stability of the circuit. In the case of the circuit of Fig. 3.1, the stage will certainly be stable with a phase margin much greater than 45°. A system is unstable if the PM is strictly less than zero degrees.

To better understand how various factors, such as G_{loop}, A(s), and β(s), affect the system's stability, it is advisable to analyze the system's Bode plots (magnitude and phase), which also helps predict the time response. Moreover, it is handy to know how to use the tool represented by the root locus. Fig. 3.2 summarizes the typical stage behavior with two separate poles (at the top) or complex conjugate poles that lead to a stable system (in the middle) or gradually to a more and more unstable one (at the bottom). In the figure, |G| is the closed-loop gain of the stage, not the loop gain G_{loop}.

Fig. 3.1

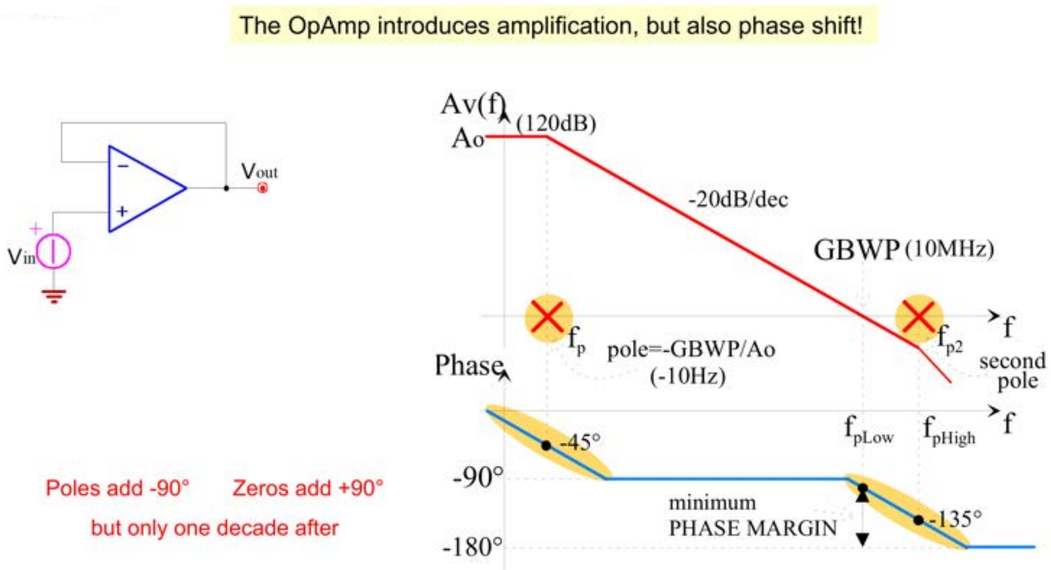
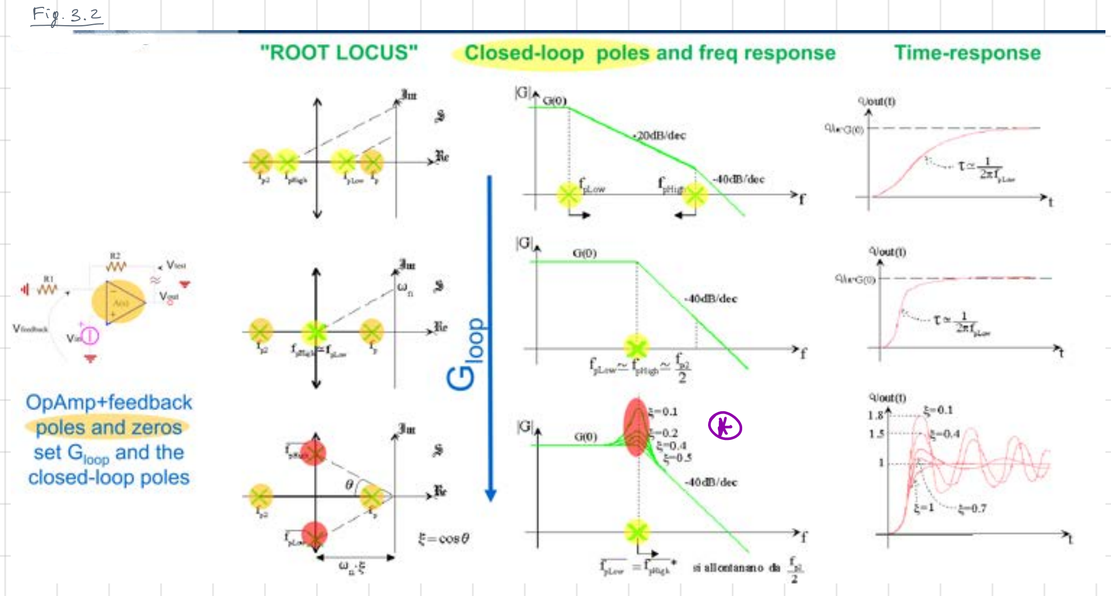
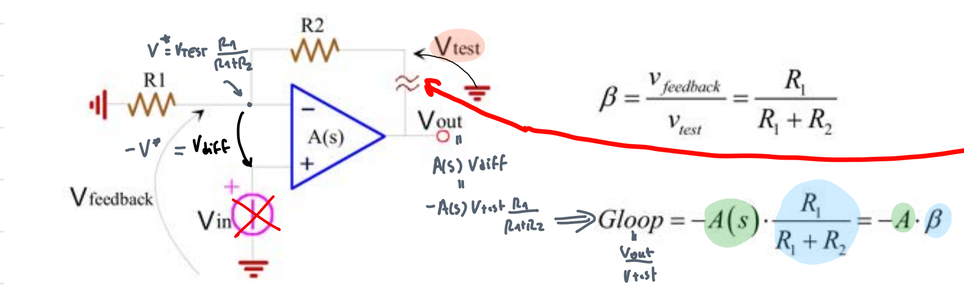


Fig. 3.2



Analysis basics

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



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

$$\Rightarrow G_{loop} = \frac{V_{out}}{V_{test}} = -A(s)\cdot\frac{R_1}{R_1+R_2} = -A\cdot\beta$$

(Non-inverting stage)

Real closed-loop gain:

$$G_{NI} = \frac{A}{1 + A\cdot\beta} =$$

1/β if $1/\beta \ll A$ (i.e. if $G_{loop} \gg 1$) :Ideal gain

A if $1/\beta \gg A$ (i.e. if $G_{loop} \ll 1$)

Steps:

1. Turn off the input sources ✕
2. Cut the loop ≶
3. Apply a V_{test}
4. Compute the open-loop output $V_{out} = A(s)\,V_{diff}$

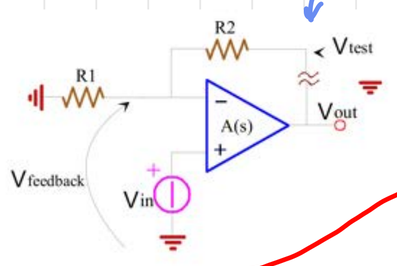
$$\Rightarrow G_{loop} = \frac{V_{out}}{V_{test}}$$

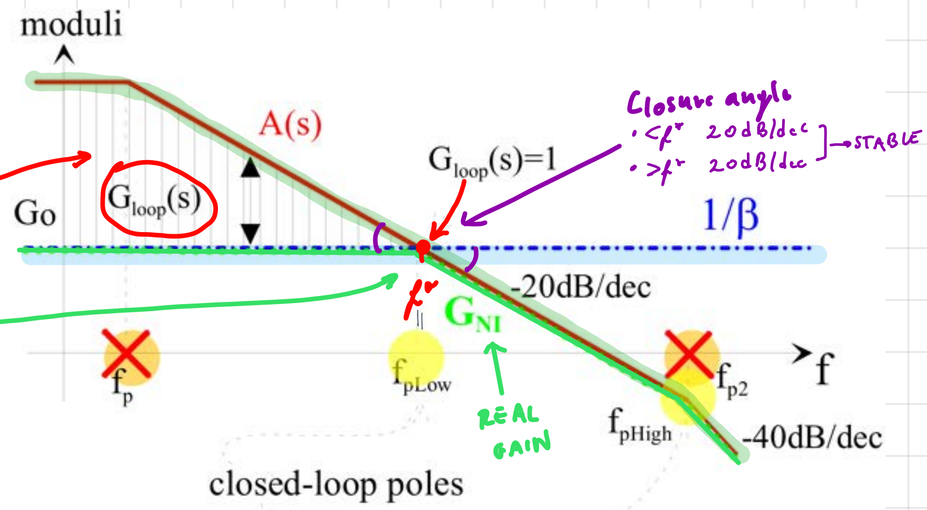
Gloop assessment and real gain extraction:

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:

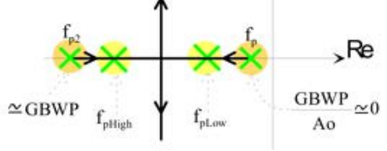
1. Draw A(s) — i.e. the frequency response of the OpAmp
2. Compute β(s) (freq. dependent if there are components like C → add poles/zeros) — NO Laplace, just asymptotic analysis at 0, medium and ∞ freq.s
   then draw 1/β(s) — remember: poles of β(s) lift up, zeroes of β(s) put down 1/β(s)
3. The split 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*

NON INVERTING stage





Root locus:



Phase Margin: $\Phi_M = 180° - 90° - arctg\frac{|f_{plow}|}{|f_{p2}|}$

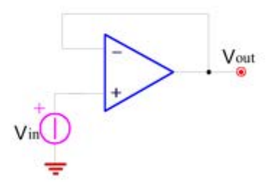
Stability assessment:

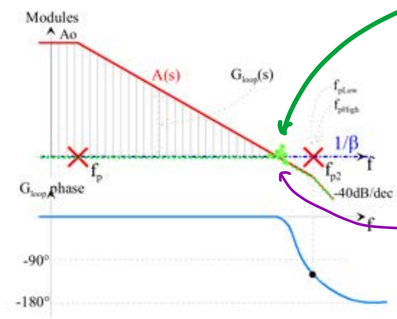
The reason why "closure angle" matters on **stability** is straighforward:

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**

Worst case NON-INVERTING stage: BUFFER

… the **BUFFER** (when G_{loop} is the highest)
↳ G=1





"Compensated" OpAmp (see later)

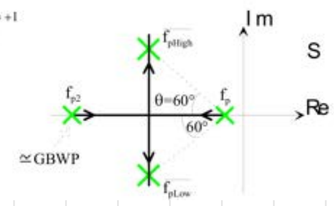
Closure angle <f* 20dB/dec >f* 20dB/dec

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

When f_{p2}=GBWP...

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

... Phase Margin=45°



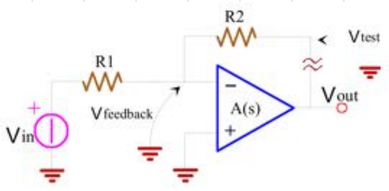
$\xi = 0.5 = \cos\theta = \cos 60° \rightarrow$ for $\theta > 60° \rightarrow$ ε↓ (too much oscillations)

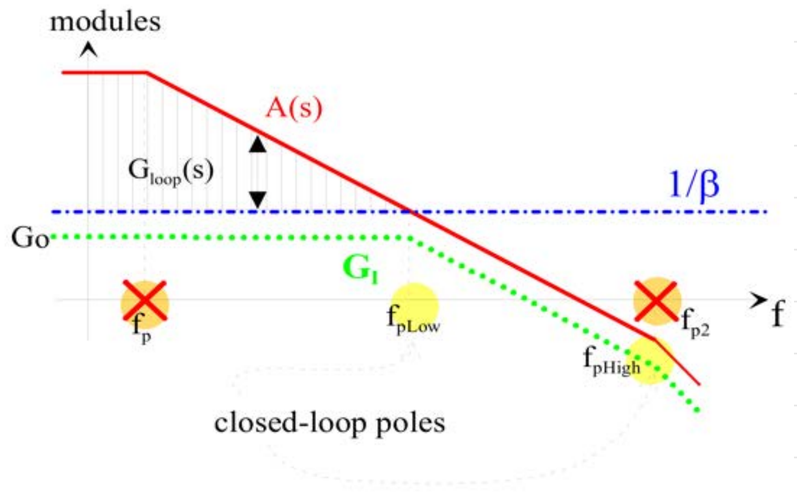
Margine di Fase = 45° = θ

BOUNDARY CONDITION (for reliable stability)

INVERTING vs NON-INVERTING Stage

INVERTING



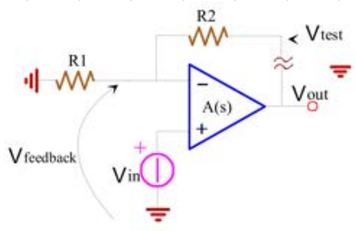


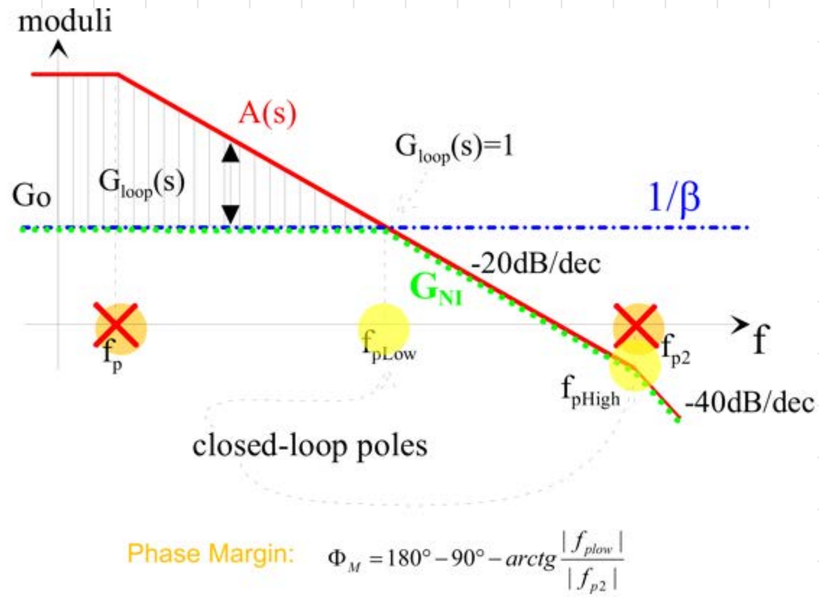
$$\beta = \frac{v_{feedback}}{v_{test}} = \frac{R_1}{R_1 + R_2}$$

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

$$G_{I,real} = \frac{\tilde{A}}{1 + A \cdot \beta} = \frac{-\frac{R_2}{R_1 + R_2} \cdot A}{1 + A \cdot \beta} = \begin{cases} -\frac{R_2}{R_1} = -\frac{1}{\beta} \cdot \left(\frac{R_2}{R_1 + R_2}\right) & \text{if } 1/\beta \ll A \quad \text{(i.e if } |Gloop| \gg 1) \\ -\left(\frac{R_2}{R_1 + R_2}\right) \cdot A & \text{if } 1/\beta \gg A \quad \text{(i.e. if } |Gloop| \ll 1) \end{cases}$$

NON-INVERTING





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

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

Phase Margin: $\Phi_M = 180° - 90° - arctg\frac{|f_{plow}|}{|f_{p2}|}$

$$G_{NI,reale} = \frac{A}{1 - G_{loop}} = \frac{A}{1 + A \cdot \beta} = \begin{cases} 1/\beta & \text{if } 1/\beta \ll A \quad \text{(i.e. if } |Gloop| \gg 1). \\ A & \text{if } 1/\beta \gg A \quad \text{(i.e. if } |Gloop| \ll 1) \end{cases}$$

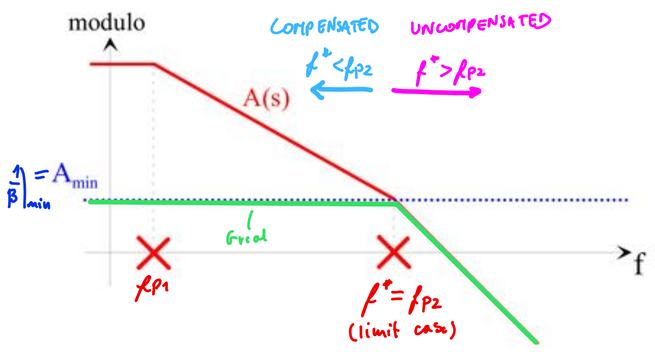
Compensation

As stated before "We say that an OpAmp is compensated if the second pole f_{p2} of it's open-loop gain A(s) is at freq. higher than the intersection of $G_{loop}(s)$ f^*"

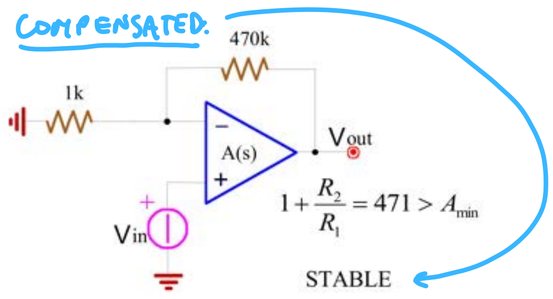
Ex. Compensate vs Uncompensated

There exists a "Minimum Gain" A_{min} and two major poles f_0 and f_1

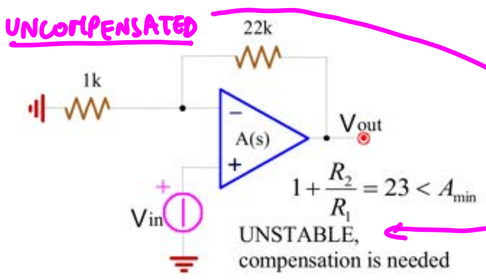
Example
- Compensated:
- OPA 27 SR = 2V/ms
- Uncompensated:
- OPA 37 SR = 17V/ms



COMPENSATED:



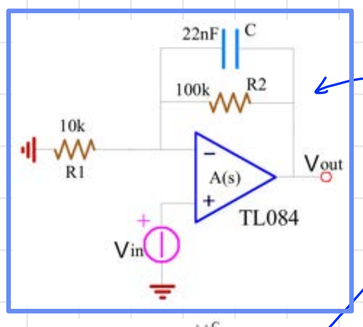
UNCOMPENSATED



← to make it stable we'll have to compensate it

• Effect of feedback capacitance

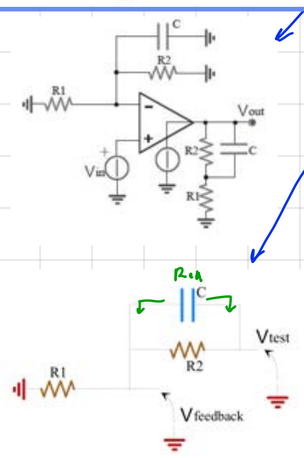
(Book p.224)



Let us study the frequency response of the non-inverting configuration with a capacitor on the feedback network (approximate integrator) which is shown in Fig. 3.26. If we want to calculate the forward block $\tilde{A}$, "turn off" the feedback (Fig. 3.27) and compute:

$$\tilde{A} = \frac{v_{out}}{v_{in}}\Big|_{openloop} = A(s)$$

Note that the output network is not involved in the calculation because it is assumed that the output impedance is zero. To compute the contribution of the feedback block, we proceed with using the circuit shown on the right of Fig. 3.27:



$$\beta = \frac{V_{feedback}}{V_{test}} = \frac{R_1}{R_1 + R_2} \cdot \frac{\left(1 + \frac{s}{|zero|}\right)}{\left(1 + \frac{s}{|pole|}\right)}$$

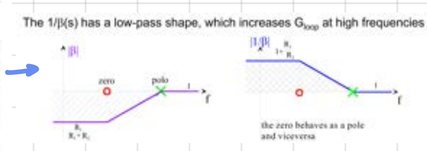
where $1/pole = -C \cdot (R_1 \| R_2)$ and $1/zero = -C \cdot R_2$.

pole: $R_{eq} = R_1 \| R_2$ $\quad pole = \frac{1}{sC(R_1 \| R_2)}$

zero: (AC-shunt along signal path) $\quad zero = \frac{1}{sCR_2}$

Because of the capacitor, β is now not constant but is low (β<1) at low frequencies and increases at high frequencies (β≅1), as shown in Fig. 3.28. Pay attention to the fact that, as shown in Fig. 3.28, in the plot of 1/β(s), the zero of β(s) behaves like a pole (introduces a slope of -20dB/dec), and the pole vice versa (introduces +20dB/dec).

Since in this case $\tilde{A}$ is not different from A(s), the closed-loop transfer function will be virtually identical to 1/β(s) at low frequencies and to A(s) at high frequencies, as shown in Fig. 3.29. If we instead want to use the general method introduced previously, we first plot the ideal gain equal to $1+Z_2/R_1$, where $Z_2 = C//R_2$. In this case, such a gain coincides with 1/β(s) just found, which is 11 at frequencies below the pole (of 1/β, which is the zero of β, which is $1/C \cdot R_2 = 72Hz$) and 1 at frequencies above the zero (when C becomes a short circuit with respect to R_2 and the stage gets buffer connected).



It is interesting to note how at "low frequencies", the ideal gain is equal to:

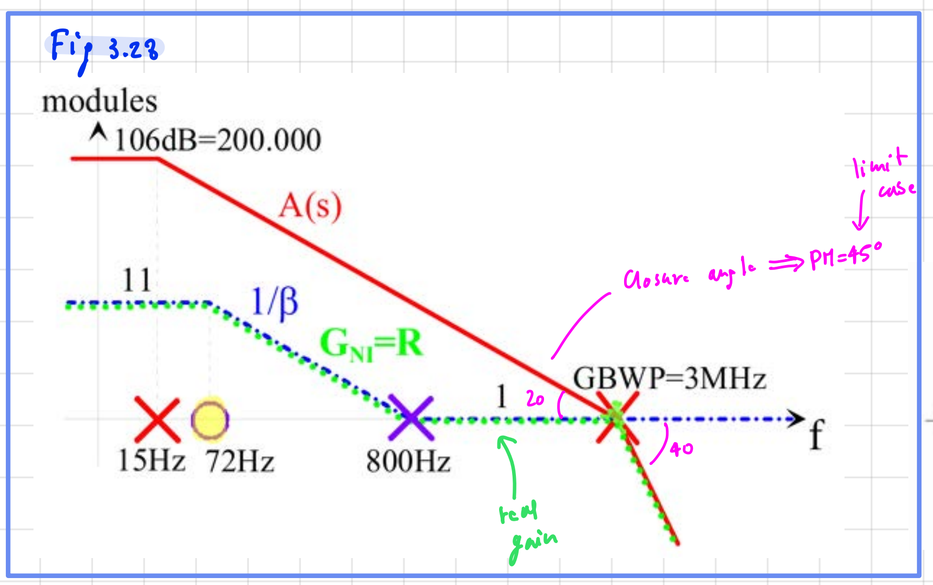
$$G_{NI} \cong \frac{1}{\beta} \cong \left(1 + \frac{R_2}{R_1}\right) \cdot \frac{1}{1 + sR_2C}$$

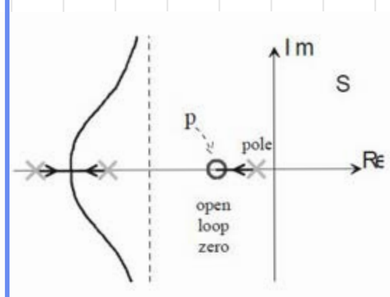
that is, how the closed-loop pole coincides with the zero of the block β. The pole of β, on the other hand, can be graphically obtained from Fig. 3.28 or from Fig. 3.29, knowing that the gain has to decrease from 11 to 1 and that the frequency of the zero is 72Hz. It is found that the pole must be placed at $11 \cdot 72 = 800Hz$. If not, it can be deduced from the electrical analysis of the circuit on the right of Fig. 3.27: $1/C \cdot (R_1 \| R_2) = 800Hz$ is found.

The ideal gain would always be equal to 1 at frequencies above 800Hz. To calculate the real gain, the frequency f* of the intersection of A(s) and 1/β(s) have to be found, which in is equal to the GBWP (3MHz) of the OpAmp this case. This means that the real gain will coincide with the ideal one up to 3MHz. From then on, it will experience the same slope of A(s), i.e. -40dB/dec. Since the closure angle is 20dB/dec before f^* and 40dB/dec after f^*, the phase margin will be equal to 45°, resulting in a slight peaking in the response of the real gain G_{NI} of the non-inverting stage, in correspondence of f^*.

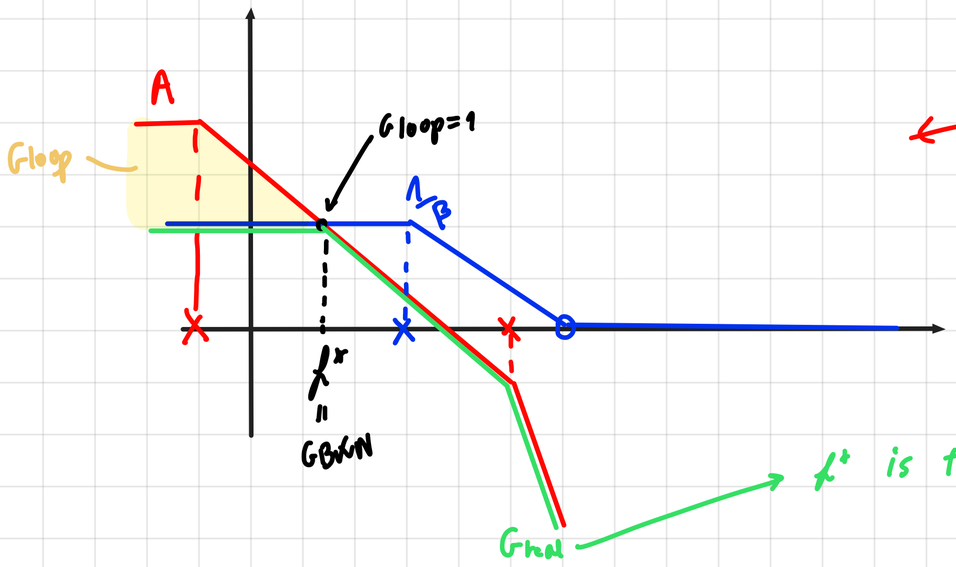
The same result, obviously, would have been obtained with the calibration of the root locus. Since $G_{loop} \gg 1$, on the root locus, the closed-loop pole migrates towards the open-loop zero (Fig. 3.30). It is interesting to note that, for $f \to \infty$ (if G_{loop} does not go to zero, i.e. if the feedback is still working), the gain G_{NI} would not go to zero, but would tend to 1 (buffer). So, there is a zero that is just the pole of β since the capacitor is placed on the feedback branch.

Fig 3.28

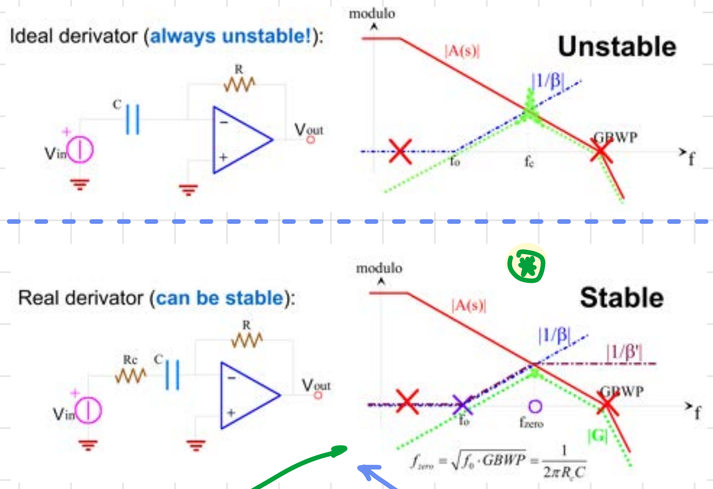




Note: If the sizing is wrong (for ex. we select the wrong OpAmp → not suitable A(s)) we can have this type of situation:

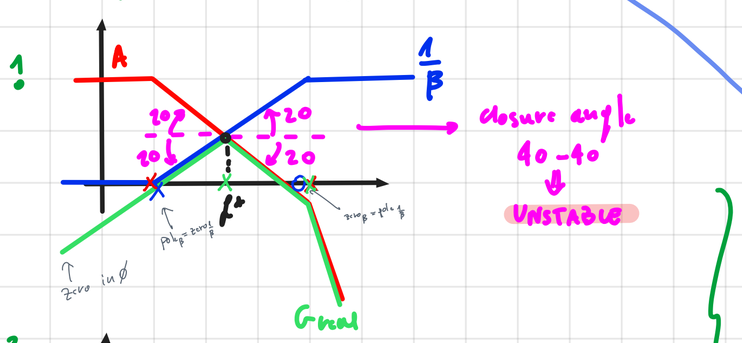


- **Derivator**

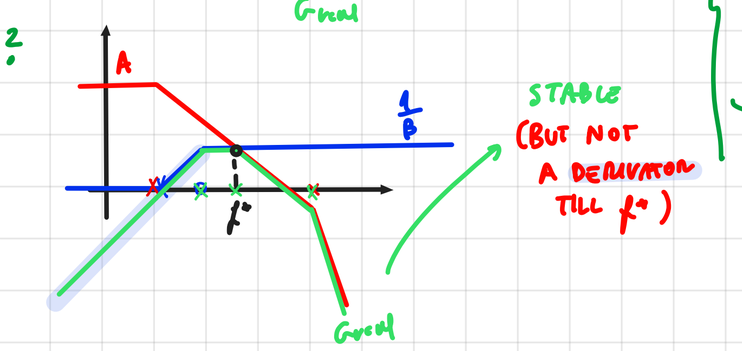


BEST CASE indeed if we analyze the other possible cases:

1.



2.



STABLE (BUT NOT A DERIVATOR TILL f*)

(Book p. 239)

The ideal derivator circuit is shown in Fig. 3.41. The ideal gain is:

$$G_{DER,ideal} = -s \cdot R \cdot C$$

Ideal has a zero in 0

and the 0dB frequency of the derivator is:

$$f_0 = \frac{1}{2\pi \cdot RC}$$

As can be seen from the intersection of A(s) and 1/β(s) at the frequency f*, the stage is definitely unstable. The real gain is equal to the ideal gain up to f*, above which, instead of continuing with +20dB/dec, the real gain suffers the collapse of A(s) (-20dB/dec) and the rise of 1/β(s) (and therefore the collapse of β(s) equal to -20dB/dec) for an overall decrease of +20-20-20=-20dB/dec. When the OpAmp GBWP is reached, the slope increases by another -20dB/dec due to the second pole of the OpAmp, which is located at the GBWP.

In conclusion, the stage acts as a derivator (i.e. with a gain proportional to s) up to f*. To determine this value, just look at the triangle that has been created in the Bode plot between the points f_0, f*, and GBWP. The midpoint, on logarithmic axes, is given by:

$$f^* = \sqrt{f_0 \cdot GBWP}$$

Unfortunately, for the stage is unstable, the Bode plot will have a marked peaking in frequency exactly in correspondence of f*, as shown in Fig. 3.41. A way to compensate the stage would be by introducing a pole in 1/β(s) (i.e. a zero in β(s)) just in correspondence of f*. A circuit that accomplishes this is shown in Fig. 3.42 and is called an approximate derivator. Sizing:

$$f^* = \sqrt{f_0 \cdot GBWP} = \frac{1}{2\pi \cdot R_C C}$$

we come to the plots demonstrated again in Fig. 3.42, on the right. The intersection leads to a phase margin equal to 45°. Now, the ideal gain does not continue with a +20dB/dec slope, but flattens upon reaching the frequency of the pole of 1/β (i.e. zero of β). In fact, at high frequencies, the capacitor is shorted, and the stage gain becomes $-R/R_C$ and not ∞ as in the previous case of Fig. 3.41. The real gain will follow the ideal one up to f*, beyond which, instead of continuing as the ideal one (flat), it will suffer the reduction in A(s) (-20dB/dec), and it will thus get an overall slope equal to 0-20=-20dB/dec.

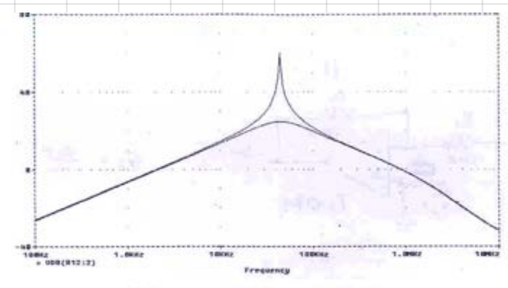
Note that the real gains of Fig. 3.41 and Fig. 3.42 coincide and, in particular, exhibit two poles in correspondence of f* (in fact, the slope changes from +20dB/dec to -20dB/dec). However, in Fig. 3.41, there is the peaking because of the instability of the stage while in Fig. 3.42, the stage is stable, and the poles are quiet (since PM=45°, poles will be complex conjugates with a 60° angle, as discussed above).

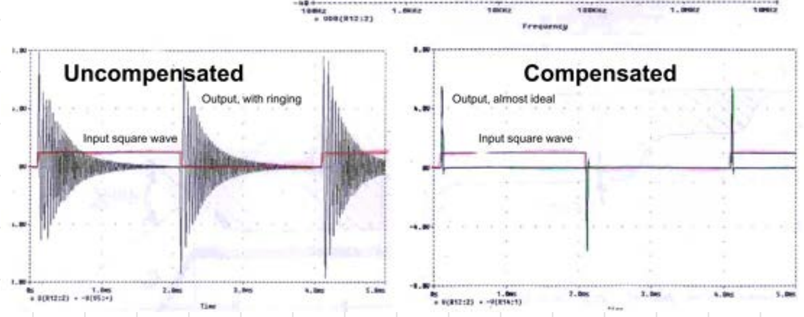
Listed here are the steps to be followed for sizing the derivator, given the bandwidth of interest to obtain (e.g. 0÷f_i) and the frequency f_0 at which the gain should be 0dB:

- choose an OpAmp with GBWP≥100·f_i and place $f^*=f_{zero}\geq 10\cdot f_i$;
- if the source has R_C>100Ω, choose $C=1/2\pi R_C f^*$ (making sure that $C \gg C_{stray}$ of the OpAmp); otherwise choose $C \gg C_{stray}$ and then $R_C=1/2\pi C f^*$;
- finally, obtain $R=1/2\pi C f_0$.

Ex. simulations

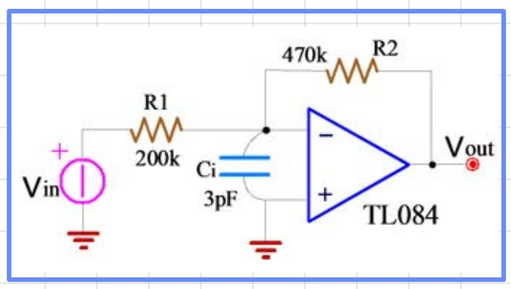
SPICE simulations:
LM101A
f_i = 1kHz
C = 16nF
R = 4.7kΩ





- **Effect of the input capacitance**

(Book p. 241)

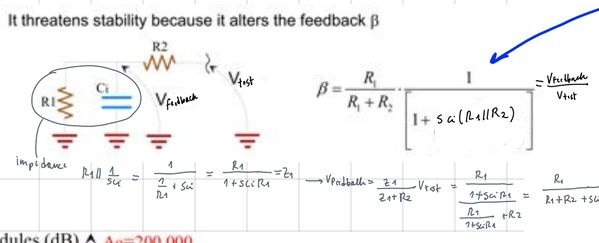


parasitic capacitor

Actually, between the two terminals of an OpAmp, there is a stray capacitor due to, among other things, the connections with the rest of the circuit. The effect of this capacitance can be disastrous if not taken into account in the design stage. Consider the circuit in Fig. 3.46, where this stray capacitance C_i is also shown.

Initially, C_C is not present. Since it is inverting, the circuit will have a real gain G≠1/β. To calculate it, we can proceed with the method that uses $\tilde{A}$, obtaining:

It threatens stability because it alters the feedback β

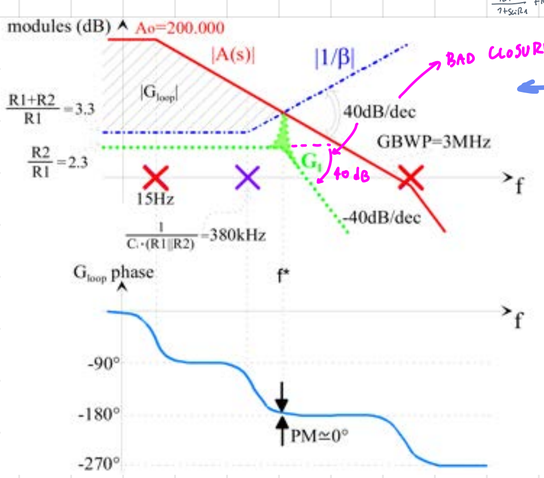


$$\beta = \frac{R_1}{R_1+R_2}\cdot\frac{1}{1+s\cdot C_i\cdot(R_1 \| R_2)} = \frac{V_{feedback}}{V_{test}}$$

$$\tilde{A}(s) = \frac{\left(R_2 \left\| \frac{1}{sC_i}\right.\right)\cdot A(s)}{\left(R_2 \left\| \frac{1}{sC_i}\right.\right)+R_1}$$

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

$$G = \begin{cases} \dfrac{R_2}{R_1} & \text{for } |G_{loop}| \gg 1 \\ \dfrac{A(s)\cdot R_2}{R_2+R_1}\cdot\dfrac{1}{\left(1+\frac{s}{pole}\right)} & \text{for } |G_{loop}| \ll 1 \end{cases}$$



BAD CLOSURE ANGLE → UNSTABLE

The expression just found is shown in the plot of Fig. 3.47, where and 1/β are also reported. The phase plot of G_{loop} is also plotted.

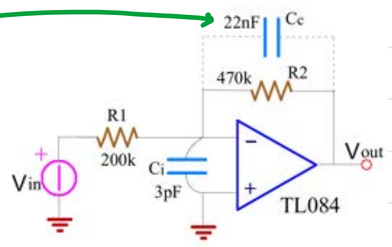
Otherwise, we can proceed with the proposed method, which is plotting only A(s) and 1/β(s). It is easy to see that at low frequencies (C_i open), we have $\beta=R_1/R_1+R_2$ and consequently $1/\beta=1+R_2/R_1$. On the other hand, at the infinite frequency, β=0, then 1/β=∞. The stage can be unstable as in the case of Fig. 3.47. To find the ideal gain, it is enough to note that no signal will flow through C_i since it is connected between the ground and the virtual ground. Then, the ideal gain is independent of its presence and is constant, $-R_2/R_1$.

The real gain will be the same, but only up to f*. After that frequency, the real gain will collapse with a -40dB/dec slope (since both A(s) and β(s) have a -20dB/dec slope). Since the stage is unstable, the real gain will have a peaking at f*, which highlights the presence of two complex conjugated poles.

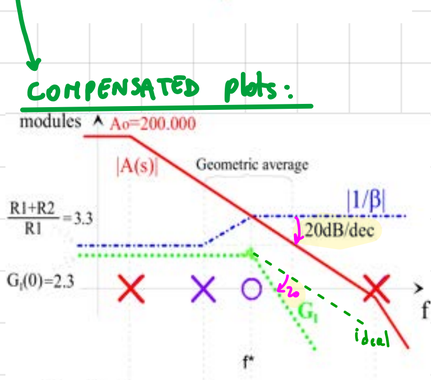
To compensate this circuit, the intersection of A(s) and 1/β has to occur with a 20dB/dec slope, not with a 40dB/dec slope, as it actually happens. At worst, it is possible to make the intersection occur exactly in correspondence of a slope change, i.e. 40dB/dec before and 20dB/dec immediately after, so as to ensure a phase margin equal to 45°. Specifically, it is sufficient to introduce a zero in the β(s), for instance through a capacitance C_C in parallel to R_2, which is midway (geometric mean) between the pole=$1/2\pi C_i(R_1 \| R_2)$ and the frequency f_C=GBWP/(1+R_2/R_1), which is a zero=√(pole·f_C)=585kHz. Since the zero of β(s) is at the pulsation $1/C_C \cdot R_2$, we obtain a value of C_C=0.6pF. The result of this compensation is shown in Fig. 3.48. Note that beyond f*, $G_{I,real}$ drops again with a -40dB/dec slope, but this time due to the simultaneous effect of the zero in β and of the condition $|G_{loop}|$=1. Also note that the introduction of C_C changes the pole, but in a negligible manner since usually $C_C \ll C_i$ holds.

Now it is worth noticing one thing. For stability, we see at which frequency f* we have $|G_{loop}|$=1. Then, we look how far G_{loop} phase is from -180°. One might say that also in Fig. 3.48, the stage is unstable because, before f*, the phase shift has already reached -180°, even with $|G_{loop}| \gg 1$. Well, we should not worry because Bode stability criterion ensures that the only point of interest to be studied is the one where $|G_{loop}|$=1. To be convinced of this, it would be worth deepening your comprehension of the theory of stability by using the Nyquist criterion.

Compensation is needed !

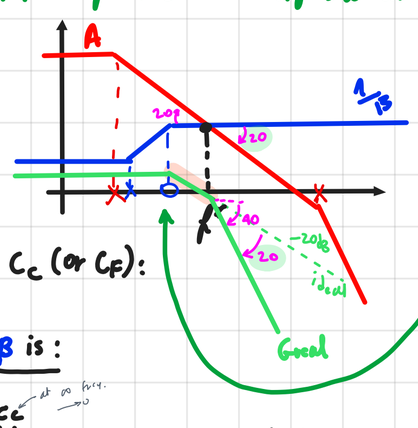


COMPENSATED plots:



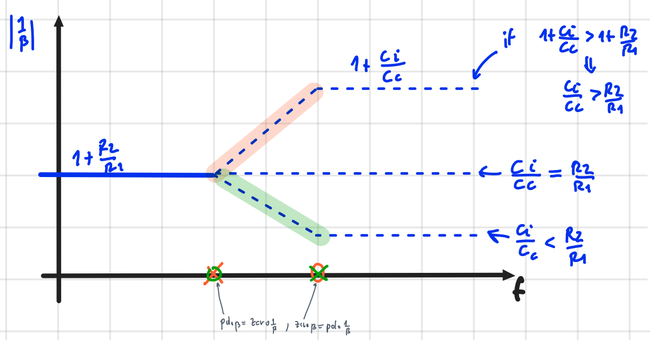
this is the BEST COMPROMISE

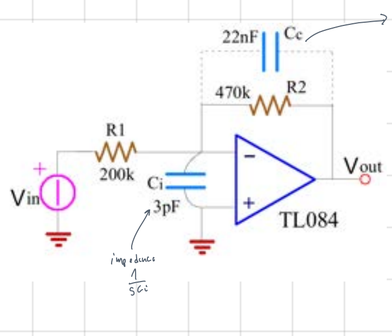
Other possible compensation case:



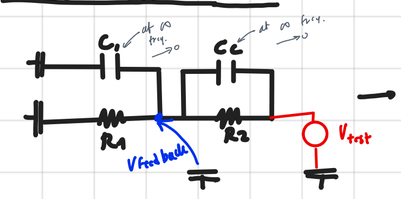
bigger f*, PM=90°, Stable BUT HERE THERE IS A LOW-PASS FILTER ACTION → UNDESIRED

Note For 1/β we can have different characteristic based on the values of C_i and C_C (or C_F):

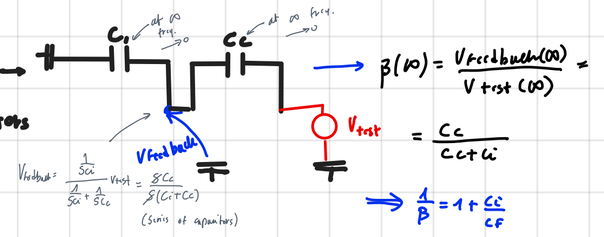




At high freq (→∞) β is:

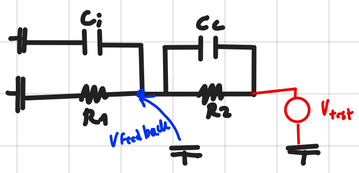


→ R1 and R2 are so high that in the parallel with Cc and Ci the 2 capacitors prevail



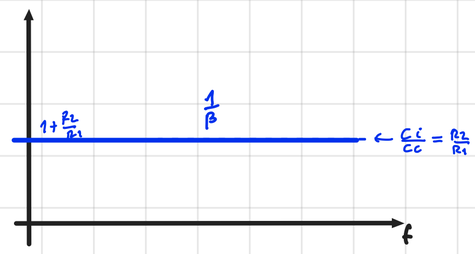
$$\beta(\infty) = \frac{V_{feedback}(\infty)}{V_{test}(\infty)} = \frac{C_C}{C_C+C_i} \Rightarrow \frac{1}{\beta} = 1+\frac{C_i}{C_F}$$

Obs. If we have our circuit:



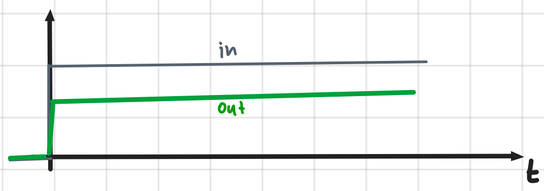
If $R_1 C_i = R_2 C_c \rightarrow \frac{C_i}{C_c} = \frac{R_2}{R_1}$ the circuit is said to be COMPENSATED

↳ meaning the $\frac{1}{\beta}$ Bode diagram is flat



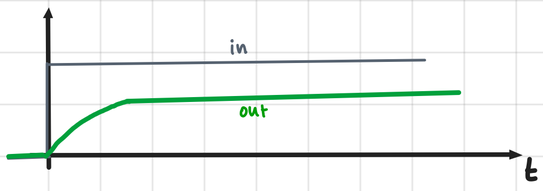
In particular we can see that the step response time depends on the cases:

1. If we don't have C_i and C_c



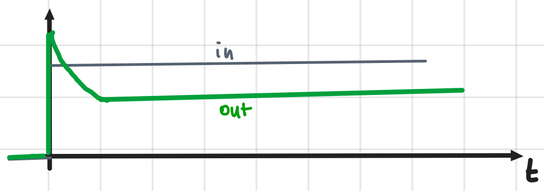
{ NO Cc
NO Ci

2. If we have C_i, but no C_c
   ↳ LOW-PASS FILTER ACTION



{ Ci
NO Cc

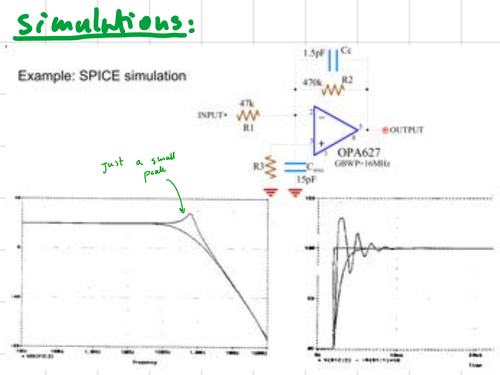
3. If we have C_c, but no C_i
   ↳ HIGH-PASS FILTER ACTION



{ NO Ci
Cc

We recover stability
(as a case with no capacitors is stable)

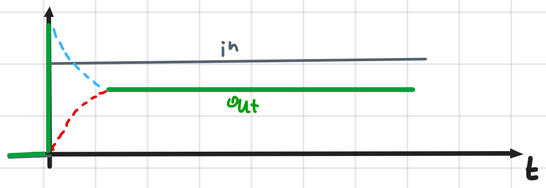
Simulations:



4. If we have C_i and C_c

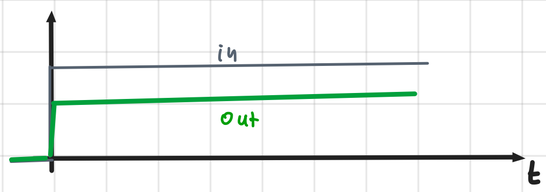
but $R_i C_i \neq R_2 C_c$

$\frac{C_i}{C_c} < \frac{R_2}{R_1}$ $\qquad$ $\frac{C_i}{C_c} > \frac{R_2}{R_1}$



$\begin{cases} \frac{C_i}{C_c} \neq \frac{R_2}{R_1} \end{cases}$

5. If we have C_i and C_c such that $\frac{C_i}{C_c} = \frac{R_2}{R_1}$



$\begin{cases} \frac{C_i}{C_c} = \frac{R_2}{R_1} \end{cases}$ → perfect compensation but high pass and low-pass

- **Transimpedance amplifier**

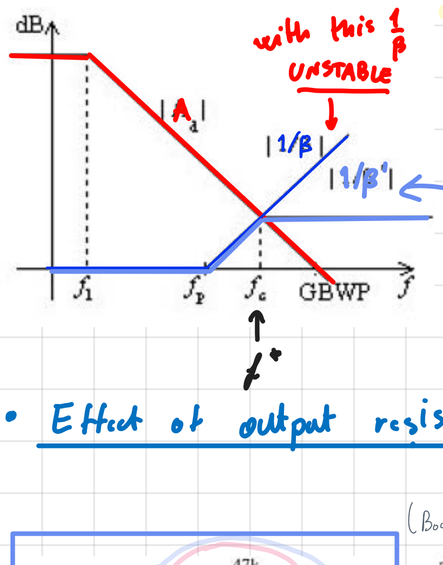
For ex. consider a photodiode amplifier (Book p. 250)

There are various types of amplifiers for photodiodes: some of them read the tension that builds up between their terminals and amplify it whereas others read the currents flowing and transform these currents into tension. The drawback of the former, if realized with active circuitry, is that part of the signal is lost in stray components of the circuit. The quality of the latter, instead, is that the current signal is injected into the virtual ground and therefore is not lost.

In the second case, i.e. current signal reading, we can see how the circuit loop gain has singularities introduced just by the stray capacitance of the photodiode. For this reason, photodiode transimpedance amplifiers can fall into the compensation method for input stray capacitance.

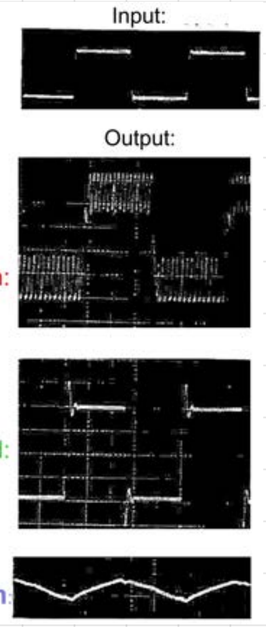
The circuit which we refer to is the one shown in Fig. 3.54. With C_i denoting the capacitance seen at the inverting terminal of the OpAmp, we note that:

$$G_{loop}(s) = -A_d(s) \cdot \frac{1 + sRC_C}{1 + sR(C_D + C_i + C_C)}$$



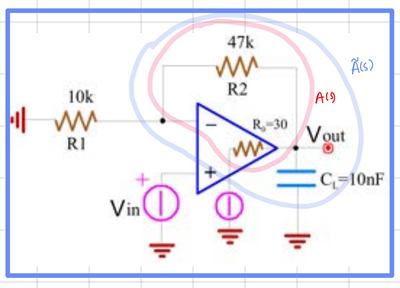
compensation with Cc → to stabilize (PM = 45°)

The effect of non-complete stability of the circuit is visible in the oscilloscope image shown in Fig. 3.56, where the output of a transresistance amplifier with compensated OpAmp can be seen. In the image on the top right, we see a strong ringing while in the figure below (notice the expanded scale) there is even the appearance of oscillations. To avoid this phenomenon, it is necessary to compensate with a feedback capacitor C_F, by paying attention to diligently choose the value of C_F; in this type of compensation a wrong choice of C_F causes an excessive reduction in bandwidth and thus a reduction in the intensity of the output frequency as the frequency of the optical signal modulating frequency increases, as shown in Fig. 3.57.



- **Effect of output resistance and load capacitance**

(Book p. 253)



There are some applications where the loop gain is unstable even if the operational amplifier is internally compensated. Possible causes of instability may be stray capacitors of the circuit. Compensation methods can modify either the forward gain, which is the case of high load capacitance, or the ideal gain, which is the case of high input capacitance. We will now examine circuits where the output impedance is highly capacitive, which can be: sample-and-hold, peak detectors, voltage references, voltage regulators, and amplifiers directly connected to coaxial cables. These circuits can be modeled as in Fig. 3.60, where C_L is the load capacitance.

Refer to Fig. 3.61. It can be noted how in G_{loop} appear A(s) and β(s), but also the additional pole at the following frequency:

$$f_{out} = -\frac{1}{2\pi \cdot C_L \cdot R_0} = 530kHz$$

It is therefore appropriate to introduce a new gain $\tilde{A}^*(s)$ as the cascade of A(s) and the pole:

$$\tilde{A}^* \cong A(s) \cdot \frac{\frac{1}{sC_L}}{R_0 + \frac{1}{sC_L}} = A(s) \cdot \frac{1}{1 + sC_L \cdot R_0} \qquad \text{and} \qquad \beta = \frac{R_1}{R_1 + R_2}$$

where the approximation is due to the fact that the two resistors R_2 and R_1 reconstructed at the OpAmp output have been considered negligible (since they are much higher than R_O). Plotting the graphs (Fig. 3.62), we proceed as usual, but using $\tilde{A}^*(s)$ instead of A(s). We note that the system is unstable since the intersection between $\tilde{A}^*$ and 1/β occurs with a closure angle of 40dB/dec, which implies a phase margin less than or equal to 45°. Precisely:

$$PM = 90° - arctg\frac{f_{pole}}{f_{out}} \cong 33°$$

The frequency of the pole corresponding to the intersection of A* and 1/β is:

$$f^* = \sqrt{f_{out} \cdot \beta \cdot GBWP}$$

One possible way for compensating the amplifier in case of large capacitive loads (e.g. for driving coaxial cables) is by introducing a decoupling resistor R_C and a feedback capacitor C_C. Fig. 3.63 shows this solution. Let us proceed with the calculation of A*(s) and the feedback block. For the former, we consider R_0 and R_C much lower than R_1 and R_2, resulting in:

$$A^* \cong A(s) \cdot \underbrace{\frac{1}{1 + sC_L \cdot (R_0 + R_C)}}_{\text{Effect of } R_0 \text{ e } C_L} \cdot \underbrace{\frac{1}{1 + sC_C \cdot (R_1 \| R_2)}}_{\text{New compensation pole}}$$

For the feedback block β, we find what is shown in Fig. 3.65. We observe that, since two loops are present, it is not trivial to deduce the trend of G_{loop} and that of the real gain through the classical approach (Fig. 3.66).

Let us see how we can proceed with the compensation of this kind of a circuit. A procedure somewhat empirical, often suggested in data-sheets, leads to choose:

$$R_C \cong R_0 \qquad C_C = C_L \cdot \frac{2R_0}{R_2} = 13pF$$

The choice is complicated by the fact that R_0 is not constant, but can vary from 100Ω to 1kΩ at the dc frequency and from 10Ω to 50Ω at high frequencies. It also depends on the instantaneous value of the output voltage V_{out}. Also note that we cannot increase R_C too much, due to the voltage drop across it that could bring the OpAmp out of its admissible voltage swings.

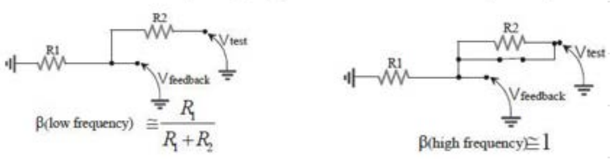
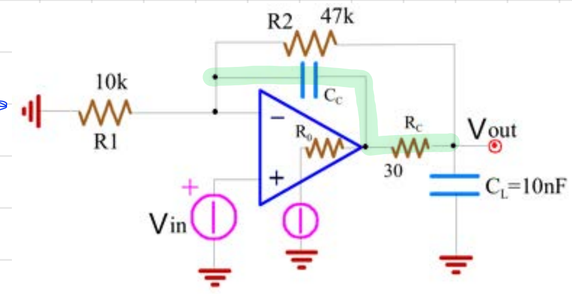
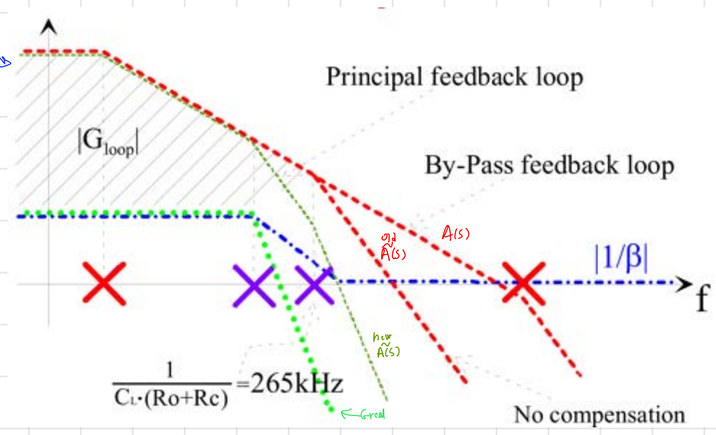


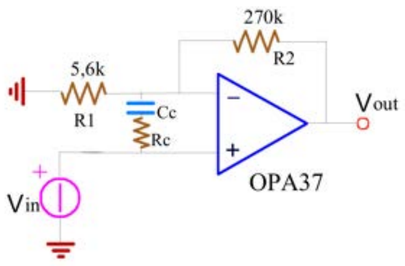
Fig. 3.65: Feedback block at low (on the left) and high frequencies (on the right).



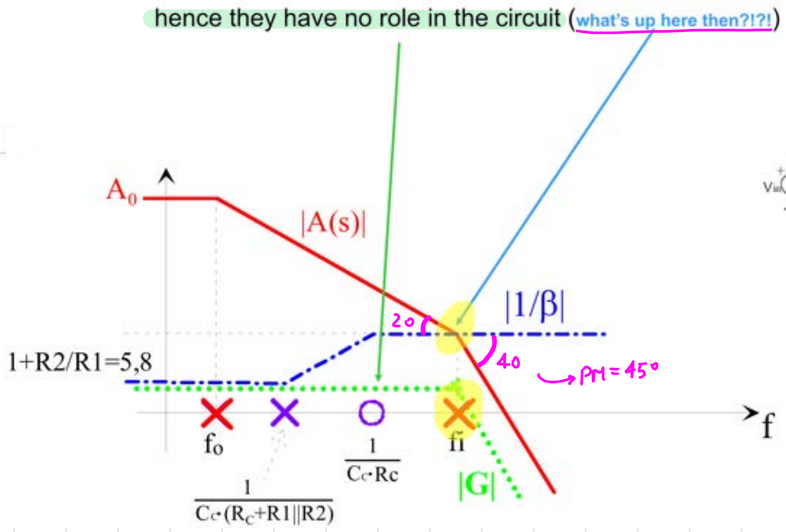


→ Other possible compensation: NEGATIVE FEEDBACK

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



R_c C_c are bootstrapped on the input signal,
hence they have no role in the circuit (what's up here then?!?!)

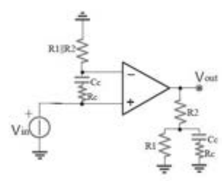


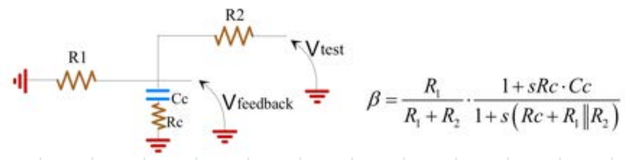
(Book p. 260)

So far, we have tried to compensate by acting on the forward gain to improve the closure angle of the intersection with 1/β. In this section, we will see some circuits in which the trend of 1/β gets modified. We will see that, among other things, this method also allows to achieve an improvement of the circuit slew-rate.

Negative feedback compensation is used to compensate any pole of A(s) or, in general, of the forward block, such as that caused by R_O. The circuit that performs this kind of compensation is shown in Fig. 3.70. For the calculation of the forward block, we refer to the circuit depicted in Fig. 3.71, on the left.

We note that R_C and C_C act at high frequencies, above $\frac{1}{2\pi \cdot C_C \cdot (R_C + R_1 \| R_2)}$ where, however, also β decreases gradually. For the calculation of the feedback block, we use the circuit depicted in Fig. 3.71, on the right. It is found:





$$\beta = \frac{R_1}{R_1 + R_2} \cdot \frac{1 + sRc \cdot Cc}{1 + s\left(Rc + R_1 \| R_2\right)}$$

Note that the series connection of R_C and C_C is bootstrapped on the input signal since the stage reads the voltage on the positive terminal and gives it back the same (at least in the case of an ideal OpAmp) to the inverting one. Then, there will be no current flow in the series connection. Since this network does not affect the real gain, but only the gain β, it will allow compensating the stage without changing its closed-loop gain.

In fact, it is obvious that, given a certain V_{out}, it will be:

$$v_{in} = v^+ = v^- = v_{out} \cdot \frac{R_1}{R_1 + R_2}$$

which yields:

$$\frac{v_{out}}{v_{in}} = \text{constant} = \frac{R_1 + R_2}{R_1}$$

The obtained magnitude Bode plot is shown in Fig. 3.72.

This kind of compensation is suitable for any OpAmp with two poles (even for the non-compensated case) whereas previous procedures required that the OpAmp was either compensated even for unity gain (buffer) or required to have a load capacitor C_L as the integral part of the compensation.

Usually, in non-compensated OpAmps, the value $\mathbf{A_{min}}$ is specified that is the minimum closed-loop gain (non-inverting) which grants stability (Fig. 3.73). In Fig. 3.74 two examples of circuits, stable and unstable, are proposed.

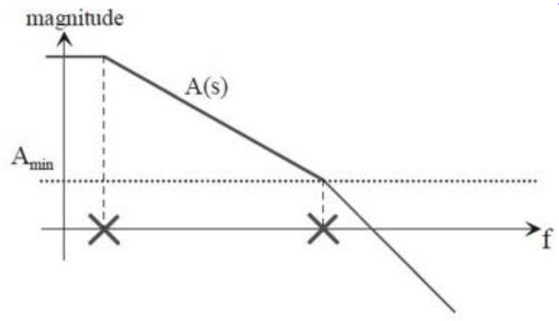
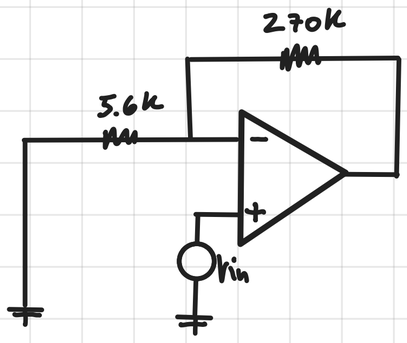
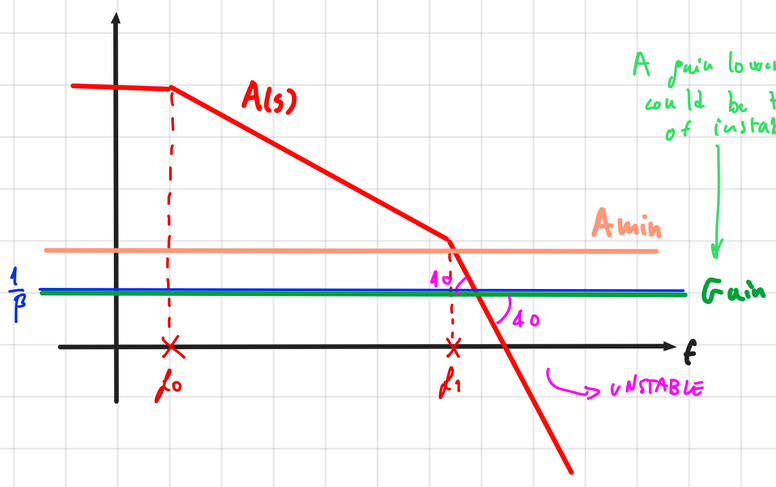


Fig. 3.73: A_{min} represents the minimum gain that can be used without needing any compensation.

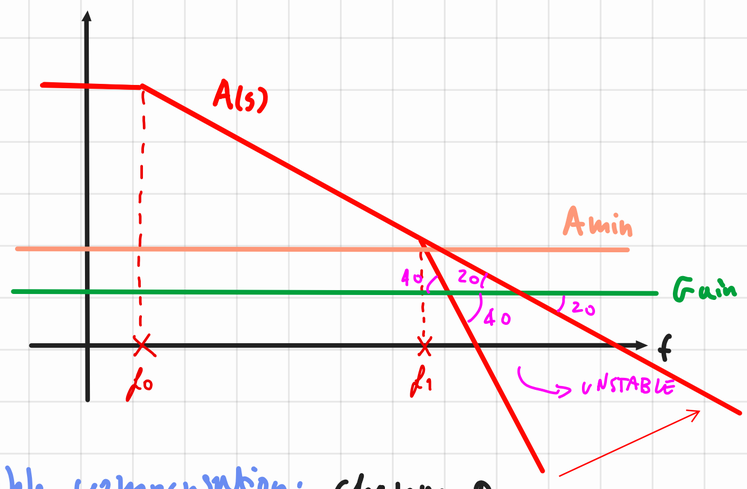
Ex. Suppose we have this circuit



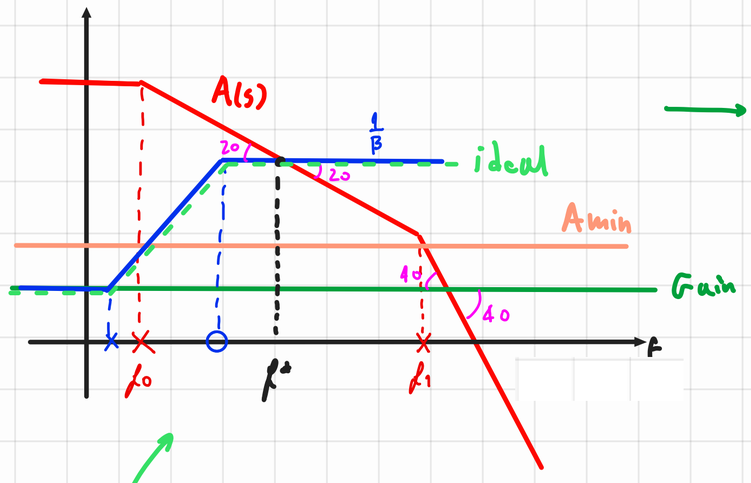


A gain lower than Amin could be the source of instability

→ Possible compensation: change OpAmp → change A(s)



↳ But usually we don't want to buy another OpAmp → other possible compensation: change β

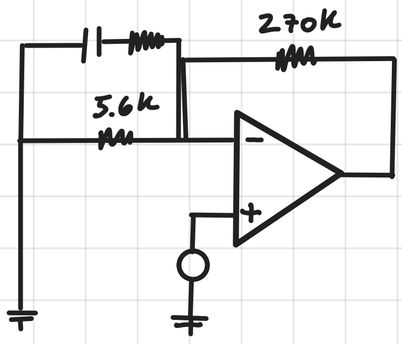


→ The stage it's compensated √

↳ In order to have a pole and zero in $\frac{1}{\beta}$ such that:

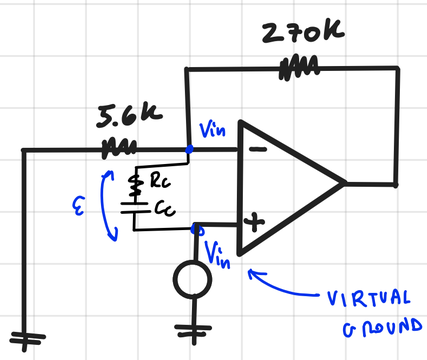
- at high freq. $\frac{1}{\beta}$ high ⇒ β low
- at low freq. $\frac{1}{\beta}$ low ⇒ β high

} use this circuit

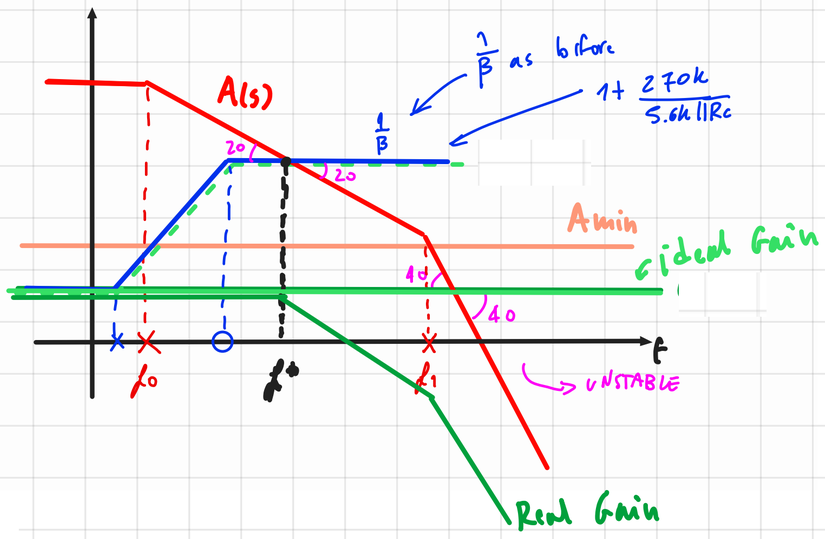


But with this solution we change the gain (real and ideal) that we wanted to remain constant

↳ We can include the β results of the previous compensation by connecting the resistor + capacitor in this way:

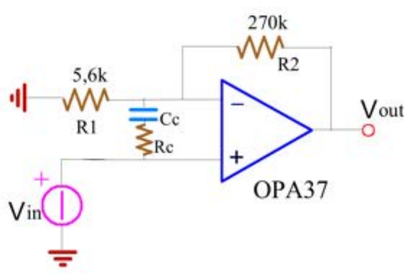


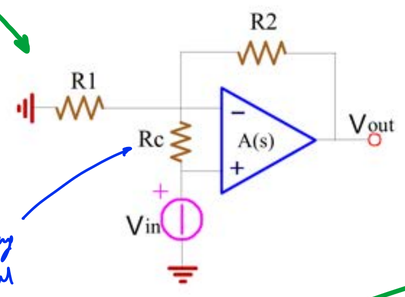
the components introduced don't change the gain



Obs.

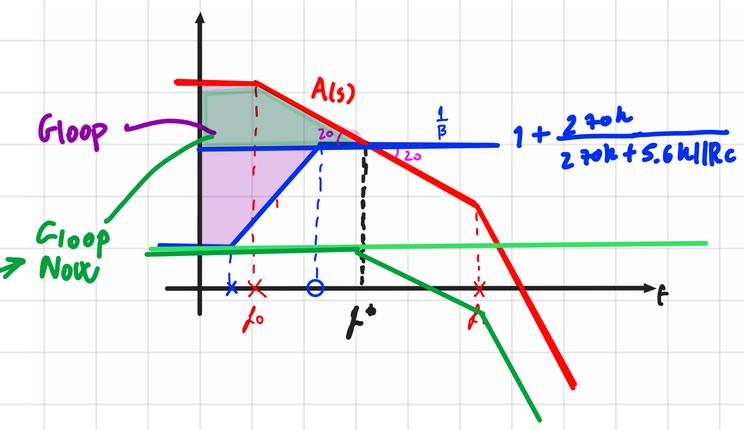
Possible simplification: removal of C_C





not changing the ideal gain

But G_{loop} lowers at all frequencies, down to DC too !



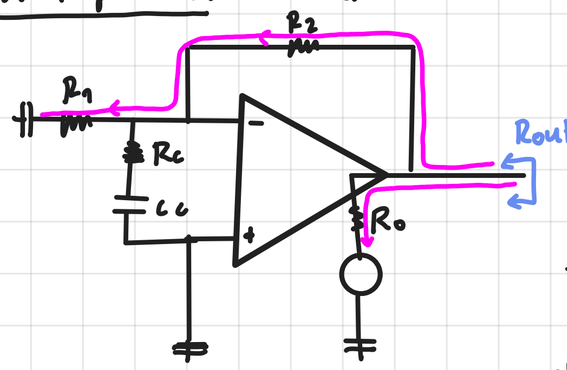
→ Gloop smaller than before

↓

but same results on the gain

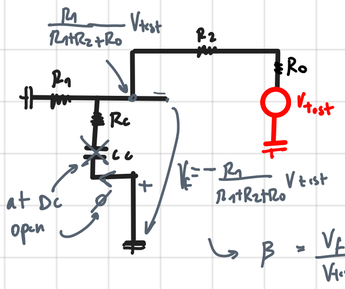
Review of negative feedback:

- **Output impedance:** Consider this circuit:



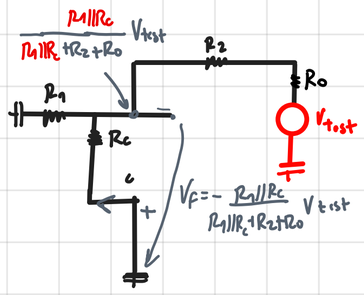
$$R_{out}(0) = \frac{R_{stupid}}{1 - G_{loop}} = \frac{(R_1+R_2)\,\|\,R_o}{1 + A_o \cdot \frac{R_1}{R_1+R_2+R_o}}$$

(DC; $A(0)=A_o$) → impedance drastically decreased if $G_{loop} \gg 1$



$$\beta = \frac{V_f}{V_{test}} = -\frac{R_1}{R_1+R_2+R_o}$$

Note: If we remove C_c then:

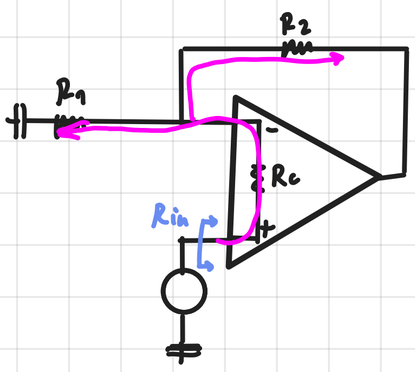


$$\beta = \frac{V_f}{V_{test}} = -\frac{R_1 \| R_c}{(R_1 \| R_c)+R_2+R_o}$$

smaller than the previous one

↳ It's better to keep C_c

- **Input impedance:**

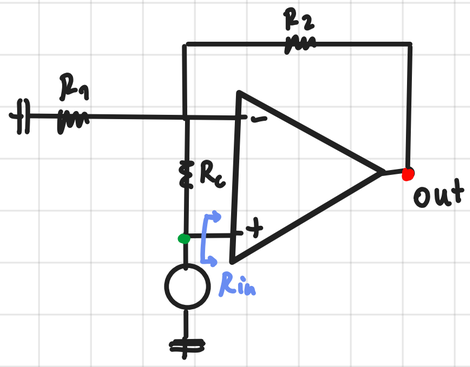


$$R_{in} = R_{stupid}(1 - G_{loop}) = [R_c + (R_1 \| R_2)] + \left(1 + A \cdot \frac{R_1 \| R_c}{R_1 \| R_c + R_2}\right)$$

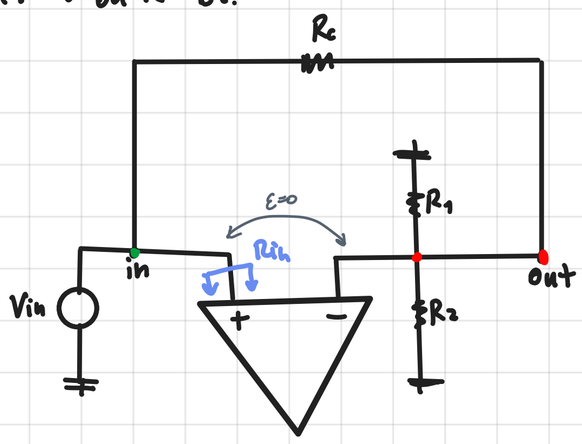
impedance increased by $G_{loop} \gg 1$

same as this for $A \to \infty \Rightarrow R_{in} \to \infty$

↳ Now another way to draw this circuit would be:



which can also be seen as →

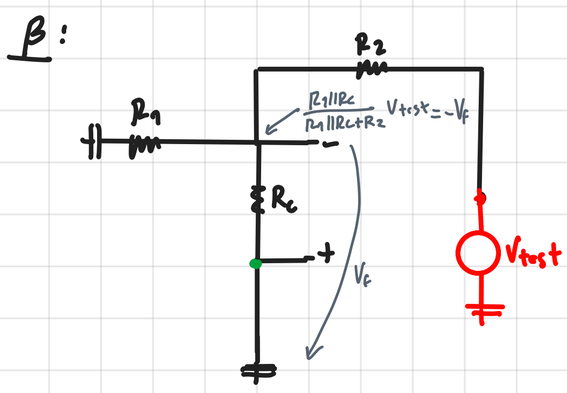


$$R_{in} = \frac{R_{stupid}}{1 - G_{loop}^*} = \frac{R_c + (R_1 \| R_2)}{1 - 1 \cdot 1} \to \infty$$

G_{loop}^* different by looking differently at configuration

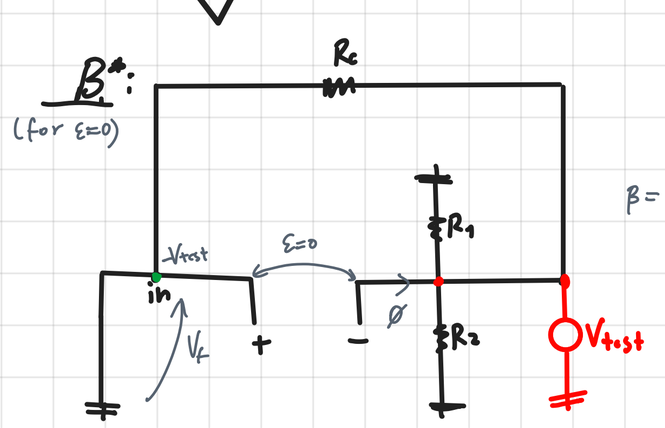
$G_{loop}^* = A \cdot 1$ (for $\varepsilon \to 0$, $A \to \infty$)

β:



$$\beta = -\frac{R_1 \| R_c}{R_1 \| R_c + R_2}$$

β*: (for $\varepsilon = 0$)

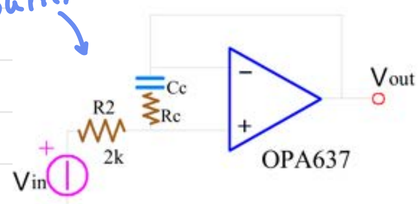


$\beta = -1$

- **Positive feedback compensation**

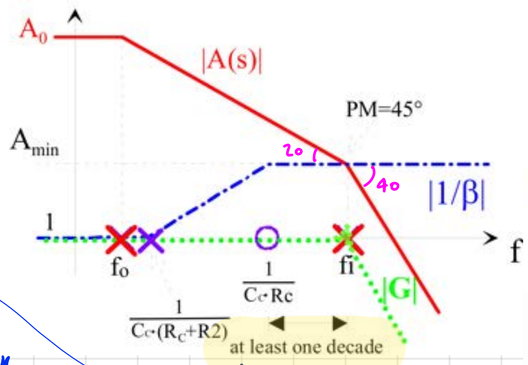
compensated buffer

For buffers, previous approaches cannot be applied... therefore

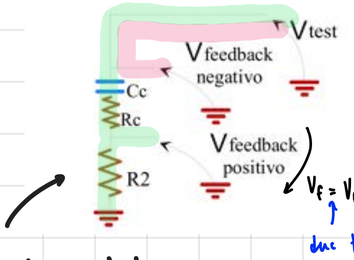


$$R_C = \frac{R_2}{A_{min} - 1}$$

$$C_C = \frac{1}{2\pi \frac{f_i}{10} \cdot Rc}$$



↳ In this way the zero will be able to compensate the PM with full +90°



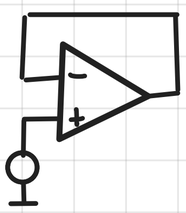
$V_f = V_{f+} - V_{f-}$ ← due to pos. feedback

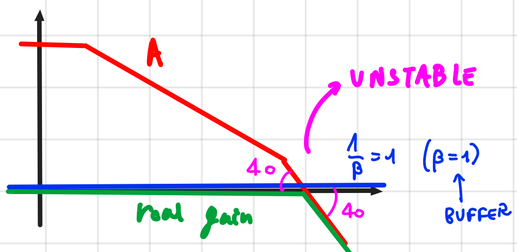
β Computations:

- DC (C open): $\beta(0) = \frac{V_f}{V_{test}} = \frac{V_{f+} - V_{f-}}{V_{test}} = 1$ (no current through R2, $V_{f-} = V_{test}$)

- AC (C close): $\beta(\infty) = \frac{V_f}{V_{test}} = \frac{V_{f+} - V_{f-}}{V_{test}} = \frac{-V_{test} + \frac{R_2}{R_2+R_C} V_{test}}{V_{test}} = -1 + \frac{R_2}{R_2+R_C}$

Buffer behaviour





In the case of a high frequency buffer, the aforementioned compensation could not work. To solve this problem, we introduce a resistor R_2 that provides positive feedback. In practice, it happens that in the magnitude Bode plots, the 1/β starts to rise at lower frequencies. It is important to note that this is the only compensation method for voltage-followers. Let us consider the circuit in Fig. 3.77.

Let us calculate, working as what has already been done up to now, forward and feedback blocks.

The expression found for the forward block is:

$$\tilde{A} = A(s) \cdot \frac{1 + sR_c \cdot C_c}{1 + s(R_c + R_2) \cdot C_c}$$

For the calculation of the feedback block, we use the circuit depicted in Fig. 3.78, on the right. It is found:

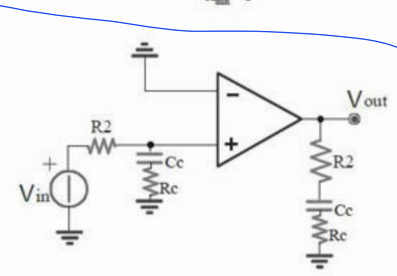
$$\beta_- = 1 \qquad \beta_+ = \frac{R_2}{R_2 + R_C + \frac{1}{sC_C}}$$

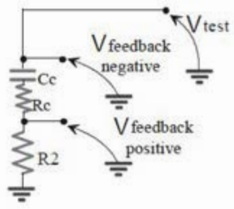
The combination of positive (β_+) and negative feedback (β_-) results in a net contribution:

$$\beta = \beta_- - \beta_+ = \frac{1 + sC_C \cdot R_C}{1 + sC_C \cdot (R_2 + R_C)}$$

As can be seen, there is a pole and a zero. To obtain Fig. 3.79, just choose:

$$R_C = \frac{R_2}{A_{min} - 1} \qquad C_C = \frac{1}{2\pi \frac{f_i}{10} \cdot R_C}$$





Observe that the input compensation network is bootstrapped and that the effect of loading the input node is then reduced. The input impedance is:

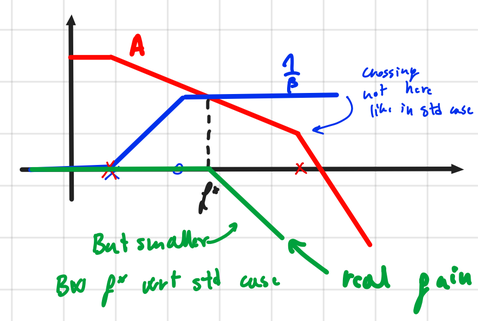
$$Z_{in} = \left(R_C + \frac{1}{sC_C}\right) \cdot A(s)$$ since $G_{loop}(s) = A(s)$ in buffer configuration

while in the previous circuit, it was only:

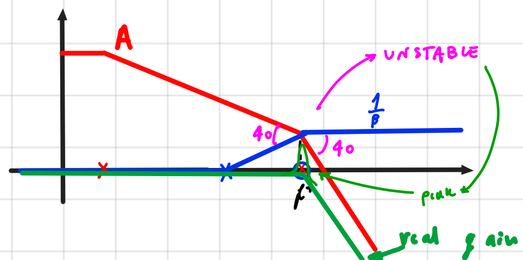
$$Z_{in} = \left(R_C + \frac{1}{sC_C}\right) \cdot G_{loop}(s)$$ where $G_{loop}(s) = A(s) \cdot \beta$ and $\beta < 1$.

The compensation just elaborated has many similarities with the previous one, but adds a new degree of freedom: the choice of R_2 which was previously determined by the desired gain.

Ex. of other compensation:



Ex. of UNSTABLE CASE:



ESOS_NOISE

Electronic Systems

Intro

(New book chapter 1)

In electronics, "noise" is defined as the random fluctuation associated to the deterministic amplitude of every electrical signal. When designing an analog electronic circuit (e.g. an amplifier), it is very important to determine its noise performance as, whenever this quantity became comparable or even higher than the input signal, the information brought by the latter could be corrupted or get completely lost. Noise also impacts other aspects of the electronic system, like the component selection and the signal conditioning partition among blocks.

In general, every electric signal x(t) can be described as the superposition of the useful electric signal, s(t), carrying useful information, the noise n(t), randomly fluctuating in time, plus potential electromagnetic disturbances d(t), caused by interferences with other circuits or devices. So the trend of a generic electric variable (voltage or current) follows this model:

$$x(t) = s(t) + n(t) + d(t)$$

While the source of disturbances, typically being specific circuits or components, can be identified and potentially attenuated by means of shielding and filtering, electronic noise is due to the random statistical movements of the charge carriers in the electronic devices and is therefore ineradicable. Examples of "noisy" electrical waveforms are shown in Fig. 1.1.

analog signal issues (also digital issue with error on timing and threshold)

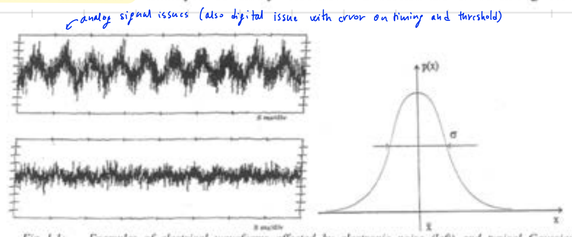


Fig. 1.1: Examples of electrical waveforms affected by electronic noise (left) and typical Gaussian distribution (right).

Being a random variable, in order to study the behavior of electronic noise it is necessary to apply statistical concepts. In most cases, noise comes as the superposition of many elementary (microscopic) processes, each one, generally, described by a Poisson stochastic process. Thus, it is possible to apply the Central Limit Theorem, which implies that the macroscopic process can be described by a Gaussian process, totally characterized by its first- and second-order moments, i.e. mean and variance. By definition, noise must always be characterized by nil mean (otherwise its mean could be considered as an offset):

WHITE NOISE

$$P(x) = \frac{1}{\sqrt{2\pi\sigma_x}} \cdot e^{-\frac{(x)^2}{2\sigma_x^2}}$$

In the case of a Poisson process, only one moment is required to describe it: the first order moment (i.e. the noise variance, or mean square value). The Poisson distribution describes single elementary processes (usually at the microscopic level). The Poisson distribution is:

$$p(n) = \overline{n}^{n} \cdot \frac{e^{-\overline{n}}}{n!} \qquad \text{with n equal to the number of "success"}$$

Similarly to the procedure used for ordinary signals, also noise can be studied in frequency domain: suppose to send a noisy signal n(t) through a bandpass filter, which cuts out any frequency outside an interval with bandwidth Δf centered around a specific frequency f_0, as shown in Fig. 1.2. Let us measure the mean square value of the signal obtained after the filter. Considering a very narrow Δf, the output of the bandpass filter is a sine wave at the frequency f_0: the mean square value is proportional to the square of the amplitude at that harmonic component, i.e. to the electrical power at that frequency. *The ratio between the mean square value and the filter bandwidth Δf is called power spectrum:*

$$S(f_0) = \frac{< n^2 >|_{f_0}}{\Delta f}$$

with units of measurement V^2/Hz or A^2/Hz depending on the nature of the noise (voltage or current).

Depending on the shape of the power spectrum, we can distinguish between white noises (when the spectrum density is constant in frequency), pink/red/brown noises (when the low-frequency components are more intense than those at higher frequencies) and blue/violet noises (when high-frequency components are stronger).

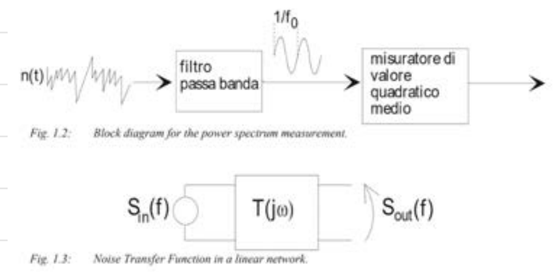


Fig. 1.2: Block diagram for the power spectrum measurement.

Fig. 1.3: Noise Transfer Function in a linear network.

The integral power spectrum is the integral of the mean square values of each harmonic components, representing the noise mean square value, i.e. the noise power:

PARCEVAL TH:

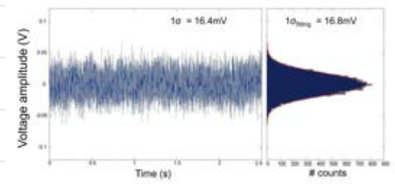
$$\sigma^2 = \int_0^\infty S(f)df$$

Now, let us study the output of an electric network with Transfer Function T(jω) and a noise generator at the input, which is characterized by a power spectrum $S_{in}(f)$ (Fig. 1.3). The linear network transfers at the output all the input harmonic components modifying their amplitude and phase. Thus, a generic harmonic component of the noise at frequency f is transferred to the output with the amplitude modified by |T(jω)|. Thus, if its mean square value at the input is $S_{in}(f)df$, the output has a mean square value equal to:

$$S_{out}(f)df = S_{in}(f) \cdot |T(j \cdot 2\pi \cdot f)|^2 \, df$$

In conclusion, the linear network transfers the power spectrum of the input noise (voltage square or current square) to the output, amplified by the square module of the (voltage or current, respectively) transfer function. Since noise is a stochastic contribution, we are not interested in its phase.

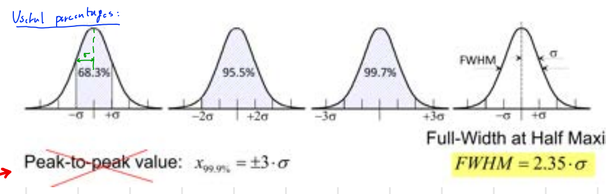
Obs. Distribution



Distribution: GAUSSIAN

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

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



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

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

It does NOT make sense to talk about peak-to-peak value if we have a GAUSSIAN, because we never have ∅ (nil) probability of having a certain value
↳ e.g. for the 0.000...01 % of time we get a value much far from the mean
↳ so we cannot talk about peak-to-peak

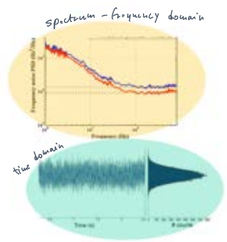
Power

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

just for this process

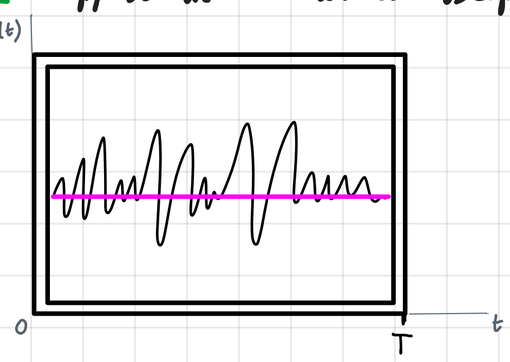
Variance, "power", mean squared value: $< x^2(t) >= \sigma^2$ (POWER = VARIANCE)



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

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

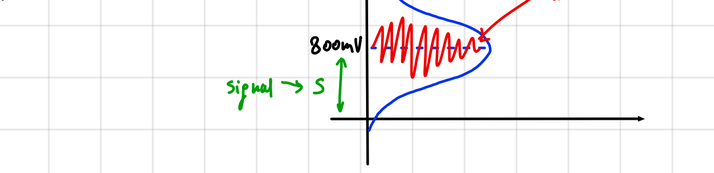
Ex. Suppose we have an oscilloscope:



→ For a noisy signal: $V_{average} = \frac{1}{T}\int_0^T v(t)\,dt = \emptyset$ ← ∅ mean

↳ Even if is a signal with a gaussian centered around a certain value

→ If we compute instead: $< v^2 > = \frac{1}{T}\int_0^T |v^2(t)|dt$ ← VARIANCE (AVERAGE POWER)



Noise correlation

Among different noise sources there can be correlation

Two noise sources: $v_t(t) = v_1(t) + v_2(t)$ (The instantaneous value has no meaning)

Mean total value: $< v_t(t) >= 0$ (consider the power!)

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

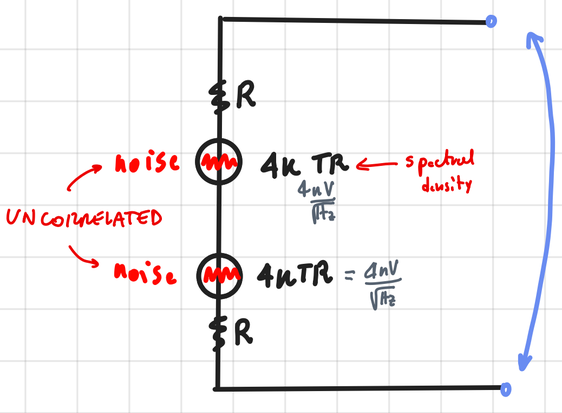
… in case of NO correlation : $< v_t^2(t) >=< v_1^2(t) > + < v_2^2(t) >$ Such as the effect overposition principle

… 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

Ex. Suppose we have 2 noisy components



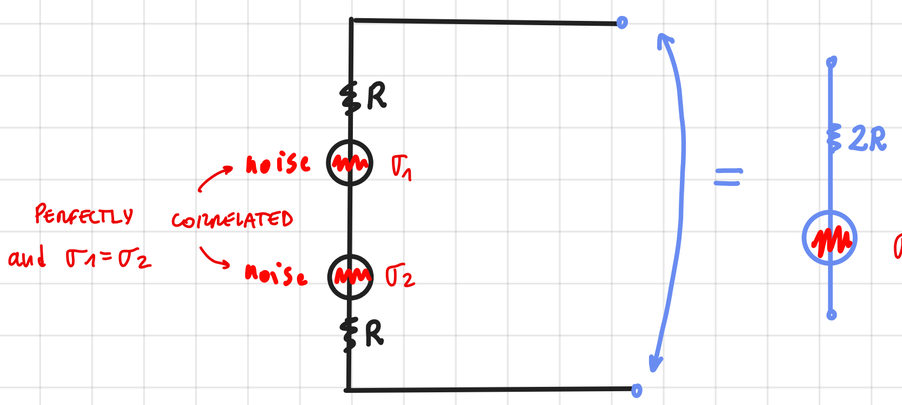
Having $\sigma = \frac{4nV}{\sqrt{Hz}}$ → Variance = σ^2

↳ $\sigma_{TOT} = \sqrt{\sigma^2 + \sigma^2} = \sqrt{\left(\frac{4nV}{\sqrt{Hz}}\right)^2 + \left(\frac{4nV}{\sqrt{Hz}}\right)^2} = 5.6\frac{nV}{\sqrt{Hz}}$

It's not the instantaneous value
↳ It's the variance
↳ power sum up

↳ Each components will have its noise, so we have to know if there's any correlation b/w these noises and take this into account → For simplicity we consider all the noise sources UNCORRELATED

Ex. Perfectly correlated resistors



$\sigma_{TOT} = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2} = \sqrt{4\sigma_1^2} = \sqrt{4}\,\sigma_1$ (2σ1σ2 = CORRELATION)

While in the uncorrelated case even if $\sigma_1 = \sigma_2$ → $\sigma_{TOT} = \sqrt{\sigma_1^2 + \sigma_2^2} = \sqrt{2}\,\sigma_1$

higher of $\sqrt{2} = 1,414$
↳ $\sigma_{TOT} = 1.414\,\sigma_{TOT}$ → 41,4% increment

"negligible" error

for 100% correlation (perfectly correlated)

Types of noise

• SHOT noise ← also called GRANULAR, or Poisson noise

→ GRANULARITY OF CHARGE CROSSING THE JUNCTION

Shot noise is a source of white noise, mainly present in diodes and in BJTs. It is associated to the DC current flowing through the component, and its power spectrum can be approximated as:

$$S(f) \approx S(0) = 2qI = \frac{\sigma^2}{\Delta f}$$

$\frac{signal}{noise} = \frac{I}{\sqrt{2qI\Delta f}} \propto \sqrt{I}$

Thus, a DC current flowing in every P-N junction has an associated noise whose amplitude is proportional to the square root of the current itself, as shown in Fig. 1.4. An important figure to keep in mind is the 4pA/√Hz introduced by a 50μA DC current.

Starting from the variance σ^2 (square value), given the system bandwidth Δf it is easy to infer the root mean square (rms) value:

$$\sigma = \sqrt{2qI_d\Delta f}$$

that is the noise amplitude can be either positive or negative, with an amplitude limited between ±3σ for the 99.7% of the time.

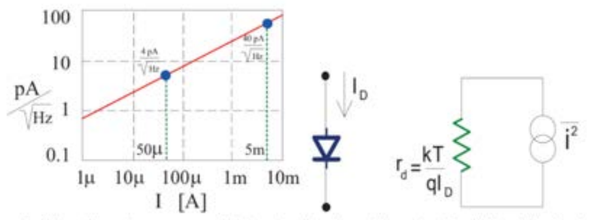
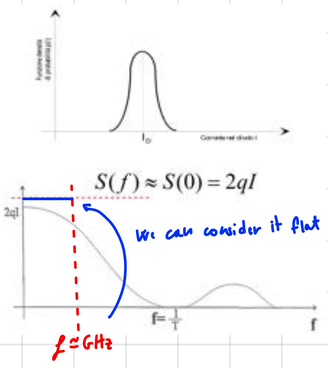


Fig. 1.4: Shot noise versus current I (left) and small signals equivalent circuit of a diode, including the shot noise (right).



SHOT-NOISE summary

- shows up only if there is a current flow
- the rms value increases with $\sqrt{I}$
- depends on f... but... almost "white" noise

$$\sigma^2 = \langle i^2 \rangle = 2 \cdot q \cdot I \cdot \Delta f$$

depend on the noisy component

BW depends on the instrument of measure

• THERMAL noise

→ DUE TO THE GRANULAR AND NEVER FIXED NATURE OF CHARGE

A second source of noise, discovered by John Johnson and often referred to as "thermal" noise, is instead caused by the random thermal motion of electrons within the electronic component (e.g. within a resistor); this random motion doesn't depend on the presence of a DC current, since the typical transport velocity of electrons in a conductor is much lower than the thermal velocity of electrons. Due to its nature, this phenomenon is of course extremely dependent on the absolute temperature, T. As for Poisson noise, thermal noise can be considered white, i.e. its noise spectral density is independent from the frequency. More accurate models, however, show that thermal noise spectral density actually starts to drop at very high frequencies (above about 10THz), outside the frequency range of typical electronic circuits and devices.

It is possible to demonstrate that, in a resistor R, the thermal noise can be modeled by a voltage generator $\langle v^2 \rangle$ in series to its terminal, or –equivalently– by a current generator $\langle i^2 \rangle$ in parallel to the resistor itself, as shown in Fig. 1.5, whose values can be computed as:

$$\langle v^2 \rangle = 4kTR\Delta f \qquad \langle i^2 \rangle = \frac{4kT\Delta f}{R}$$

where k is the Boltzmann constant. An important figure is the noise generated by a 1kΩ resistor at room temperature (293K), which is equal to $v = 4nV/\sqrt{Hz}$.

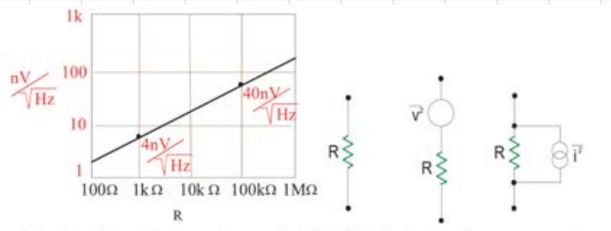


Fig. 1.5: Thermal noise versus resistance value (left) and thermal noise of a resistance representation.

→ Remember that $4kT = 1.66 \cdot 10^{-20} \frac{V^2}{Hz \cdot \Omega}$

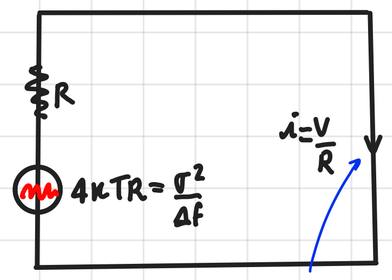
THERMAL-NOISE summary

- independent of current flow
- rms value depends on $\sqrt{T}$ and $\sqrt{R}$
- independent of frequency ("white" noise)

$\langle v^2 \rangle = 4kT \cdot R \cdot \Delta f$ ← Voltage noise increases with R

$\langle i^2 \rangle = \frac{4kT}{R} \cdot \Delta f$ ← current noise decreases with R

Ex.



$i^2 = \frac{4kT}{R}$ ← $i = \frac{\sqrt{4kTR}}{R} = \frac{\sqrt{4kT}}{\sqrt{R}}$

• FLICKER noise

← MANY (UNKNOWN) ORIGINS

Flicker noise is typical of all active devices and also of some discrete passive elements, such as carbon resistors. Its amplitude is very dependent on the nature of the phenomena taking place inside the device: for instance, it has a strong impact on MOSFET transistors, while being almost negligible for BJT transistors, in which it is mainly caused by traps with contaminations and crystalline defects in the emitter-base depletion region, whose energy is concentrated at low frequencies.

Flicker noise can be modeled with a parallel generator and is always associated to a current flow; moreover, differently from the two kinds of electronic noise described up to now, flicker noise has a very distinctive frequency-dependent behavior, shown in Fig. 1.6:

$$\langle i^2 \rangle = h\frac{I^a}{f^b}\Delta f$$

where Δf is a small bandwidth around the frequency f; I is a DC current; a and b are constant values that depend on the device, typically around one (in general, in the range 0.5-2); h is a constant value typical of the device.

Since its spectral distribution is substantially inversely proportional to the frequency (see Fig. 1.7), it is often called "1/f noise" or "pink" noise, and its contribution is particularly important at low frequencies. As shown in Fig. 1.6, its amplitude is often quoted in terms of "noise corner frequency", i.e. the frequency at which the 1/f noise intersects the white noise; above this frequency, the 1/f component can typically be neglected. Depending on the specific device considered, the noise corner frequency could range from few Hz or kHz (typical for BJTs or JFETs) up to MHz or even GHz (typical for MOSFETs). The amplitude distribution of the flicker noise is usually not Gaussian.

Note that if we consider a frequency interval with boundaries f_1 and f_2, the total noise power in such interval only depends on the ratio between f_1 and f_2, and not on their absolute values:

$$P = \int_{f_1}^{f_2} \frac{H}{f} df = H \ln\left(\frac{f_2}{f_1}\right)$$

For instance, the flicker noise power is the same if we consider an interval between 0.01 Hz and 0.1 Hz and a second interval between 100 Hz and 1000 Hz. This is very different with respect to the case of white noise, which instead depends on the absolute width of the considered interval (in our example, the white noise power would be a factor 10^4 higher in the latter case than in the former).

An important observation: mathematically, the 1/f noise would result in an infinite spectral density at 0Hz, thus resulting in infinite noise power, which is absurd. This paradox can be easily solved by considering that the minimum frequency is determined by the observation time. For instance, let us consider an amplifier with 1 kHz bandwidth and assume to observe the behavior for 1 day, i.e. to have an inferior limit set by:

$$f_2 = \frac{1}{24hours \cdot 60min \cdot 60sec} \cong 10^{-5}Hz$$

the observed bandwidth spans across 8 decades. If, instead, we observe it for 100 days the bandwidth would span across 10 decades (since the inferior limit would be 10^{-7} Hz), its spectral density approaching infinity at low frequency, but only giving a power increase of 25% with respect to the former case.

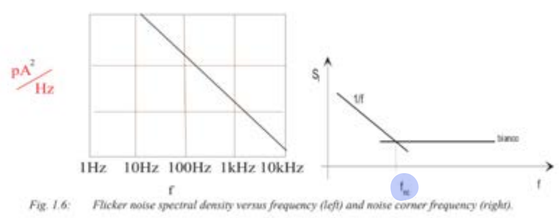


Fig. 1.6: Flicker noise spectral density versus frequency (left) and noise corner frequency (right).

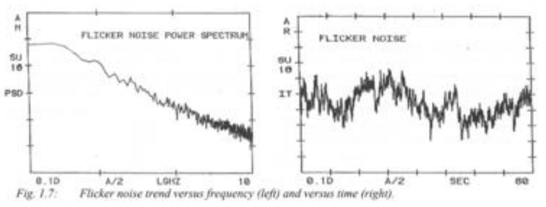


Fig. 1.7: Flicker noise trend versus frequency (left) and versus time (right).

FLICKER-NOISE summary

- proportional to current flow
- power spectrum depends on 1/f
- hence is a "coloured" noise

$$\langle i^2 \rangle = k\frac{I^a}{f^b}\Delta f$$

- constant power for each decade

$$P = \int_{f_1}^{f_2} \frac{k \cdot I}{f} df = k \cdot I \cdot \ln\left(\frac{f_2}{f_1}\right)$$

$P_{one\ decade} = k \cdot I \cdot 2.3$

- does it diverge? $f_{lower} = 0$?
- not for real... $f_{(real\ lower)} = \frac{1}{24h \cdot 60min \cdot 60sec} \cong 10^{-5}Hz$

so 10^{-5}-10^{-4}=10^{-4}-10^{-3}=...=10^{2}-10^{3}= just 8x

- comparable to white noise at the → Noise Corner Frequency

• BURST noise

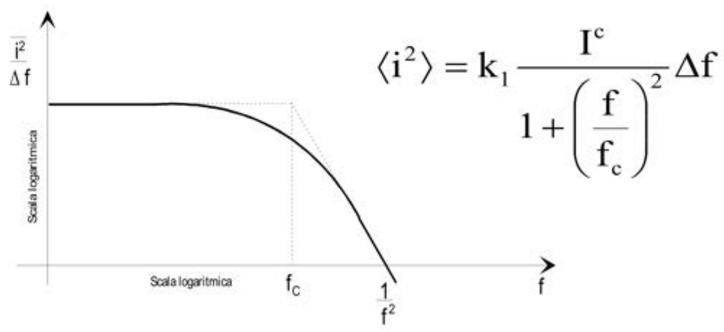
→ DUE TO LATTICE DEFECTS, CHARGE TRAPPING CENTERS...

The burst noise is another type of low-frequency noise typical of some integrated circuits or discrete transistors, which owes its name to the bursts on few discrete levels (two or more levels) that appear when watching this signal at an oscilloscope, as shown in Fig. 1.8 (where some bursts appear superimposed to a seemingly-white noise). The sources of this noise are not completely known yet, although it has been proved that it is associated to the presence of heavy metal ion contaminations and that it is linked to the simultaneous release of many electrons from different traps. In particular, golden-doped devices present a very high burst noise.

The repetition frequency of the bursts is usually in the audio band (some kilohertz or lower), causing a typical popping when transmitted with a speaker. The spectral density of the burst noise is:

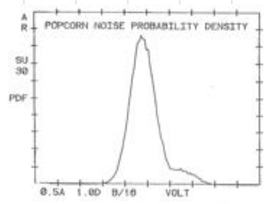
$$\langle i^2 \rangle = k_1 \frac{I^c}{1 + \left(\frac{f}{f_c}\right)^2}\Delta f$$

where k_1 is a constant value distinctive of each device, I is a DC current, c a constant between 0.5 and 2 and f_c is a characteristic frequency associated to a specific noise process. The spectrum is shown in Fig. 1.9; at higher frequencies the noise spectrum decreases as $1/f^2$. The burst noise processes often occur with different time constants, causing the presence of many humps in the spectrum. Furthermore, often the burst noise combines with the flicker noise resulting in the spectrum shown in Fig. 1.9. The amplitude distribution is not a Gaussian. Fig. 1.8 reports the burst noise probability density, highlighting the non-Gaussianity of such process, which results in the superposition of two (or more) Gaussian probability densities.

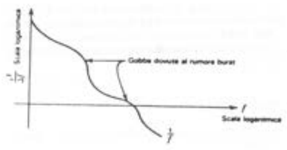


BURST-NOISE summary

- not Gaussian

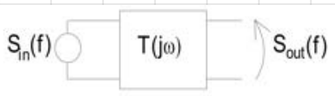


- many different contributions



NOISE in-out TRANSFER

We'll study the TF from an input to an output considering:

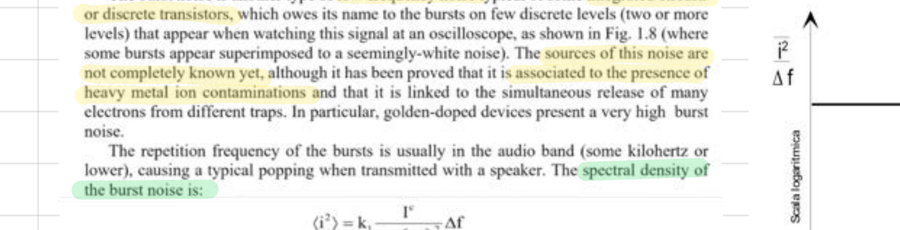


$$S_{out}(f)df = S_{in}(f) \cdot |T(j\omega)|^2 df$$ ← POWER GAIN

Identical Bode diagram, but for noise only "power" transfer matters (independent of phase)

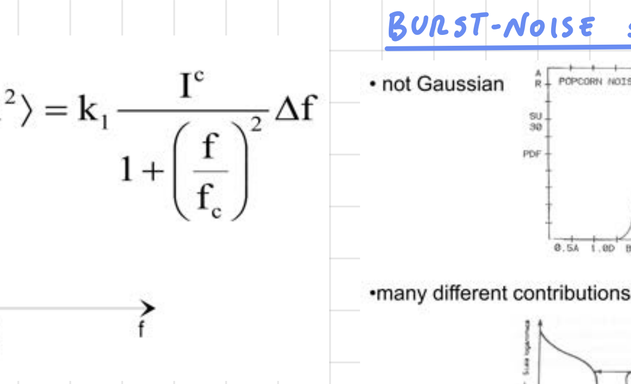
Input noise: $S_i(f) = \frac{\overline{v_i^2}}{\Delta f}$ = S_{io} ← Sin

Output noise: $S_o(f) = \frac{\overline{v_o^2}}{\Delta f}$ = $A_{vo}^2 S_{io}$ ← gain of the amplifier



The total output noise power is $\langle v_{oT}^2 \rangle = \int_0^\infty S_0(f)df = \int_0^\infty |A_v(f)|^2 \cdot S_{io}df = S_{io} \cdot \int_0^\infty |A_v(f)|^2 df$

Ex. Consider a perfectly noiseless amplifier



* R = 1kΩ → $\frac{\sigma^2}{\Delta f} = \frac{4nV}{\sqrt{Hz}}$

↳ we won't have as output noise contribution:

→ We will consider the output noise in terms of power

Just a case it's the same

$\sigma_{out}^2 = \left(\frac{4nV}{\sqrt{Hz}}\right)^2 \cdot (11)^2$ → $\sigma_{out} = \sqrt{\left(\frac{4nV}{\sqrt{Hz}}\right)^2 \cdot (11)^2} = \frac{4nV}{\sqrt{Hz}} \cdot 11$

$S_{out}(f)$ = $S_{in}(f)$ · $|T(jw)|^2$

Fig. 1.10: Circuit for the equivalent noise bandwidth computation.

Fig. 1.11: Linear electric network with its input noise generators.

Note:

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: $f_N = \int_0^\infty \frac{df}{1 + \left(\frac{f}{f_c}\right)^2} = \pi \cdot \frac{f_c}{2} = 1.57 \cdot f_c$

2 poles: $f_N = 1.22 \cdot f_c$

EQUIVALENT NOISE BW

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

Ex. Diff. cases based on # of poles

1 pole: $f_N = \frac{\pi}{2}$ pole

2 poles: f_N = 1.22 pole

3 poles: f_N = ...

4 poles: $f_N \simeq$ pole

↳ If we have multiple poles we can consider the dominant one → the overestimation is negligible → indeed if we consider the worst one $f_N = 1$

NOISE EQUIVALENT BANDWIDTH

Typically, when speaking about noise the most important characteristic is its total power, or rms amplitude.

In principle, it is easy to compute the total noise power by integrating its spectrum density across the bandwidths of interest Δf. However, real circuits don't have a defined bandwidth with ideally-steep limits, but instead complex transfer functions with one or multiple poles. If the equivalent input noise density is **independent from the frequency** (white noise), it is possible to simplify the computations using the concept of noise equivalent bandwidth.

Consider an amplifier with constant noise spectral density $\langle v_i^2 \rangle / \Delta f = S_{ui}$, and a single-pole transfer function with voltage gain $A_v(f)$. The spectral density of the output voltage noise $S_o(f) = \langle v_o^2 \rangle / \Delta f$ is the product of the input voltage noise spectral density and the square of the voltage gain, decreasing by 40dB/dec beyond the pole. The total output voltage noise is obtained by integrating the output spectral density:

$$\langle v_{oT}^2 \rangle = \sum_{f=0}^{\infty} S_0(f)\Delta f = \int_0^\infty S_0(f)df = \int_0^\infty |A_v(f)|^2 S_{ui}df = S_{ui}\int_0^\infty |A_v(f)|^2 df$$

We can express this result as the product of the circuit DC gain A_{vo} and introducing the definition of **noise equivalent bandwidth**, i.e. an equivalent Δf so that $\langle v_{oT}^2 \rangle = S_{ui}A_{vo}^2 f_N$. By comparing the two equations, we obtain:

$$\Delta f = \frac{1}{A_{vo}^2}\int_0^\infty |A_v(f)|^2 df$$

Considering a single pole transfer function, given by:

$$A_v(f) = \frac{A_{vo}}{1 + j\frac{f}{f_{pole}}}$$

where f_{pole} is the -3dB frequency, and solving the integral, we find the following noise bandwidth, which is a general solution valid for every single pole circuit:

$$\Delta f = \int_0^\infty \frac{df}{1 + \left(\frac{f}{f_{pole}}\right)^2} = \pi \cdot \frac{f_{pole}}{2} = 1.57 \cdot f_{pole}$$

Note that this equivalent bandwidth is broader than the -3dB one, by a factor 1.57, *thus a circuit with single pole transfer function causes a noise equivalent to a clean frequency cut at the frequency $\Delta f = 1.57 f_{pole}$*. For circuits with more than one pole, the noise equivalent bandwidth decreases, ideally reaching the -3dB bandwidth when the circuit has an infinite number of poles. For instance a transfer function with two conjugate poles at 45° has a noise bandwidth that is only the 11% broader than the -3dB bandwidth ($\Delta f = 1.11 \cdot f_{pole}$), whereas a circuit with a coincident double pole has an equivalent noise bandwidth equal to $\Delta f = 1.22 \cdot f_{pole}$. *Thus the noise bandwidth doesn't coincide with the amplifier bandwidth, because we are interested in the power transfer function, and not in the voltage one.*

Let's analyze a special case: the simplest noisy circuit, with limited bandwidth, i.e. the resistor with a capacitor in parallel. Consider the resistor equivalent parallel noise shown in in Fig. 1.10 and proceed as above with the normal computations. The transfer functions is:

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

thus the module is: $|T(j\omega)|^2 = \frac{R^2}{1 + \omega^2 C^2 R^2}$

the output spectral density is :

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

the overall output noise voltage (mean square value) is:

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

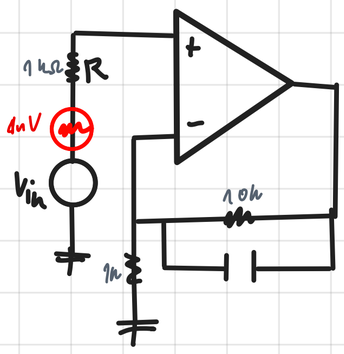
$$= 4kTR \cdot \left[\frac{\text{arctg}(\omega CR)}{2\pi RC}\right]_0^\infty = \frac{4kTR}{2\pi RC} \cdot \left(\frac{\pi}{2}\right) = \frac{kT}{C}$$

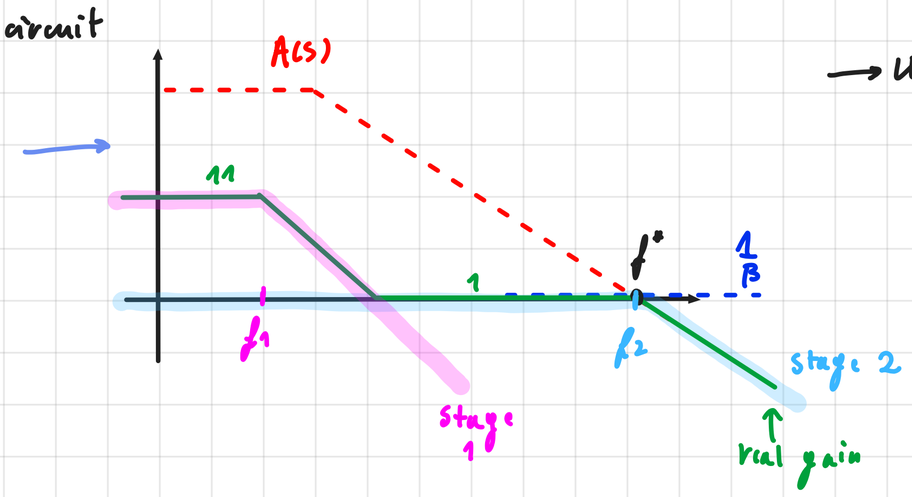
This last equation corresponds to multiplying the 4kTR spectral noise by a term

$$\Delta f = \frac{\pi}{2} \cdot \frac{1}{2\pi RC} = \frac{1}{4RC}; \qquad \langle v_{0T}^2 \rangle = 4kTR\Delta f = 4kTR\frac{1}{4RC} = \frac{kT}{C}$$

The output noise is therefore independent from the resistance, thus the overall output noise voltage only depends on capacitance and temperature, despite the resistor being the noisy component. The noise bandwidth of the circuit in Fig. 1.10 is equal to $f_n = 1.57 \cdot 15000 = 23562Hz$, considering that the output voltage loses 3dB at 15 kHz.

Ex. Consider this circuit





→ We can consider the real gain as 2 stages in this way their plots will resemble the simple plots seen before

↳ $\sigma_{out} = \sqrt{\left(\frac{4nV}{\sqrt{Hz}}\right)^2 \cdot 11^2 \cdot \frac{\pi}{2} f_1 + \left(\frac{4nV}{\sqrt{Hz}}\right)^2 \cdot 1^2 \cdot \frac{\pi}{2} f_2}$

INPUT

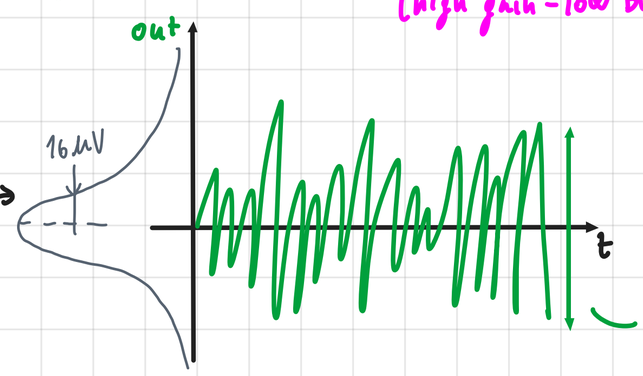
first contribution (high gain - low BW)

second contribution (low gain - high BW)

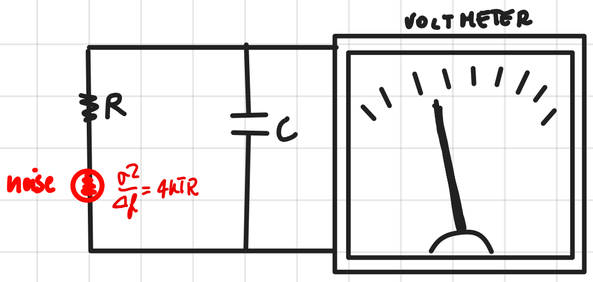
depending on the amplification or the BW one of the 2 can prevail

↳ for ex. $f_1 = 1kHz$ $f_2 = 1MHz$

↳ $\sigma_{out} = \sqrt{(4n \cdot 436)^2 + (4n \cdot 3962)^2} \simeq 16\,\mu V_{rms}$ ← this prevail



Ex. Resistor ∥ Capacitor:



→ $rms = \sigma = \sqrt{\sigma^2} = \sqrt{4kTR \cdot f_N} = \sqrt{4kTR \cdot \frac{\pi}{2} \cdot \frac{1}{2\pi RC}} = \sqrt{\frac{kT}{C}}$

1 pole given by capacitor: $pole = \frac{1}{2\pi RC}$

→ the output noise will depend on C, although the noisy component was the R

NOISE EQUIVALENT GENERATORS

Since every active electronic component introduces its own sources of white and non-white noise, when dealing with complex circuits it is often useful to model the noise of each single sub-circuit (i.e. of every generic linear circuit) with a couple of "equivalent" noise generators, one for modelling current noise sources and one for voltage ones, as shown Fig. 1.11. If our circuit has a generic input resistance R_{in}, thanks to the Thevenin theorem we can easily derive the power spectrum of the voltage generator, $S_v(f)$ by shorting out the network input terminals and measuring the noise spectrum at the output terminals $S^0(f)$. Being $T(j\omega)$ the transfer function v_{out}/v_{in}, the output noise spectrum is due only to $S_v(f)$ and it is equal to $S_{out}(f) = S_v(f) \cdot |T(j\omega)|^2$. Thus, the value of the equivalent input noise voltage generator is given by the equation $S_v(f) \cdot |T(j\omega)|^2 = S^0(f)$.

Consequently: $S_v(f) = \frac{S^0(f)}{|T(j\omega)|^2}$ ← for R=0

$S_v(f)$ this is the equivalent power spectrum which, applied to the circuit input and in absence of other noise sources, causes the same output noise spectrum when the input terminals are shorted.

A similar definition applies to the equivalent current noise generator. In this case, the noise spectrum at the real network output, $S^{ii}(f)$, is measured when the input terminals are open. The transfer function v_{out}/i_{in} is $R_{in}T(j\omega)$, thus the output noise spectrum of the equivalent network is $S_{out}(f) = S_i(f) R_{in}^2 |T(j\omega)|^2$; consequently, we find:

$S_i(f) = \frac{S^{ii}(f)}{R_{in}^2 |T(j\omega)|^2}$ ← for R = ∞

The definition of equivalent input noise generators allows to compute the effects caused by all the noise sources introduced by each electronic component. In the general case, the network can be sourced by a generator with generic impedance R_S, as in the case of Fig. 1.12. The noise spectrum at the output terminals can be evaluated referring to Fig. 1.12 bottom, in which both noise equivalent generators have their own transfer function to the output, giving two separate noise contributions which must be quadratically summed to compute the total output noise power:

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

Observing this equation, we note that with low source resistances, $R_S \ll R_{in}$, the output noises is dominated by the voltage generator, whereas for $R_S \gg R_{in}$, the input current noise dominates. *Thus, when selecting the source impedance* (which of course gives its own contribution to the total output noise), *we must pay attention to the values of the two noise generators.*

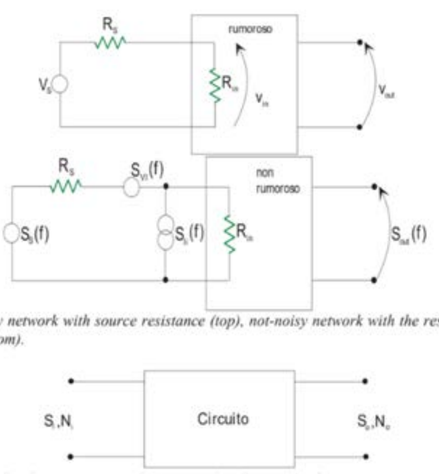


Fig. 1.12: Noisy network with source resistance (top), not-noisy network with the respective noise sources (bottom).

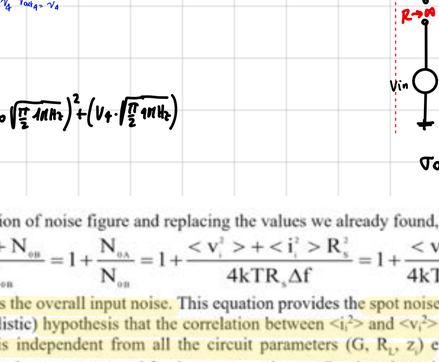


Fig. 1.13: Signal and noise power at the input and at the output of a circuit.

NOISE FIGURE NF

In order to characterize the noise performance of a two-pole network, whose generic representation is shown in Fig. 1.13, we can define a "noise factor" F, which expresses the degradation of the signal-to-noise ratio introduced by the network:

$$F = \frac{S/N \text{ input ratio}}{S/N \text{ output ratio}}$$

Even thought this method is only limited to cases where **the input source is resistive**, its exploitation is a very powerful tool, especially for telecommunications systems.

Referring to Fig. 1.13, where S represents the signal power and N the noise power, we have as only source of input noise (N_i) the source resistor R_S. At the output, instead, we find a noise power N_0 given by all circuit components and by the source resistor. We obtain:

$$F = \frac{S_i}{N_i} \cdot \frac{N_0}{S_0}$$

Considering an ideal noise-less amplifier, the output noise is only due to the input source resistor, and thus, if the circuit power gain is G, the output signal S_0 and the output noise N_0 are given by $S_0 = GS_i$ and $N_0 = GN_i$. Replacing these equations in the definition of noise figure, we obtain in this case F=1, i.e. 0 dB.

Another useful definition of F comes from the original one:

$$F = \frac{S_i}{N_i} \cdot \frac{N_0}{S_0} = \frac{N_0}{GN_i}$$

that can be written as: $F = \frac{\text{overall output noise}}{\text{portion of the output noise due to the source resistor}}$

Another figure of merit often used to characterize the noise performance of a network is the "noise figure" NF, defined in the same way as the noise factor but expressed in dB: $NF = 10 \cdot \log_{10} F$. The noise figure is usually defined for a small bandwidth Δf around a frequency f with $\Delta f \ll f$ (Δf =1 Hz at the test frequency commonly used of 1kHz). We can exploit the equivalent input noise generators, as in Fig. 1.14, to evaluate the effects of each circuit parameter on the noise figure: a generic circuit, with input impedance z_i and voltage gain $G = v_o/v_x$, is driven by a source resistor R_S and feeds a load R_L. The source resistor introduces a thermal noise $\langle i_s^2 \rangle$, while the circuit noise is represented by its equivalent (uncorrelated) generators $\langle i_i^2 \rangle$ and $\langle v_i^2 \rangle$. In order to compare the noise introduced by the circuit to that of the source resistor, we can consider the Thevenin equivalent at the input, at the left of the proper circuit (i.e. **without** considering Z_i).

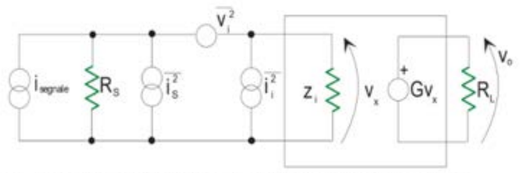


Fig. 1.14: Representation of the equivalent input noise, for the noise figure computation.

Using the definition of noise figure and replacing the values we already found, we obtain:

$$F = \frac{N_{oA} + N_{oB}}{N_{oB}} = 1 + \frac{N_{oA}}{N_{oB}} = 1 + \frac{\langle v_i^2 \rangle + \langle i_i^2 \rangle R_s^2}{4kTR_s\Delta f} = 1 + \frac{\langle v_{rms\,in}^2 \rangle}{4kTR_s\Delta f}$$

where $\langle v_{rms\,in}^2 \rangle$ is the overall input noise. This equation provides the spot noise factor, in the (pretty much realistic) hypothesis that the correlation between $\langle i_i^2 \rangle$ and $\langle v_i^2 \rangle$ is negligible. Observe that F is independent from all the circuit parameters (G, R_L, z_i) except for the equivalent input noise generators and for the source resistance R_S: therefore, *the noise factor of different circuits must always be compared at the same R_S.* Moreover, *for small values of R_S the noise is dominated by the $\langle v_i^2 \rangle$ generator, whereas the $\langle i_i^2 \rangle$ generator is dominant for large R_S;* in the between, there is an "optimal" value of R_S which minimizes the noise factor. This optimal resistance can be computed by differentiating the previous equation:

$$\frac{dF}{dR_s} = 0 \Rightarrow -\frac{4kT\Delta f \langle v_i^2 \rangle}{(4kTR_s\Delta f)^2} + \frac{\langle i_i^2 \rangle}{4kT\Delta f} = 0 \Rightarrow -\langle v_i^2 \rangle + \langle i_i^2 \rangle R_s^2 = 0 \Rightarrow R_{s,ott}^2 = \frac{\langle v_i^2 \rangle}{\langle i_i^2 \rangle}$$

The optimal noise figure is:

$$NF_{opt} = 10\log_{10}\left(1 + \frac{\langle v_i^2 \rangle}{4kT\Delta f\sqrt{\frac{\langle v_i^2 \rangle}{\langle i_i^2 \rangle}}} + \frac{\langle i_i^2 \rangle \sqrt{\frac{\langle v_i^2 \rangle}{\langle i_i^2 \rangle}}}{4kT\Delta f}\right) = 10\log_{10}\left(1 + \frac{\sqrt{\langle v_i^2 \rangle \cdot \langle i_i^2 \rangle}}{2kT\Delta f}\right)$$

thus the smaller the product $\langle i_i^2 \rangle \cdot \langle v_i^2 \rangle$, the smaller the NF_{opt}.

Let us suppose to have the two trends shown in Fig. 1.15, associated to two different devices A and B. From a superficial analysis of the two graphs, the device A could appear to be less noisy than B, since its NF is lower. Nevertheless, this comparison strongly depends on the source resistance value: indeed, in case the source resistance is equal to R_S', the device B is much better than A! Therefore, comparing two devices only basing on their NF_{opt} doesn't make any sense, as its circuit must be evaluated in its actual operating conditions. *In case the source resistance is large ($R_S > R_{S,opt}$) and a selection has to be performed between two amplifiers with different equivalent noise generators, the choice should fall on the one with the smaller parallel generator. The opposite applies in case of small source resistance ($R_S < R_{S,opt}$).* Finally, it is possible to compute the total input noise voltage for a generic two-port network (see Fig. 1.14) as:

$$\frac{v_{s\,rms\,totale}}{\Delta f} = \sqrt{\frac{\langle v_s^2 \rangle}{\Delta f} + \frac{\langle i_s^2 \rangle}{\Delta f} \cdot R_s^2 + 4kTR_s}$$

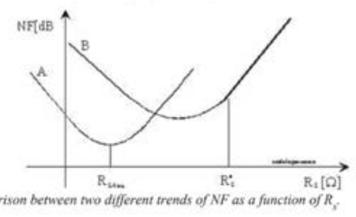
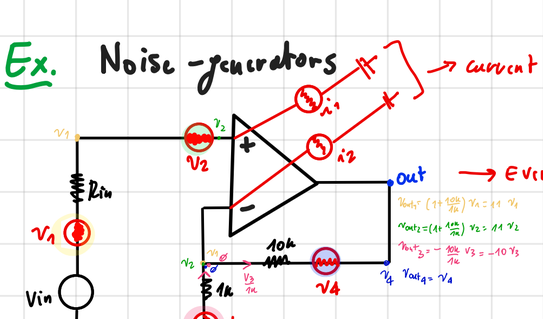


Fig. 1.15: Comparison between two different trends of NF as a function of R_s.

Ex. Noise-generators



→ every component has noise → many noise gen.

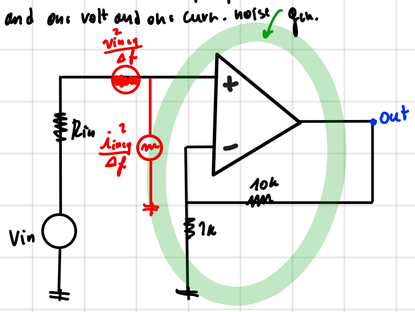
$BW = 1MHz$ → $\sigma_{out} = rms_{out} = \sqrt{(v_4 \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1M})^2 + (v_2 \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1M})^2 + (v_3 \cdot 10 \cdot \sqrt{\frac{\pi}{2} 1M})^2 + (v_4 \cdot \sqrt{\frac{\pi}{2} 1M})^2}$

Note
- If we change Rin → v_1 contribution changes
- If we change 10k → v_4 contrib. changes
- If we change 1k → v_3 contrib. changes

→ Difficult to analyze and control all these contributions

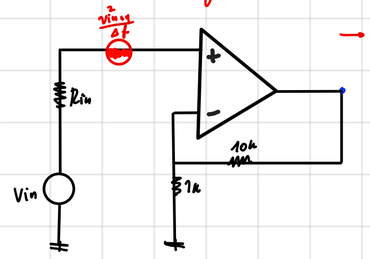
↳ It's better to consider noise equivalent sources!

We consider the OpAmp noiseless and one volt and one curr. noise gen.



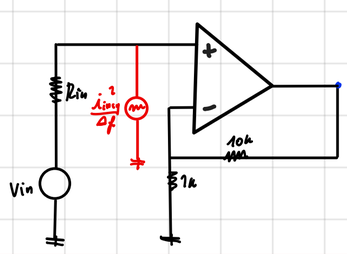
→ We have to consider both the Volt and the curr. equiv. gen., indeed:

- just Volt. noise gen.:



→ If Rin → ∞ (open) ↳ the noise $\frac{\overline{v_{ineq}^2}}{\Delta f}$ would have NO EFFECT

- just current noise gen.:

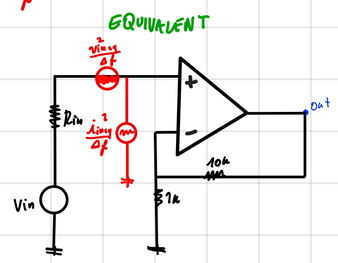


→ If Rin → 0 (short) ↳ no current → noise $\frac{\overline{i_{ineq}^2}}{\Delta f}$ has NO EFFECT

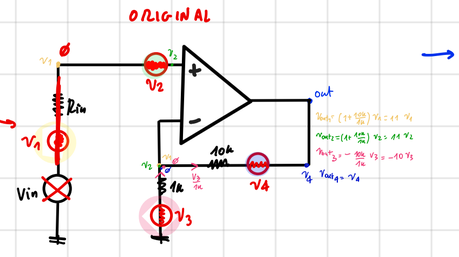
↳ How to compute the current and voltage generators?

- $\frac{\overline{v_{ineq}^2}}{\Delta f}$ (Volt. gen.): Put Rin = 0 (short) → curr. gen. has no effect → all volt. gen.

EQUIVALENT



ORIGINAL



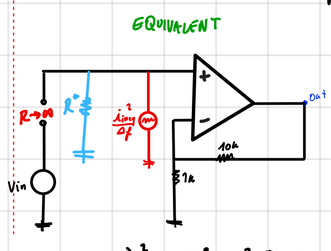
$\overline{\sigma_{outrms}^2}_{eq} = \frac{\overline{\sigma_{outrms}^2}_{ideal}}{(11 \cdot \sqrt{\frac{\pi}{2} 1M})^2} = \frac{\overline{v_{ineq}^2}}{\Delta f}$

↳ $Out_{rms\,eq} = \frac{\overline{v_{ineq}^2}}{\Delta f} \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1MHz}^2 \equiv Out_{rms\,original} = (v_2 \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1MHz})^2 + (v_3 \cdot 10 \sqrt{\frac{\pi}{2} 1MHz})^2 + (v_4 \cdot \sqrt{\frac{\pi}{2} 1MHz})^2$

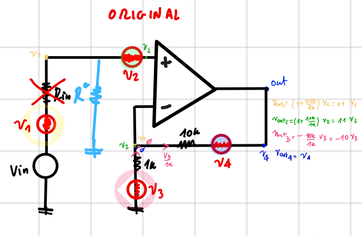
- $\frac{\overline{i_{ineq}^2}}{\Delta f}$ (curr. gen.): If we put R → ∞ → no volt. gen BUT introduces ∞V → ∞ effect of curr noise gen.

↳ We put instead a high value R^*

EQUIVALENT



ORIGINAL



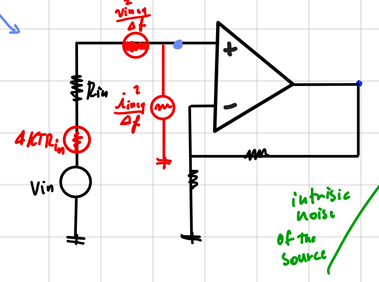
$\overline{\sigma_{outrms}^2}_{ideal} = \frac{\overline{i_{ineq}^2}}{\Delta f} \cdot R^{*2} \cdot 11^2 \cdot \frac{\pi}{2} 1MHz \simeq Out_{rms\,original} = (v_2 \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1MHz})^2 + (v_3 \cdot 10 \sqrt{\frac{\pi}{2} 1MHz})^2 + (v_4 \cdot \sqrt{\frac{\pi}{2} 1MHz})^2$

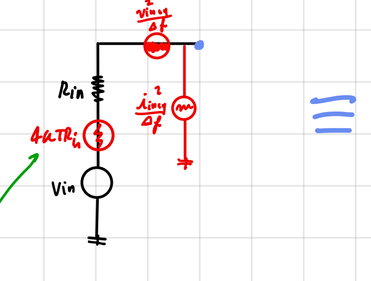
→ $\overline{\sigma_{outrms}^2}_{ideal} = \frac{\overline{i_{ineq}^2}}{\Delta f} = \frac{\overline{\sigma_{outrms}^2}_{original}}{(R^* \cdot 11 \cdot \sqrt{\frac{\pi}{2} 1M})^2} \xrightarrow{R^* \to \infty} 0$

in this ex. no effect for curr. eq. gen.

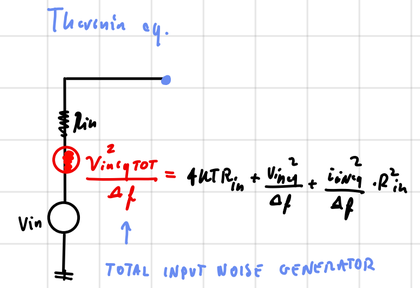
could be diff. cause of R^*

Ex. Thevenin eq. generator and its NF





=



IDEALLY (only source)

$\frac{\overline{v_{ineqTOT}^2}}{\Delta f} = 4kTR_{in}$

ACTUALLY

$\frac{\overline{v_{ineqTOT}^2}}{\Delta f} = 4kTR_{in} + \frac{\overline{v_{ineq}^2}}{\Delta f} + \frac{\overline{i_{ineq}^2}}{\Delta f} \cdot R_{in}^2$

↳ $NF = 10\log\left[\frac{\text{tot. noise of the real circuit}}{\text{tot. noise of just source}}\right] = \left[\frac{4kTR_{in} + \frac{\overline{v_{ineq}^2}}{\Delta f} + \frac{\overline{i_{ineq}^2}}{\Delta f} \cdot R_{in}^2}{4kTR_{in}}\right] = 10\log\left[1 + \frac{\frac{\overline{v_{ineq}^2}}{\Delta f}}{4kTR_{in}} + \frac{\frac{\overline{i_{ineq}^2}}{\Delta f} \cdot R_{in}}{4kT}\right]$

Ex.

Let us now make a practical example: consider the operational amplifier LM394, whose noise trends are represented in Fig. 1.16, with bi-logarithmic axes. The NF (highlighted area in the figure) is the difference between overall input noise and source resistor noise (in log scale). From this figure we can observe that R_{Sopt}, i.e. the point at which the equivalent voltage and current generators have the same value, is equal to 15 kΩ. The total noise coincides with the OpAmp equivalent voltage noise for low source resistances, being then dominated by the source resistor noise beyond 200 Ω and by the equivalent current generator beyond 10MΩ .

Given a voltage noise generator $\langle v_i^2\rangle=(500nV)^2$ and a current noise generator $\langle i_i^2\rangle=(50pA)^2$, we find that for 2 kHz bandwidth and 293 K temperature, the optimum source resistance is equal to $R_{Sopt}=\sqrt{(500nV)^2/(50pA)^2}=10k\Omega$, whereas the optimum noise figure is:

$$NF_{opt}=10\log\left(1+\frac{\sqrt{(500nV)^2\cdot(50pA)^2}}{2\cdot1.38\cdot10^{-23}\cdot300\cdot2000}\right)=4dB$$

For matched circuits (source resistance equal to Thevenin input resistance), the smallest possible NF is 3dB, unless particular techniques are used for further noise suppression; the NF for a practical low-noise circuit can range between one and few dBs.

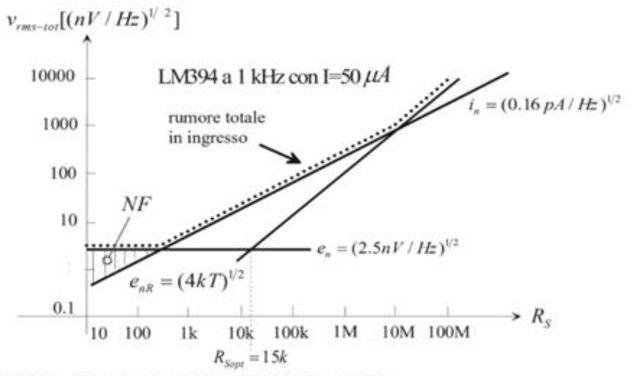


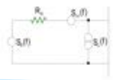
Fig. 1.16: Equivalent input voltage noise of the OpAmp LM394.

NF RECAP

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

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

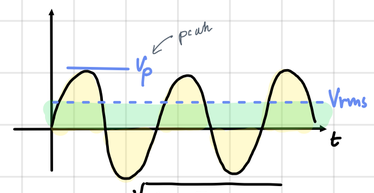
- (Power ratio)

$SNR=10\log_{10}\frac{(V_{pin}/\sqrt{2})^2}{V_{rms}^2}$ c.f. $=10\log_{10}\frac{(100mV/\sqrt{2})^2}{(3mV_{rms})^2}=27dB$

- (rms ratio) ← we can also compute SNR from the rms bringing the power 2 in front of the expression

$SNR=20\log_{10}\frac{V_{pin}/\sqrt{2}}{V_{rms}}=27dB$

Ex. Consider a sinusoidal signal

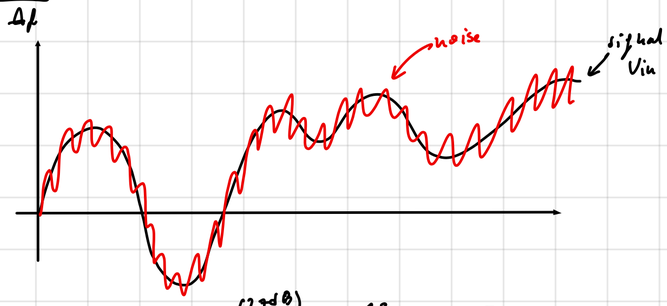


$v(t)=V_p\sin(2\pi f t)$

↳ $V_{rms}=\sqrt{\frac{1}{T}\int_0^T v^2(t)dt}=\frac{V_p}{\sqrt{2}}$

The power of v(t) is equal to the power of a DC volt V_{rms}

↳ $SNR=\frac{(V_{pin}/\sqrt{2})^2}{V_{rtotal}^2/\Delta f}$ (rms values)



So if we have a signal

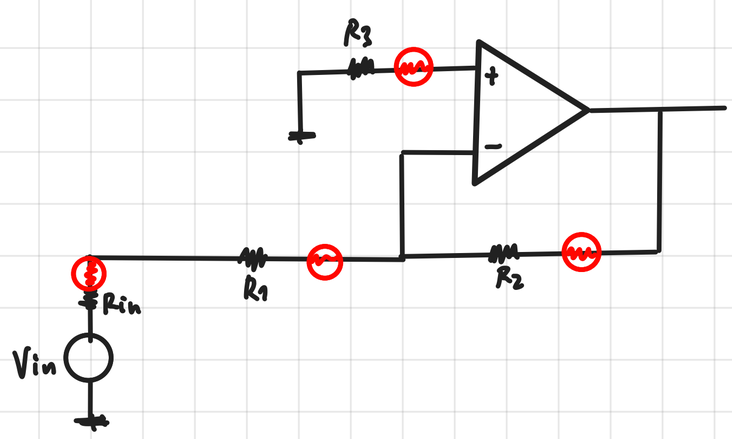
↳ the signal rms is $10^{\frac{27dB}{20}}=10^{1.3}=23\ \left[\frac{V_{rms}}{V_{rms}}\right]$ higher than the rms noise

↳ the power of the signal is $10^{\frac{27dB}{10}}=555=23^2$ higher than the noise power

Note: The SNR doesn't tell me how bad my signal is because it consider the total noise.

↳ To understand better how the contribution of the intrinsic noise at the source (the one on Rin) compares to the contribution of the noise of the OpAmp, we have to consider the NF:

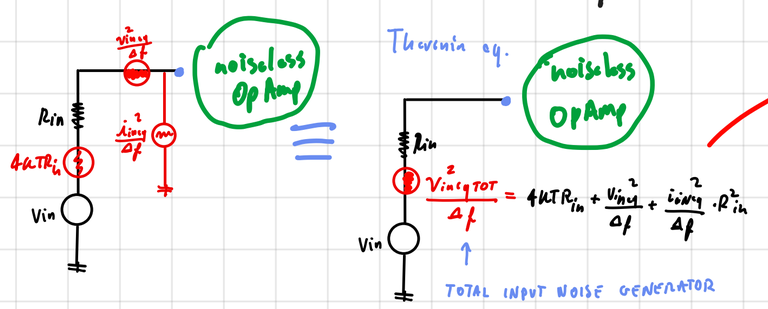
Ex.



$\rightarrow NF=\frac{SNR_{in}}{SNR_{out}}=\frac{S_{in}}{N_{in}}\cdot\frac{N_{out}}{S_{out}}=\frac{N_{out}}{N_{in}\cdot G}=\frac{\text{tot. out. noise}}{\text{Out. noise due to just Rin}}$ ($S_{out}=S_{in}\cdot G$)

intrinsic source noise

Consider also that in the Thevenin eq. we have

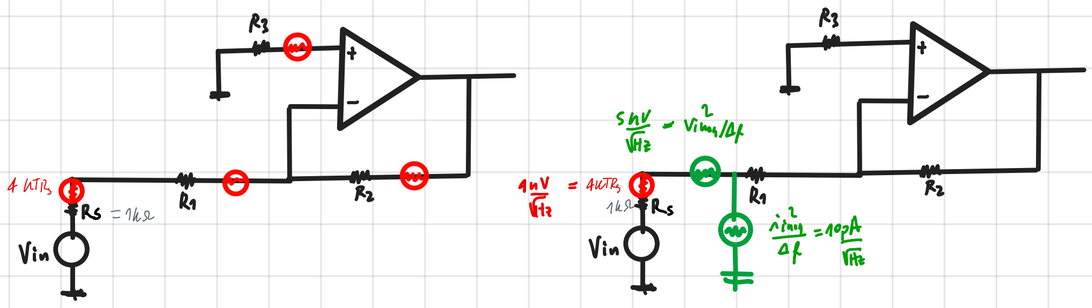


So now the resistor seems to be noisier than the intrinsic resistor because we also consider the noise of the OpAmp

↳ This mean that if the thermal noise is higher → the equivalent temperature T_{eq} will be higher

↳ $T_{eq}=(F-1)T_0$ ← before the eq.

Ex. Consider the circuit:



↳ • $R_{sopt}=\sqrt{\frac{(5nV/\sqrt{Hz})^2}{(10pA/\sqrt{Hz})^2}}=\frac{5nV/\sqrt{Hz}}{10pA/\sqrt{Hz}}=500\Omega$

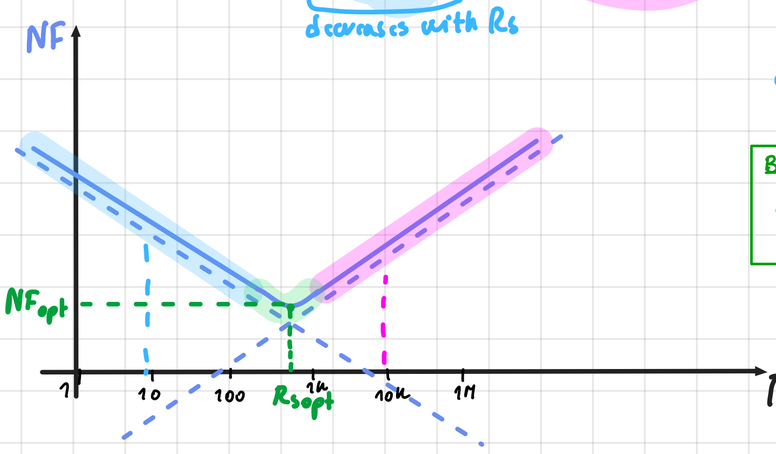
• $NF_{opt}=10\log_{10}\left(1+\left(\frac{5nV/\sqrt{Hz}}{2.8nV/\sqrt{Hz}}\right)^2+\left(\frac{10pA/\sqrt{Hz}\cdot500\Omega}{2.8nV/\sqrt{Hz}}\right)^2\right)=10\log_{10}(1+3.2+3.2)=8.7dB$

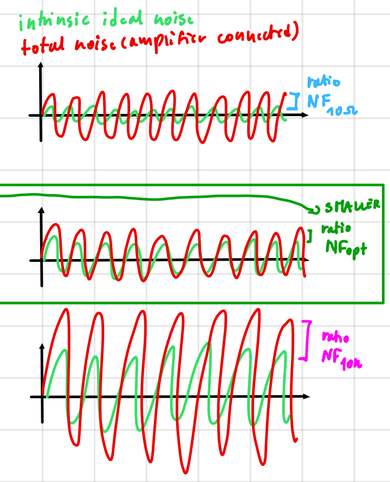
at the opt. these 2 are equal

Noise Figure

$NF=10\log_{10}\left(1+\frac{\left(\frac{5nV}{\sqrt{Hz}}\right)^2}{4kTR_S}+\frac{\left(10pA/\sqrt{Hz}\right)^2\cdot R_S}{4kT}\right)$

(first term: decreases with Rs; second term: increase with Rs)



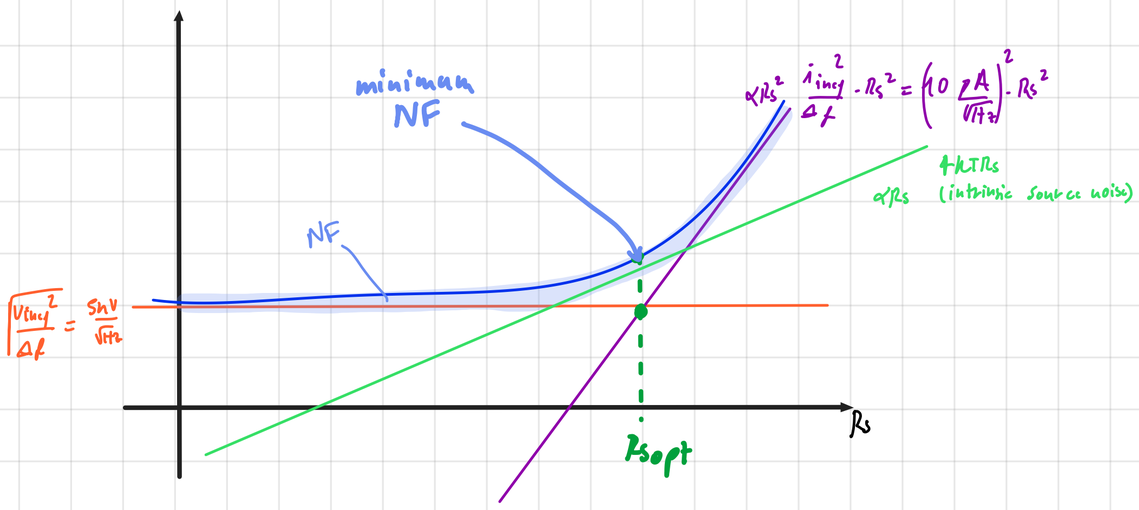


• at 10Ω

• at 500Ω

• at 10k

↳ We could also see that NFopt is the BEST COMPROMISE by plotting the trends of the noise contribution:



$\rightarrow NF=\frac{\text{total-noise}}{4kTR_s}$

Obs. We can also see that for:

- $R_s\to0$ → the $\frac{v_{ineq}^2}{\Delta f}$ prevails
- $R_s\to\infty$ → the $\frac{i_{ineq}^2}{\Delta f}\cdot R_s^2$ prevails

Noise in DIODES

Differently from resistors, which mainly introduce thermal noise, a diode mainly introduces shot and flicker (1/f) noise; its equivalent noise circuit is represented in Fig. 1.17 (where $r_d=1/g_m$ is not a physical resistance, thus it doesn't introduce any noise). Both flicker and shot noise can be represented by means of current generators in parallel to r_d, leading to the following expression for equivalent noise generators:

$$< v_s^2 >= 4kTr_s\Delta f \qquad < i^2 >= 2qI_D\Delta f + K\frac{I_D^a}{f}\Delta f$$

The shot noise can also be expressed as:

$$< i_s^2 >= 2qI_D\Delta f = 2q\frac{kT}{kT}I_D\Delta f = 4kT\left(\frac{1}{2}\cdot g_m\right)\Delta f$$

($r_d=\frac{1}{g_m}$ is NOT noisy but it seems so)

In addition, a secondary noise source associated to a diode is the thermal noise due to its series resistance r_S (caused by the non-nil resistivity of the semiconductor); even though this noise source is often very negligible with respect to shot noise, it is worth to spend a few words discussing its contribution.

Note that increasing the diode bias current brings to an increased thermal noise. This trend is due to the fact that increasing the biasing the electric fields become significant also in the ohmic regions and the present carriers are not anymore in balance with the reticle at the temperature T.

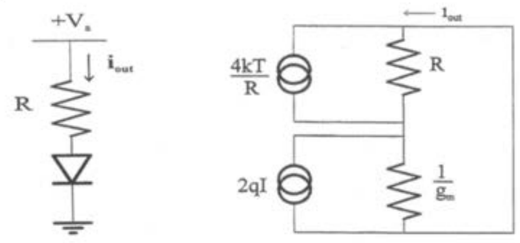


Fig. 1.19: Computation of the current noise in a biased junction (left) and an equivalent circuit for small signals (right) with the noise generators.

The average kinetic energy, E, of the charges in the resistive paths can by formally written as $E=3/2kT_e$, with $T_e>T$ thus it is not surprising if also the thermal noise of these charges is $4kT_e/Re$ and it increases at higher biasing. Let's consider r_S=100 Ω, I_d equal to 0.01 mA, 0.1 mA, 1 mA, 10 mA obtaining the following white noise:

$$g_m=0.4mS,\ r_d=2.5k\Omega \Rightarrow S_{out}=3\cdot10^{-24}+2.5\cdot10^{-25}=3.25\cdot10^{-24}\ A^2/Hz$$
$$g_m=4mS,\ r_d=250\Omega \Rightarrow S_{out}=1.6\cdot10^{-23}+1.36\cdot10^{-23}=1.96\cdot10^{-23}\ A^2/Hz$$
$$g_m=40mS,\ r_d=25\Omega \Rightarrow S_{out}=1.28\cdot10^{-23}+1.06\cdot10^{-22}=1.19\cdot10^{-22}\ A^2/Hz$$
$$g_m=0.4S,\ r_d=2.5\Omega \Rightarrow S_{out}=1.9\cdot10^{-24}+1.6\cdot10^{-22}=1.62\cdot10^{-22}\ A^2/Hz$$

thus with r_S=100Ω, in order to minimize the overall noise, we will chose the lowest current possible even if it causes an increase of $(1/g_m)^2$ and an increase of the flicker noise.

Note that both shot noise ($2qI_D\Delta f$) and small signal impedance ($1/g_m = kT/qI_D$) depend on I_D: increasing I_D has the double effect of reducing the diode small-signal impedance (beneficial effect) and of increasing its noise contribution (detrimental effect). The equivalent noise generator is equal to :

$$2qI_D\Delta f\left(\frac{1}{g_m}^2\right)=2qI_D\Delta f\left(\frac{kT}{(qI_D)^2}\right)=\frac{2kT}{q}\cdot\frac{1}{I_D}\Delta f$$

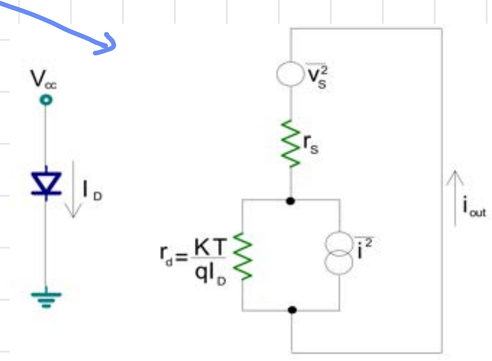
from which we can notice that increasing I_D has a global beneficial effect to the (voltage) noise, due to the square power associated to it.

Let us now consider the network in Fig. 1.19 (right) and analyze the current power spectrum associated to i_{out}. We can consider the equivalent small signal circuit (Fig. 1.19 left) which also schematizes the two noise equivalent generators, associated to both the resistor R (thermal noise 4kT/R) and to the junction (shot noise 2qI).

The noise current that flows between power supply and ground, i_{out}, is the superposition of two independent contributes; its spectral density equals:

$$S_{out}=2qI\cdot\frac{(1/g_m)^2}{(1/g_m+R)^2}+\frac{4kT}{R}\cdot\frac{R^2}{(1/g_m+R)^2}$$

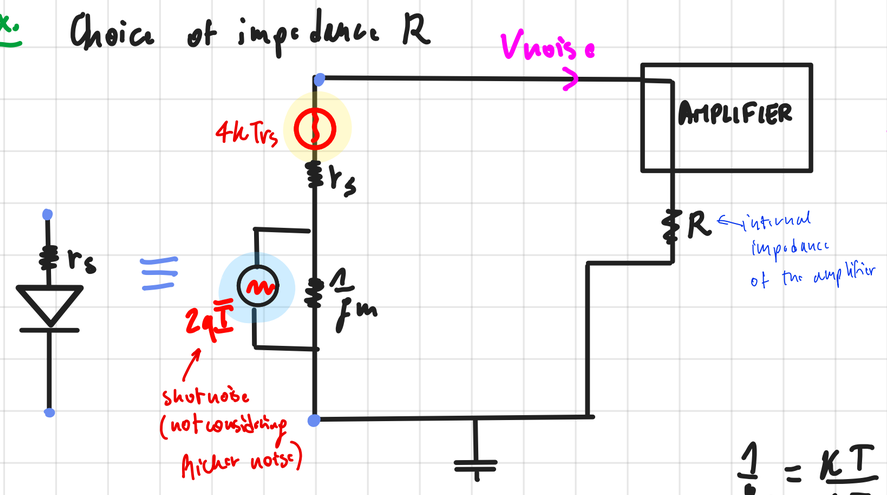
If $1/g_m >> R$, the dominant contribution is 2qI; conversely, if $1/g_m << R$, the contribution 4kT/R becomes dominant. At low currents, i.e. when the junction resistor is much higher than the resistor R, the current noise spectrum has an intensity equal to 2qI; when the bias current increases, the series resistance dominates the overall noise and thus the spectrum intensity tends to the constant value 4kT/R.



④ can be modeled as a thermal noise on an eq. resistor

Equivalent Noise Resistors: $R_{eq,I}=2\cdot\frac{1}{g_m}$ $\qquad R_{eq,V}=\frac{1}{2}\cdot\frac{1}{g_m}$

Ex. Choice of impedance R



$$\frac{1}{f_m}=\frac{KT}{qI}$$

$$\frac{v_{noise}^2}{\Delta f}=2q\bar{I}\left(\frac{\frac{1}{f_m}}{\frac{1}{f_m}+r_s+R}\right)^2\cdot R^2+4kTr_s\left(\frac{R}{R+r_s+\frac{1}{f_m}}\right)^2=$$

highlight this

$$\simeq\frac{2q}{\bar{I}}\left(\frac{kT}{q}\right)^2+4kTr_s\left(\frac{1}{1+\frac{r_s}{R}+\frac{kT}{q\bar{I}R}}\right)^2$$

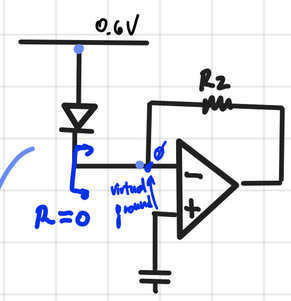
decreases for increasing $\bar{I}$ (does NOT depend on R)

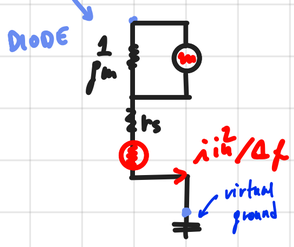
decreases for decreasing $\bar{I}$ (decreases for decreasing R)

↳ So for the impedance R:

PHOTO-CURRENT MODE

Low R





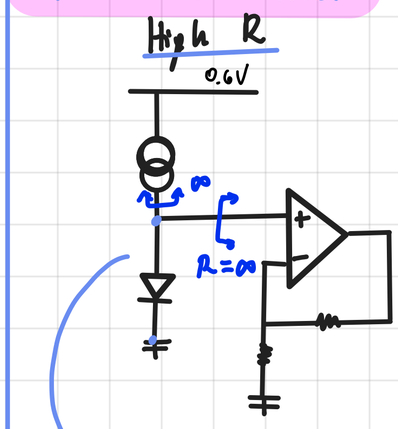
$$\frac{i_{in}^2}{\Delta f}=2q\bar{I}\left(\frac{\frac{1}{f_m}}{\frac{1}{f_m}+r_s}\right)^2+\frac{4kTr_s}{\left(\frac{1}{f_m}+r_s\right)^2}$$

- increase $\bar{I}$ to kill this shot noise
- increase r_s to kill this contribution

If $\bar{I}\uparrow$ then $\frac{1}{f_m}\propto\frac{1}{\bar{I}}$ becomes negligible wrt r_s

PHOTO-VOLTAIC MODE

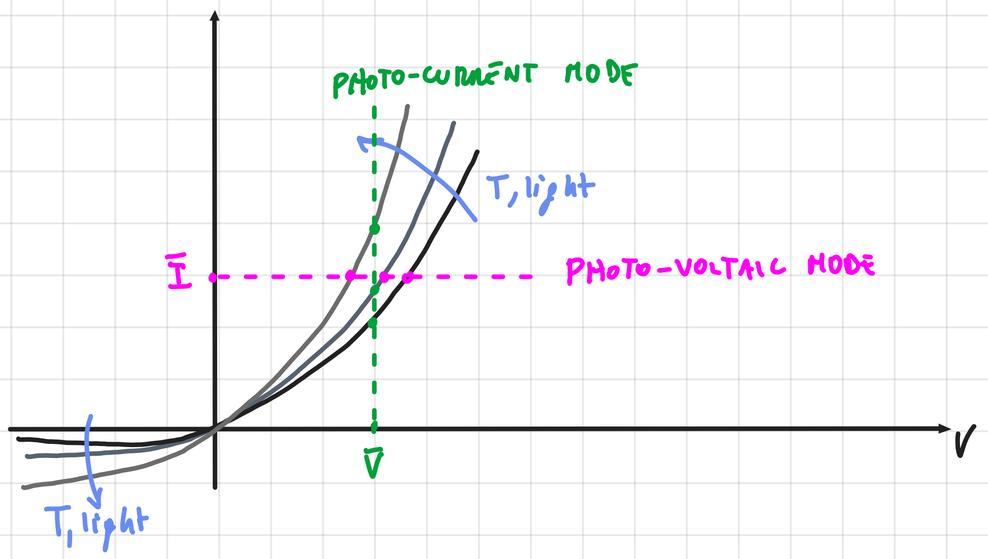
High R



$$\frac{V_{in}^2}{\Delta f}=4kTr_s+2q\bar{I}\left(\frac{1}{f_m}\right)^2$$

- decrease r_s to decrease this contribution
- increase $\bar{I}$ to decrease this

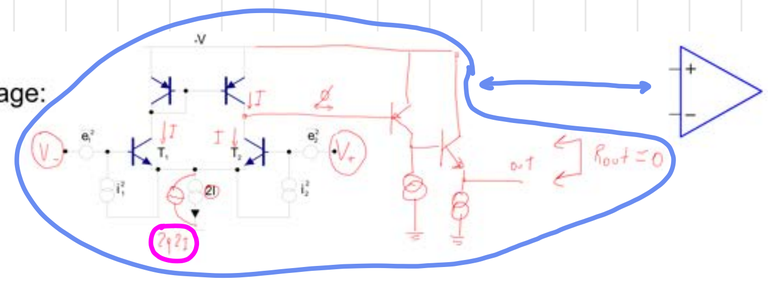
→ Diode characteristic



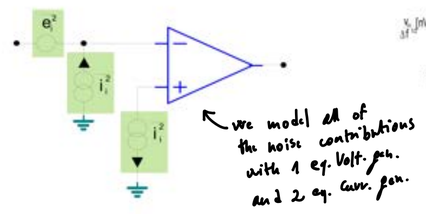
Noise in OpAmps

We actually did not see all the possible noise contributions, but we can certainly say that the OpAmps (their components) have a lot of noise contributions.

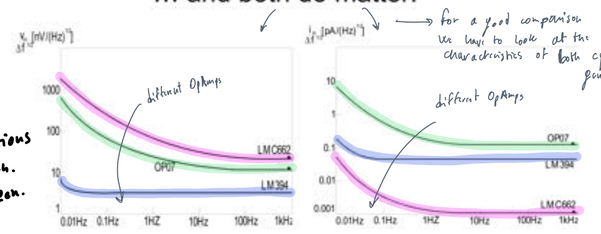
Differential input stage:



Input equivalent generators: … and both do matter!

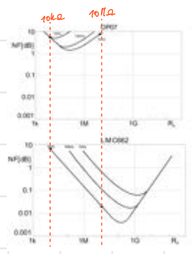


we model all of the noise contributions with 1 eq. Volt. gen. and 2 eq. curr. gen.



for a good comparison we have to look at the characteristics of both equivalent generators

and compare their NF



Note: The 2 currents generators will take into accounts contributions like the shot noise of the tail, so given that this value is accounted for in both equiv. generators, these two will actually be CORRELATED.

But we can neglect this correlation, meaning that we'll actually take into account an overestimate of the noise.

Which one is better? Either lower current or voltage noise?

Comparisons:

OP07 e_{in}=9.6nV/√Hz i_{in}=120fA/√Hz

LMC662 e_{in}=22nV/√Hz i_{in}=0.113fA/√Hz

For R_s=10kΩ: ← Low Rs → Voltage noise is relevant

OP07 $v_{n\,rms\,totale}=\sqrt{(9.6\cdot10^{-9})^2+(128\cdot10^{-15})^2}=9.7nV/\sqrt{Hz}$ NF=4dB ← BETTER

LMC662 $v_{n\,rms\,totale}=\sqrt{(22\cdot10^{-9})^2+(0.113\cdot10^{-12})^2}=22nV/\sqrt{Hz}$ NF=11dB

For R_s=10MΩ: ← High Rs → current noise is relevant

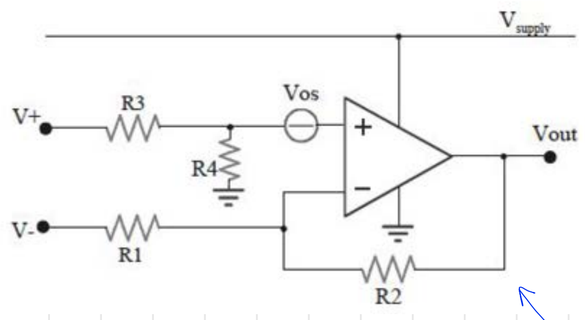
OP07 $v_{n\,rms\,totale}=1.3\mu V/\sqrt{Hz}$ NF=19dB

LMC662 $v_{n\,rms\,totale}=22nV/\sqrt{Hz}$ NF=0.025dB ← BETTER

Intro

(Book p. 273)

The term Instrumentation Amplifier refers to a special class of amplifiers which, on account of their nature, are widely used when it is needed to acquire weak differential signals from high impedance sources in the presence of large common mode interferences. We shall see their basic features, when they are employed, and how they are designed for achieving an ideally infinite CMRR, negligible offset, high input impedance.



Voltage difference amplifier

(Book p. 273)

A simple solution to the problem arising from the need that has just been mentioned might be the so-called Differential Amplifier depicted in Figure 3.1. In this configuration, many non-idealities impair the Common Mode Rejection, and the high common mode gain is the heaviest limit of the Differential Amplifier. In order to have an infinite CMRR, the following two conditions must be met:

(CMRR = Common Mode Rejection Ratio)

- the Op-Amp must have an infinite CMRR;
- R3/R4 must be equal to R1/R2.

In practice, the second condition is the dominant one: the Op-Amp is built with a CMRR much higher than that attributable to the mismatches between the ratios of pairs of elements, which are typically more than 100dB, while the mismatches (if the resistors are discrete and standard) are high and lead to degradation of the CMRR by 40dB. However, even in the field of integrated resistors where the degrees of tolerance are high, this condition dominates the other.

Regarding the Common Mode Rejection Ratio, another issue is related to the input impedance values of the inverting and non-inverting nodes which, in addition to being different from each other, are relatively low. Therefore, the CMRR undergoes further degradation due to the inevitable mismatches of resistance sources. In fact, even if the input resistance values of the inverting and non-inverting nodes were the same, they would not be much larger than the source resistances. Since the source resistance of the positive input will be probably different from the source resistance of the negative input, a common mode signal input become a differential signal and would be treated as such. Unfortunately, R1 and R3 cannot be chosen too large since they must be less than R2 and R4, respectively, in order to obtain a gain of the structure higher than unity whilst R2 and R4 cannot be chosen too large. This structure is sometimes made with discrete components to perform a single measurement and has a fixed gain, and the CMRR is increased to acceptable levels by trimming.

Ex.

In Fig. 4.2, there is an example of how to use an INA: we want to measure the temperature variation through the use of a thermal resistor and an OpAmp with A_0=100dB and CMRR=80dB. We choose R_T=30Ω(@0°C)+0.15Ω/°C. The differential signal is V_{diff}=150µV/°C and V_{out}=70.5mV/°C. We observe that the distributed resistance along the cable introduces 2·ΔV, i.e. a systematic error. For example, if this resistance is 1Ω distributed, we have a systematic error of 2·1Ω·1mA=2mV and a resulting inaccuracy of 13.3°C. (*)1 That is why you often use the 4 wires (2 to force the current I and 2 to measure V_{diff}) connection that is called *Kelvin*. (¥)2

The differential voltage v_{diff}=30Ω·1mA= 30mV causes in the non-inverting stage an input common mode voltage (because V-=V+) equal to $v_{in,cm} = v_{diff} \cdot 470k/(1k + 470k) \simeq 30mV$ and 30mV·471=14.1V output. Because of the finite A_0, at the OpAmp input remains a voltage error $v_\varepsilon = V_{out}/A_0$=100'000=141µV, corresponding to an error of 141µV/150µV/°C=+0.94°C. The presence of a finite CMRR will alter this error because we will have $v_{out} = v_\varepsilon \cdot A_0 + v_{in,cm} \cdot A_{cm}$ because:

$$v_\varepsilon = \frac{v_{out}}{A_0} - v_{in,cm} \cdot \frac{A_{cm}}{A_0}$$

thus, a further contribution of 30mV/CMRR=3µV, which is negligible in this case.

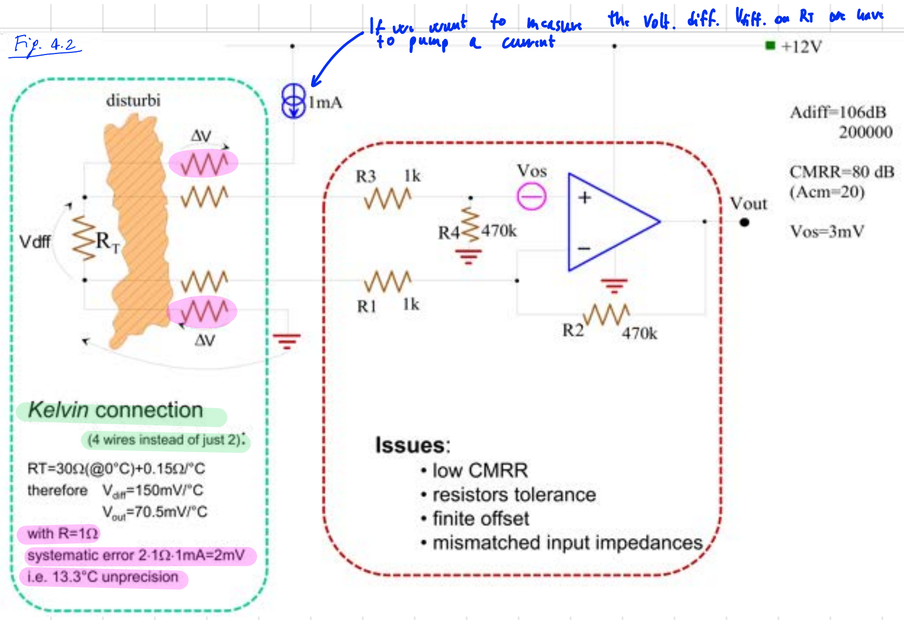
Even resistive degrees of tolerance give a contribution to the common mode:

$$v_{out_{tolerances}} = v_{in} \cdot \left[\frac{R_4}{R_4 + R_3} \cdot \left(1 + \frac{R_2}{R_1}\right) - \frac{R_2}{R_1} \right]$$

tolerance of 10% (R3=R1·0.9 and R2=R4) gives v_{out}/v_{cm}≈10%, i.e. an equivalent CMRR of 470/0.10=73dB, worse than the intrinsic CMRR of the OpAmp.

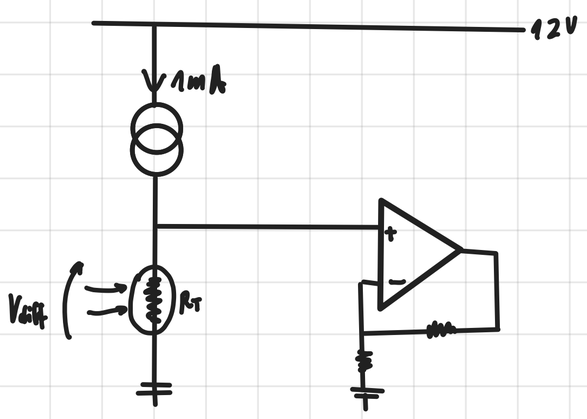
Furthermore, the impedance values seen by the input terminals are different: $Z_{neg}=R_1$=1kΩand $Z_{pos}=R_3+R_4$=471kΩ. Finally, the offset V_{os} cannot be canceled unless you use an external trimmer, dedicated inputs, or through circuital shrewdness. In this case, it introduces an error of 20°C.

Fig. 4.2



Note on the ex.:

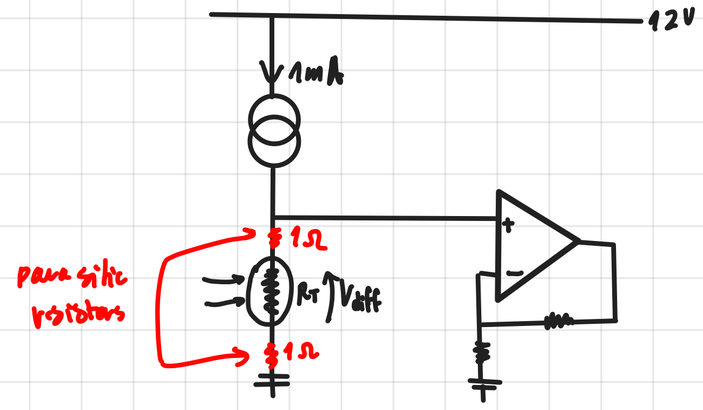
→ We could have think to use a configuration like this to measure V_{diff}:



→ But this has several issues:

- R_T has a very low value → any possible parasitism in series with R_T would cause an error in the read-out

↳ e.g.



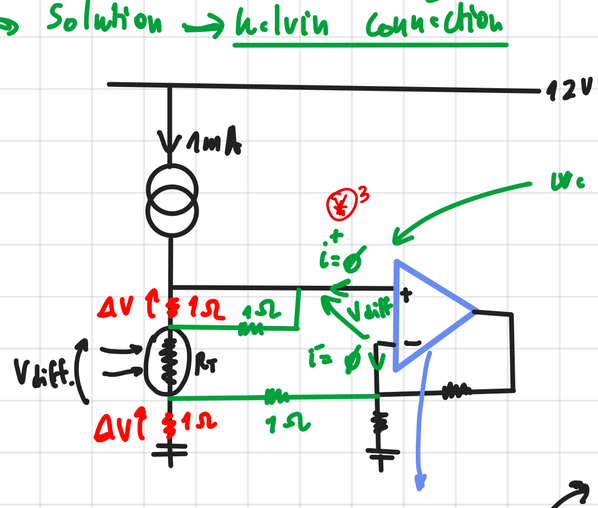
→ Indeed at 0°C → R_T = 30 Ω

↳ $R_T + R + R$ = 32 Ω

↳ with a sensitivity of 0.15 Ω/°C → 2 Ω error

(*)1 $\frac{2}{0.15}$ = 13.3° error

→ Solution → Kelvin connection

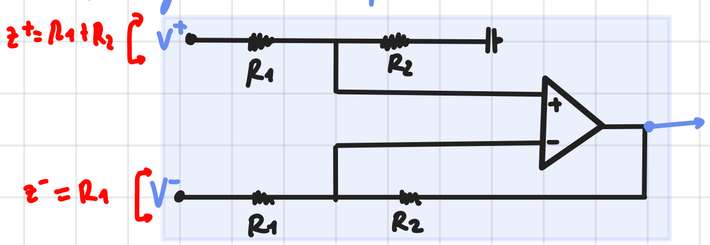


we see that the diff. $v^+ - v^- = V_{diff}$ (supposing that the OpAmp doesn't drink any current ($i = i^+ = i^- = 0$)

↳ so no voltage drop on the 2 additional wires (¥)2

→ To amplify voltage differences ($v^+ - v^- = v_{diff}$) we can then use the following steps (Subtractor → seen at ESO3):

↳ Voltage differences amplifier



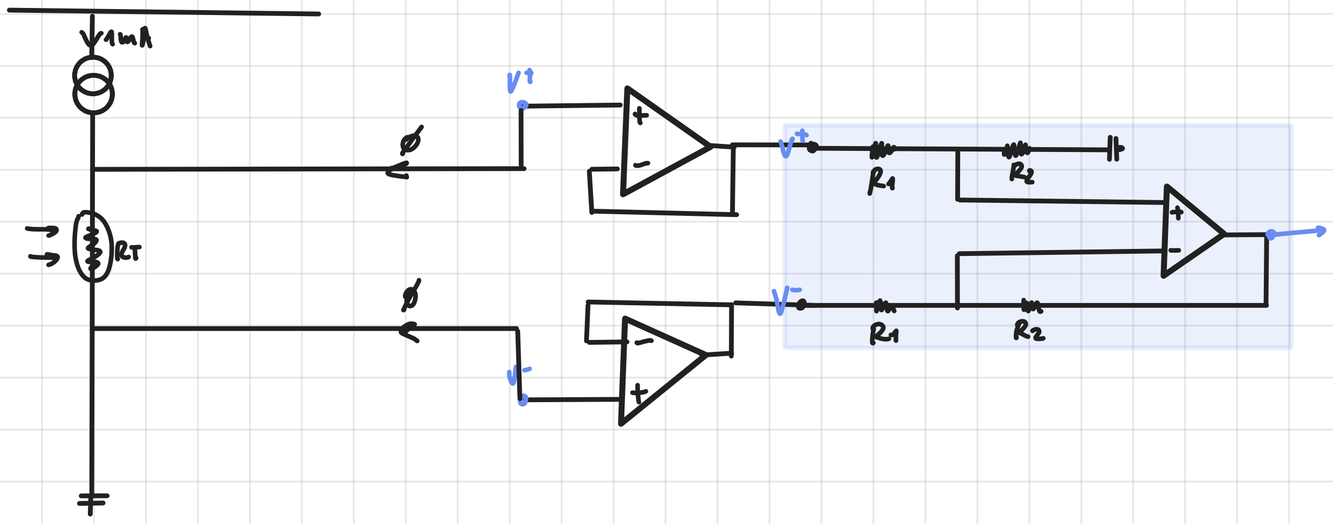
Issues

- Input impedance $z^- = R_1$
- Input impedance $z^+ = R_1 + R_2$

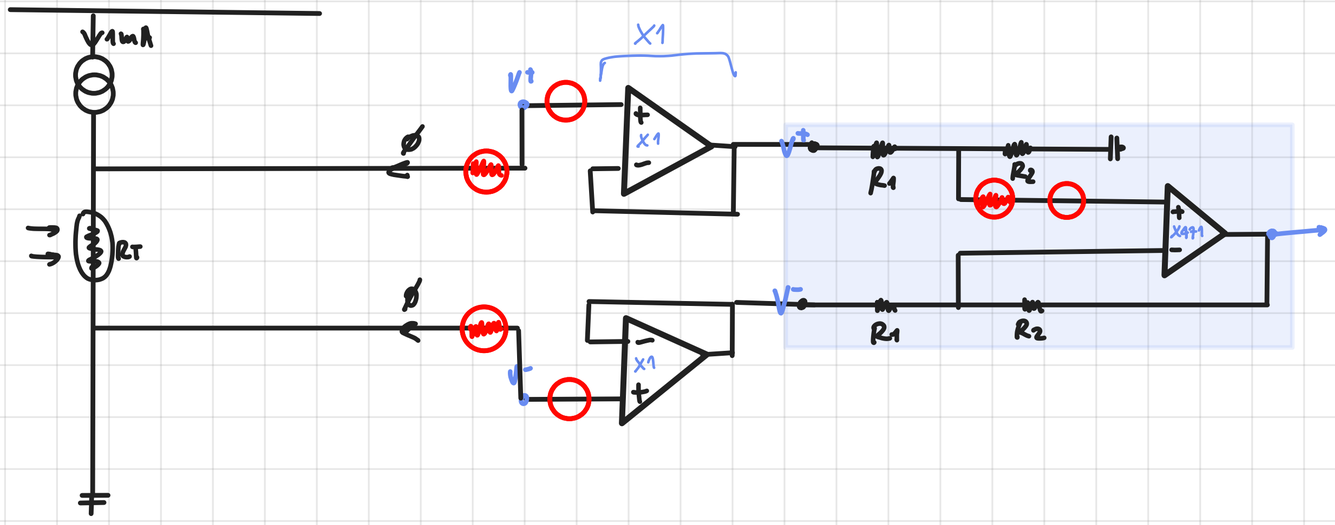
≠ → so i is actually ≠0 (*)3

and $i^+ \neq i^-$ → the voltage drop across the 2 additional cables added with Kelvin connection won't be 0 → and the two voltage drops won't compensate each other

→ To improve the previous stage → we can introduce buffers

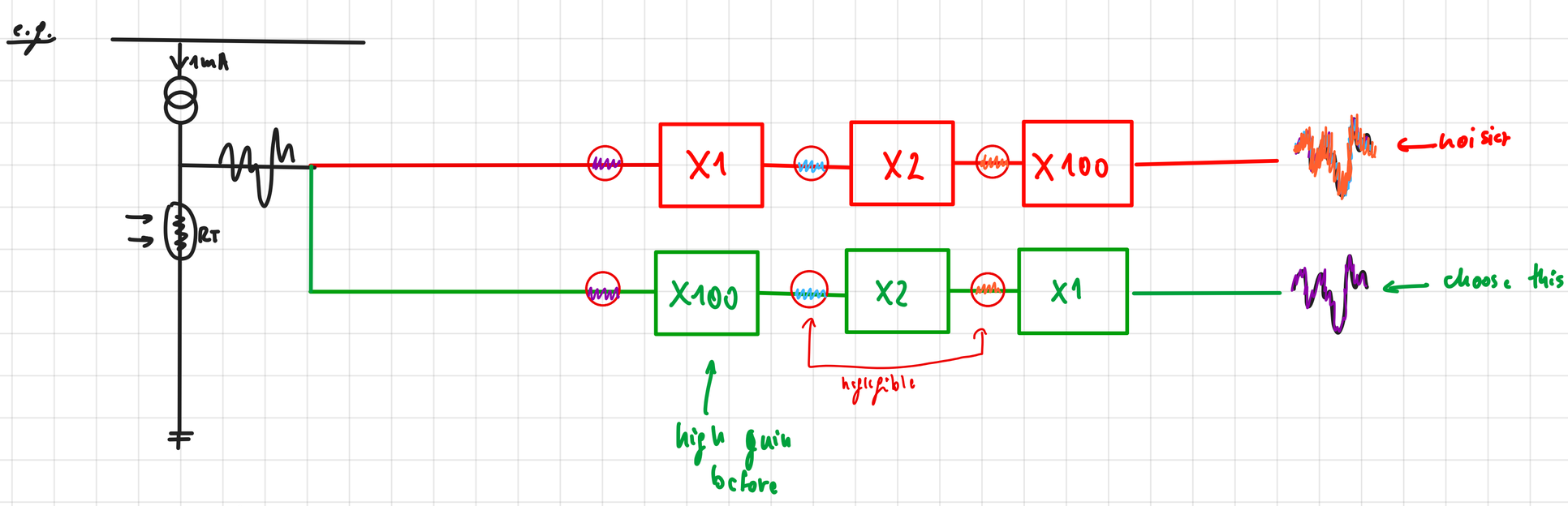


↳ But this circuit also have problems → OpAmps' offsets and noises → too many components, they add too much noise

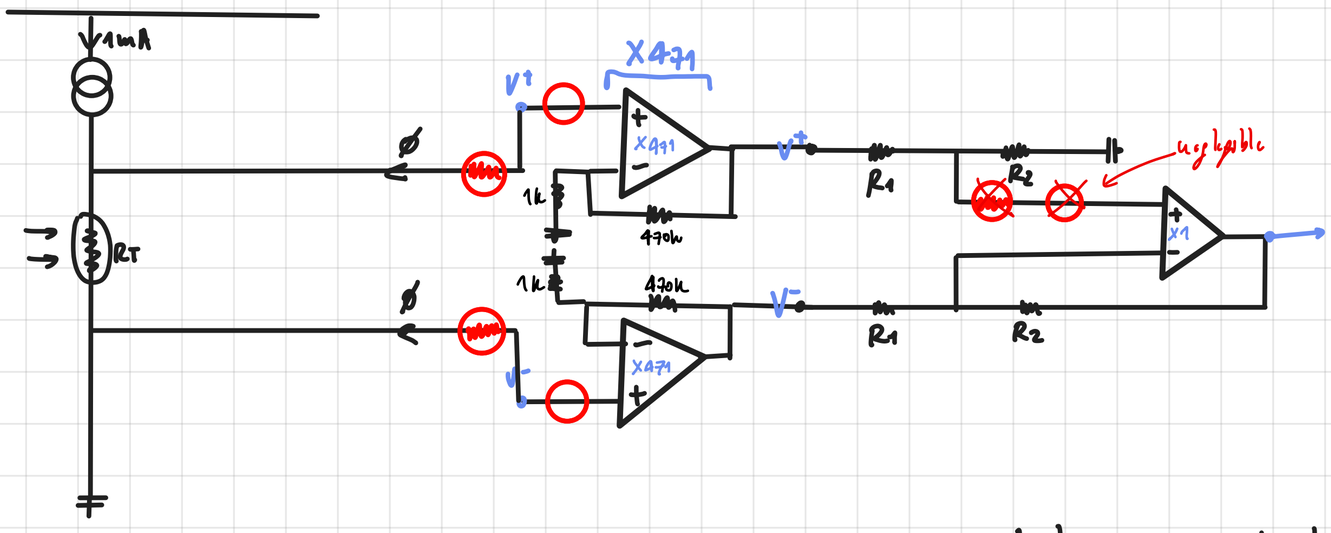


↳ Obs. If we put many stages in series is better to put the one with the higher stage before in this way only its noise contribution will affect the out-signal and the contributions of following stages with small noise will be negligible. If we instead put in series stages with small gain before the ones with higher gains all noise contributions will be collected going to the out signal.

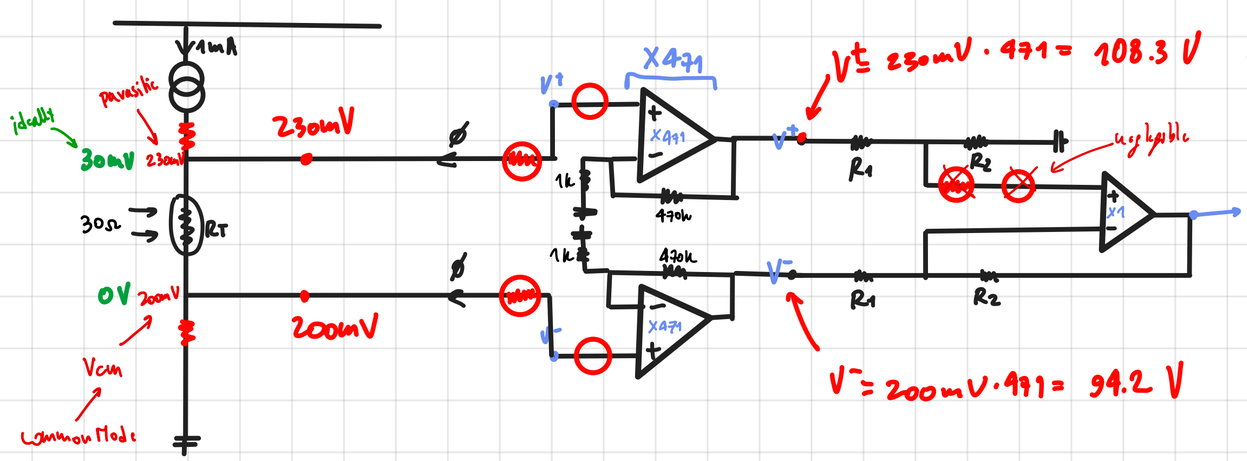
c.f.



→ So in our case its better to put the stage with gain x1 after:



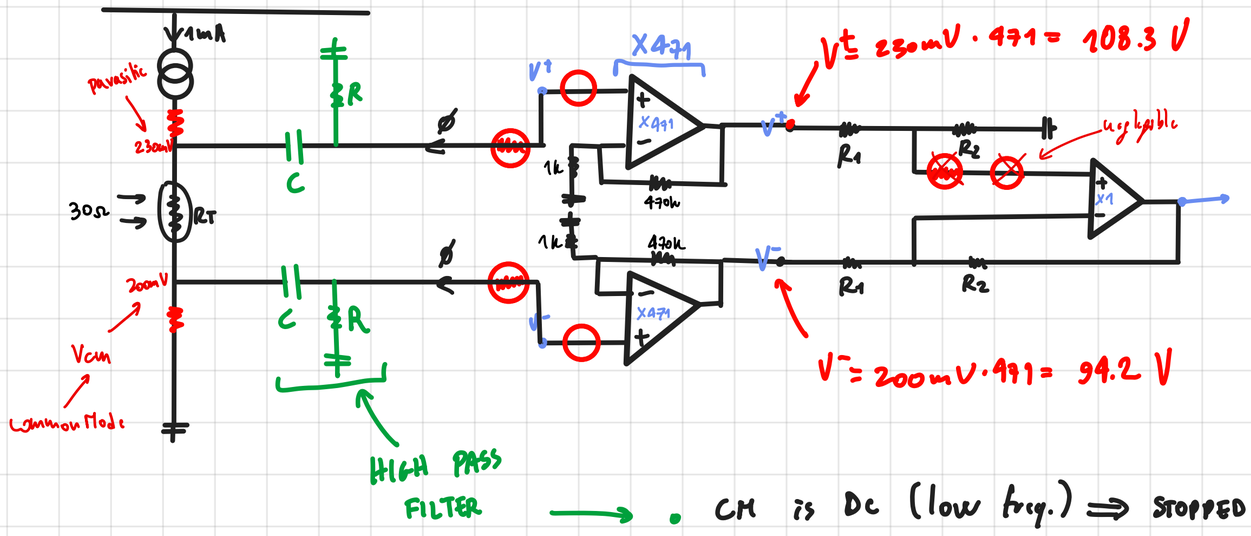
→ But there could be anyway some issues, indeed we could have some parasitic res. that will also change the common mode gain.



The OpAmp can't reach these value → it will SATURATE to the same value
↳ so V+ - V- = 0
↳ no more Vdiff to amplify

↳ For this reason I will like to have • Gcm = 0 - 1 (low)
• Gdiff = 471

→ To amplify only the diff. signal and not the common mode a solution might be:

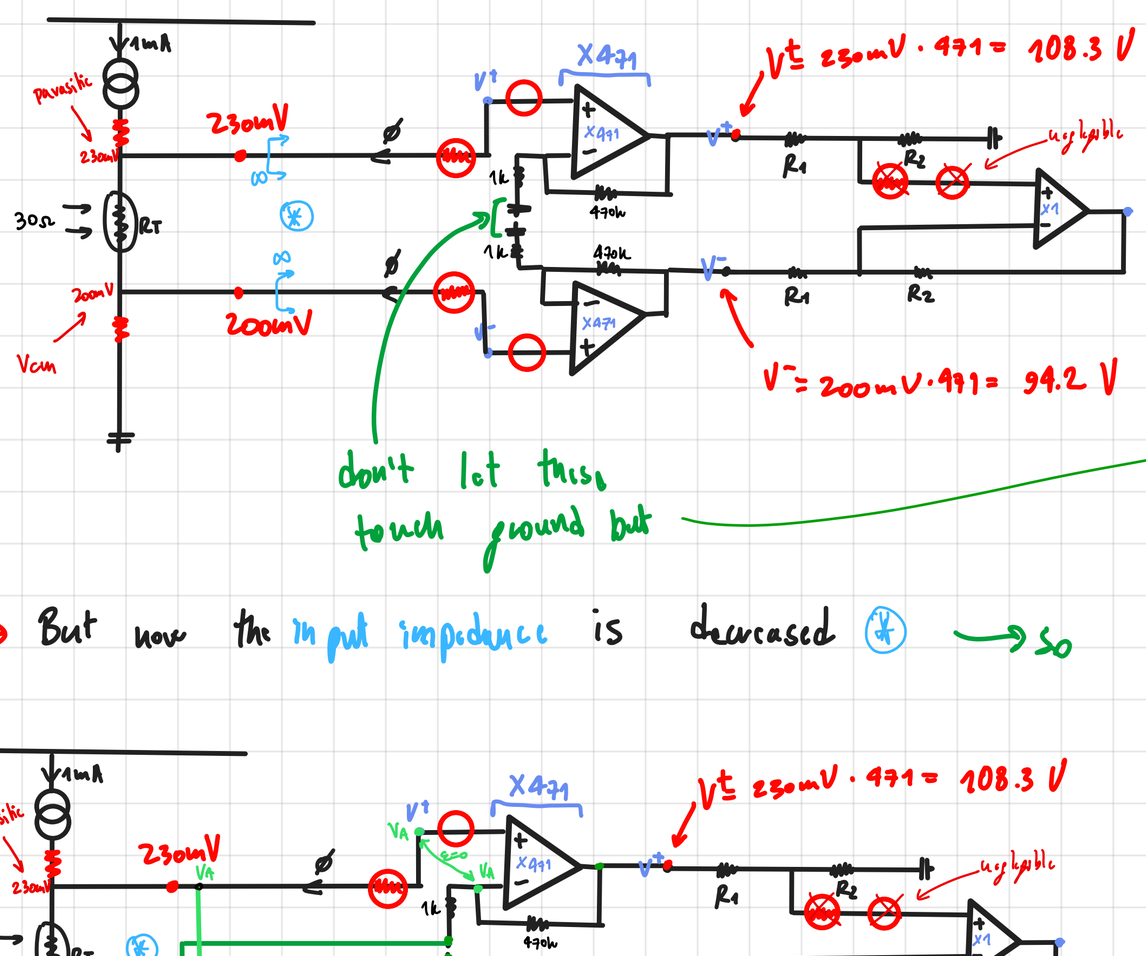


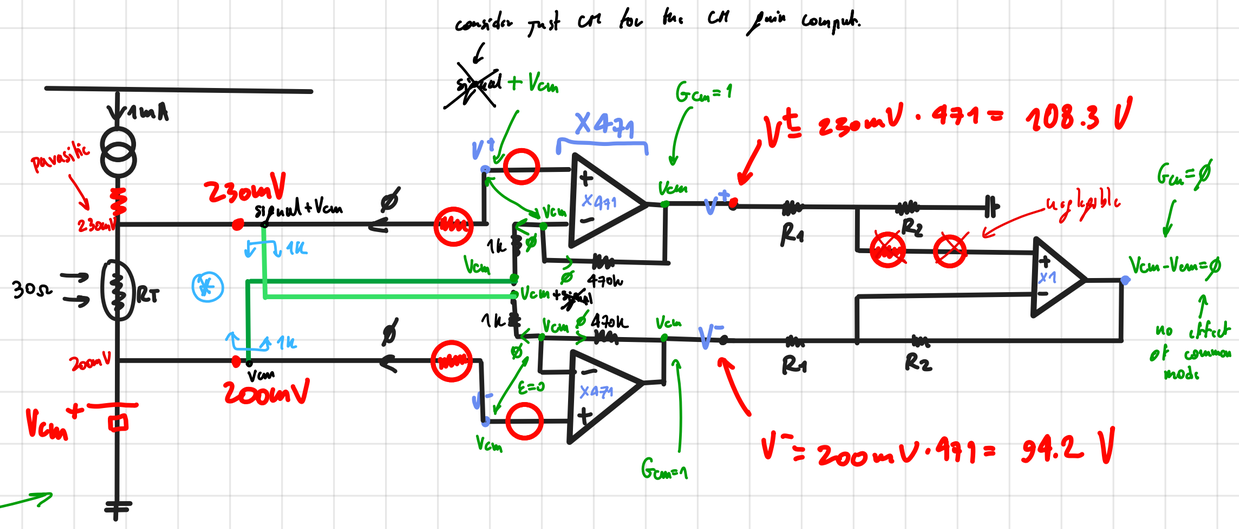
- CM is DC (low freq.) ⇒ STOPPED
- diff signal is AC (high freq.) ⇒ PASS

but our signal can also be DC (e.g. const. temp.)

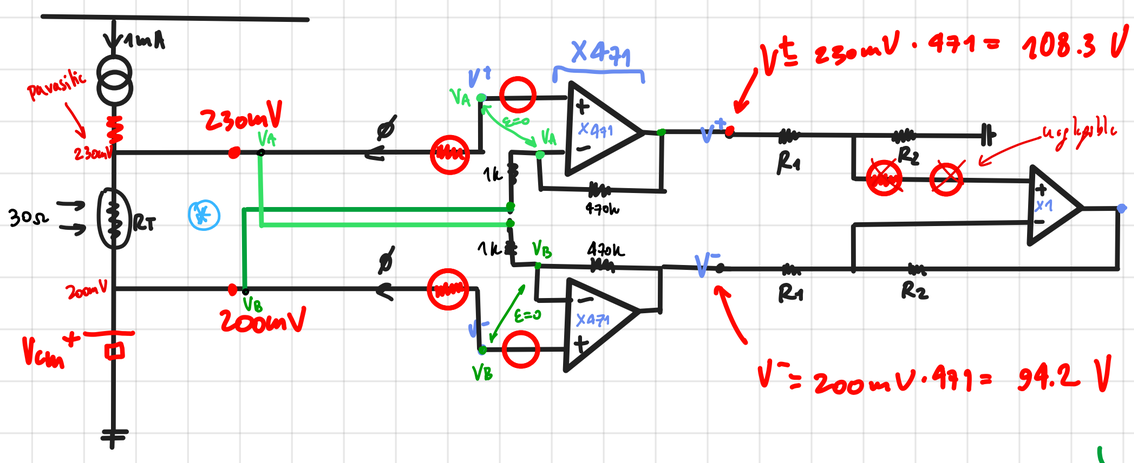
We want also to amplify DC values

→ To amplify only the diff. signal (also the DC part) we can try with this solution:

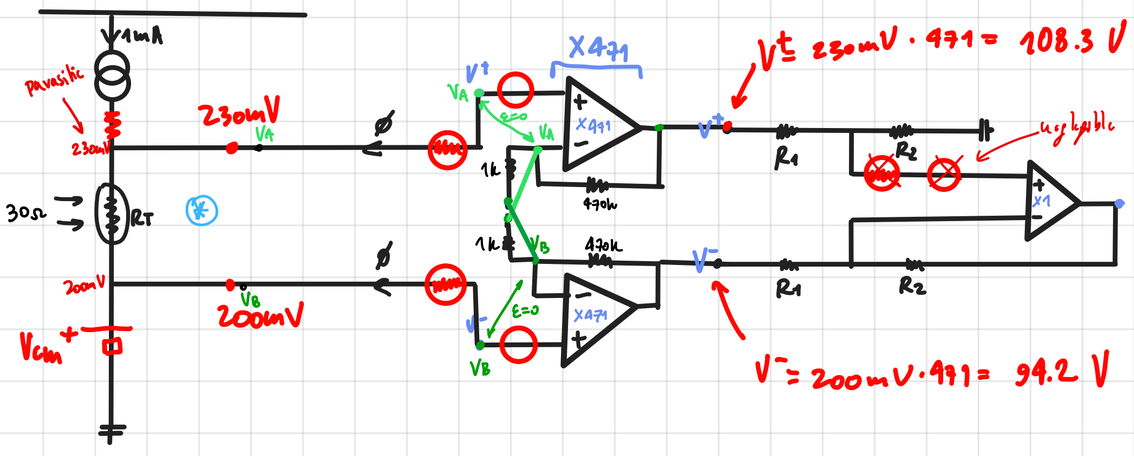




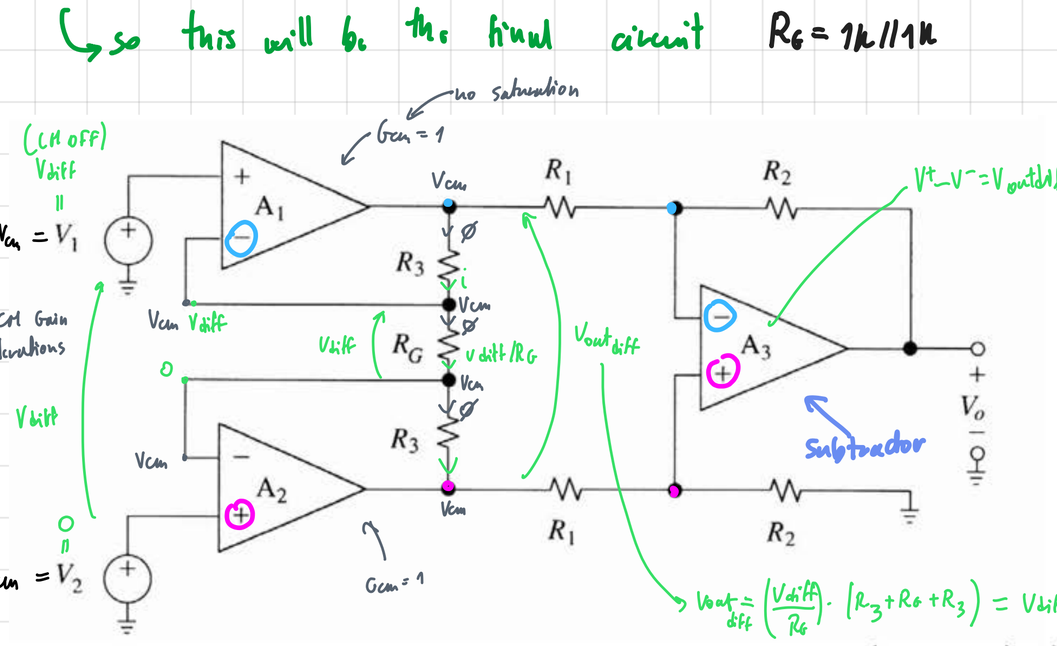
→ But now the input impedance is decreased (*) → so we must change something



So consider



↳ so this will be the final circuit $R_G = 1k//1k$



Differential gain: → $A = A_I A_{II} = \left(1 + \frac{2 \cdot R_3}{R_G}\right) \cdot \left(\frac{R_2}{R_1}\right)$ ($V^+ - V^-$)

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

Instrumentation amplifier

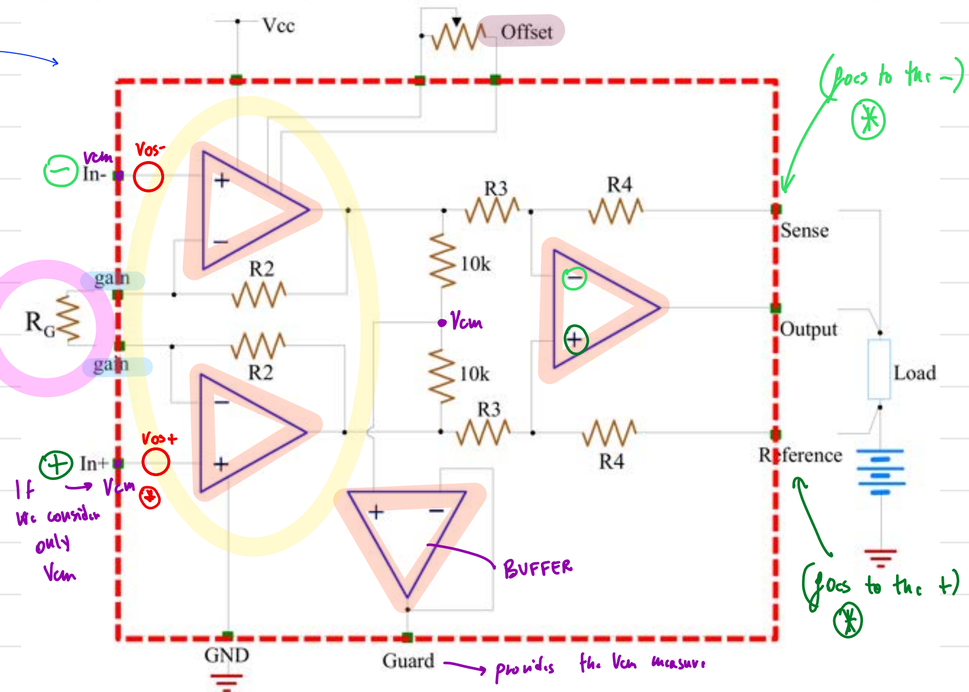
(Book p.296)

Let us see how, with some smart changes, we may be able to solve some problems of the differential amplifier by obtaining an instrumentation amplifier with performance adequate for our needs. We refer to Fig. 4.3. To solve the problem due to the low input resistance values, connecting two buffers to the input terminals could be sufficient. If we have two gain stages in our design, we could not only solve the preceding problem, but also increase G_{diff} and thus the whole CMRR.

The connection between the virtual grounds of the two input amplifiers ensures $G_{cm}=1$ for this first stage. In fact, if the input is a common mode signal, also the two virtual grounds will follow this signal, so no current can flow into R_G and R_2, and the OpAmp output will follow the common mode input. In conclusion, the stage preceding the difference amplifier has $G_{cm}=1$ and $G_{diff}>1$, thus improves the CMRR of the whole circuit.

Compared to the simple differential amplifier, this structure has also the great advantage of being able to vary the gain simply by tuning the resistor R_G. In practice, two external terminals (gain pins) are made available for connection to an external resistor (which runs in parallel to the internal one if present), and you can vary the gain without appreciably changing the basic behavior of our system.

The overall CMRR is equal to the product of the CMRR of the difference amplifier and G_{diff} of the first stage, and the CMRR depends directly on the gain. The noise of the system depends on the gain setting: it is easy to see that the higher the first stage's gain is, the more negligible the noise of the next stage (compared to that of the first one) will be.



$$G_{diff}=1+\frac{2R_2}{R_G}$$

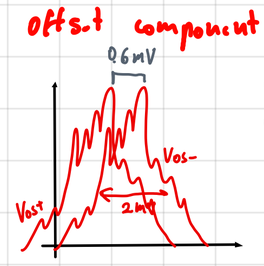
$G_{cm}=1$

$$G_{diff}=\frac{R4}{R3}$$

G_{cm} almost nil

(*) → If $V_{os+}=V_{os-}$ (deterministic offset) they can compensate each other

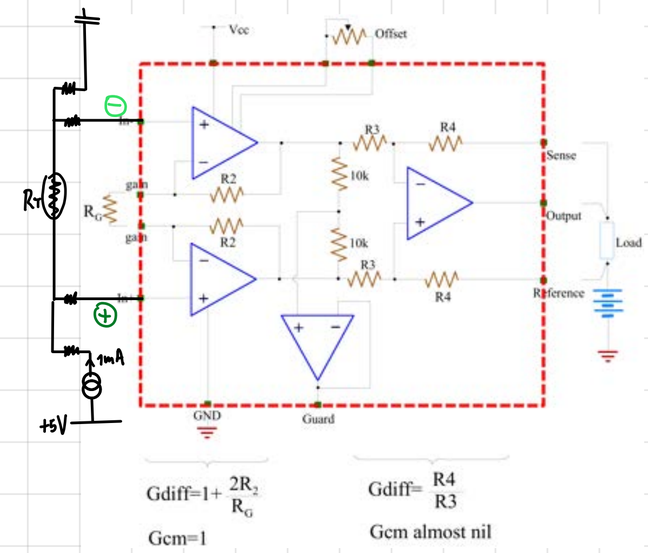
↳ they could also have a random offset component



for both Vos+ and Vos-
- determ. = 2mV
- random = ±3 mV

↳ so there will be an offset

- **Input connection**



$$G_{diff}=1+\frac{2R_2}{R_G}$$

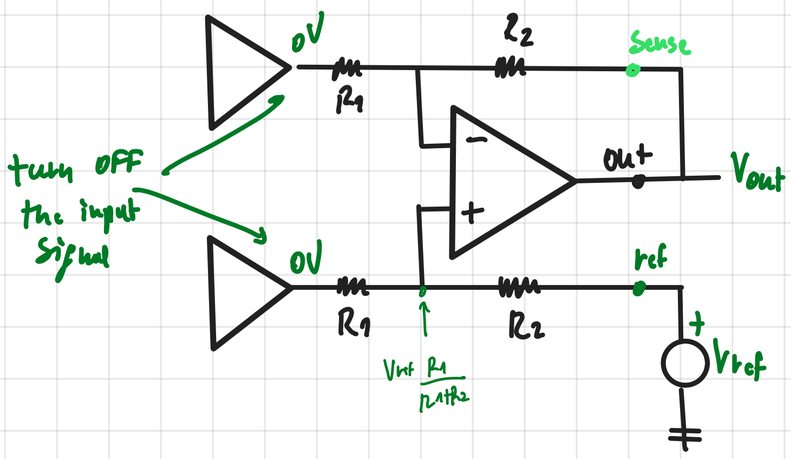
$G_{cm}=1$

$$G_{diff}=\frac{R4}{R3}$$

G_{cm} almost nil

(*) REFERENCE

It's like studying this circuit:



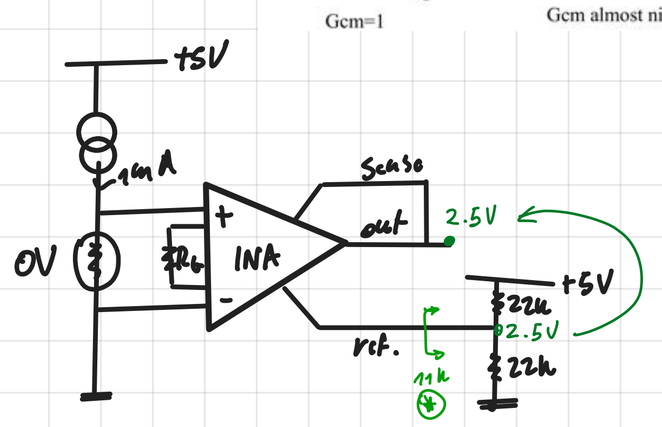
$$V_{out}=\left(V_{ref}\frac{R_1}{R_1+R_2}\right)\cdot\left(1+\frac{R_2}{R_1}\right)=V_{ref}$$

→ so if we connect ref to ground

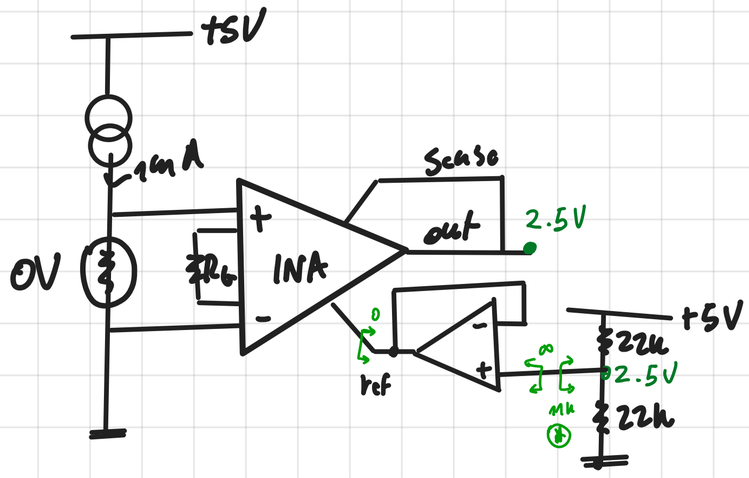
↳ $V_{ref}=V_{out}=0V$

(So it indeed act like a reference)

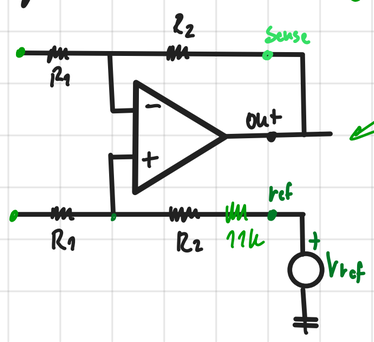
Ex.



Note on ex. It would actually be better if we put a buffer on the ref. like this.



→ In this case considering that inside the INA there will be the resistors that must see a voltage source (so with ↳ instead of ↳)

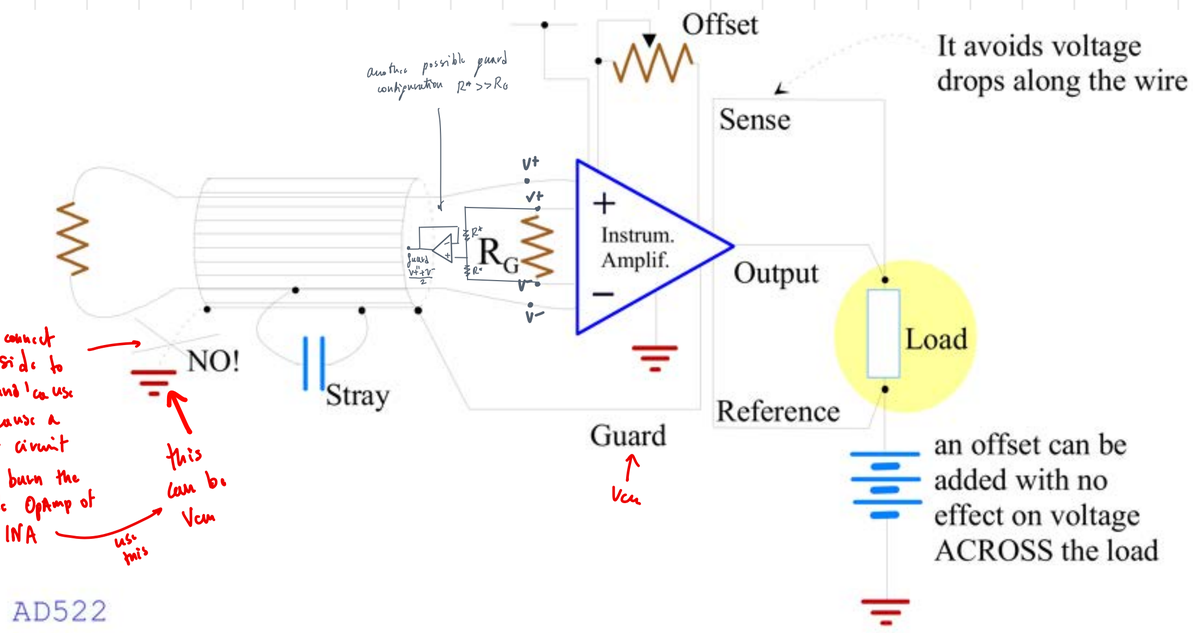


If this is not negligible wrt the other resistors

↳ making the voltage amplifier not ideal

(*) SENSE → To avoid the signal lost along the wire, due to the stray resistances

Ex.



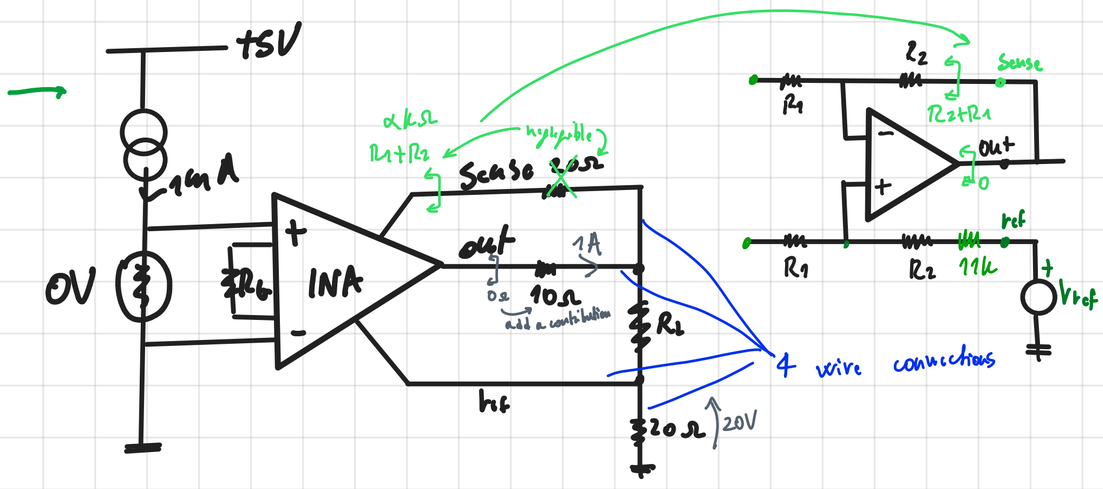
AD522

Vcc=10÷36V	Isupply=10mA
Vos=1mV	IB=25nA
Rin=1GΩ	
GBWP=300kHz	
$V_{noise\ in}=15\mu V_{rms}$	(10÷15kHz)

For example the AD522 INA has the following features: Vcc=10÷36V Isupply=10mA Vos=1mV IB=25nA Rin=1GΩ GBWP=300KHz Vnoisein=15μVrms (10÷15kHz)

Fig. 4.4 shows the correct usage of the INA, with the proper exploitation of the extra pins (reference, sense and guard) to minimize errors due to connections and parasitisms.

In addition to the shown typologies, there are several others with better performance than those obtainable with these circuit topologies, such as the class of Instrumentation Current Feedback Mode.

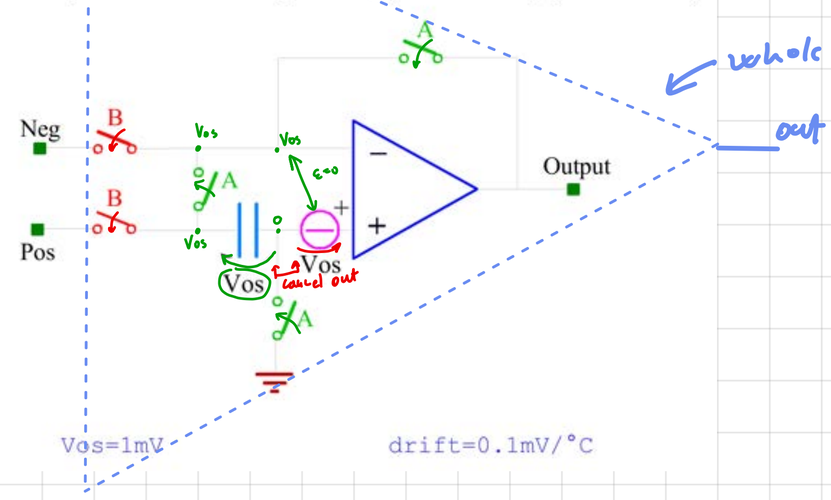


Auto zeroing

(Book p. 279)

The offset voltage in CMOS technology is worse than that in bipolar technology. For this reason, in order to improve the situation, circuit topologies, such as Commutating Auto Zeroing (CAZ) Amplifiers which allow to create acceptable values of Offset: V_{os}=1μV, drift=0.1μV/°C, are used. Fig. 4.5 shows the operating principle of such an amplifier. The CAZ is produced with CMOS technology because the switches are better, and the offset is worse than that of BJT. In the INAs, this solution is not used because the repeated switching can cause great disturbance.

Can compensate the voltage offset (first order approximation):



whole opAmp → first close A → then close B

out

↳ Vos cancel out

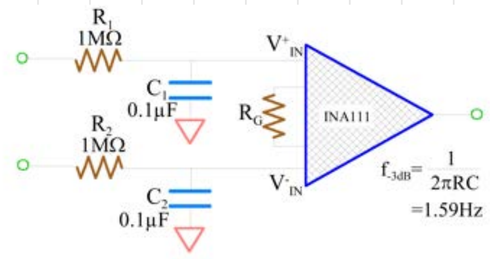
Possible causes of mismatches / errors

(Book p. 280)

INAs have a very high CMRR which should not be lowered. For example, if you want a low pass filter, it is not correct to use the scheme shown in Fig. 4.7. In fact, the degree of tolerance of resistance values is a factor of primary importance (high degree of tolerance means expensive devices!), but the capacitance values have degrees of tolerance of at least 1%. Thus, a branch has a pole 1.01% higher and the other 0.99% lower than the expected value 1/(2πRC). Hence the filter differently attenuates the input frequency, determining a differential signal that cannot be rejected despite the INA's very high CMRR.

One can use the circuit shown in Fig. 4.8 to improve the differential filtering (on the left) or the common mode filtering (on the right). For the circuit shown on the left, input resistance values must be matched with 1kΩ to avoid the waste of the CMRR performance of the INA. For the circuit shown on the right, the capacitance C3 "mitigates" the mismatch of C1 and C2 as long as C3>>C1 and C3>>C2.

Tollerances of R and C degrade CMRR !!!



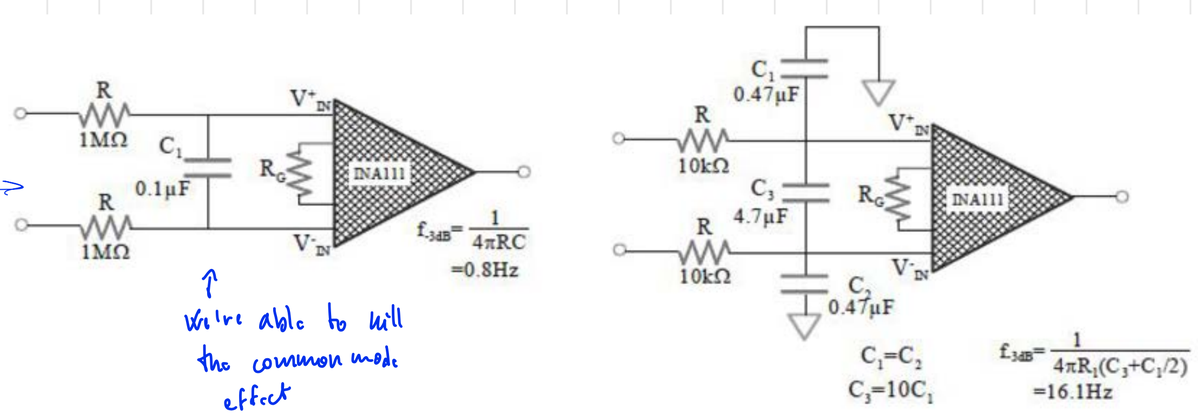
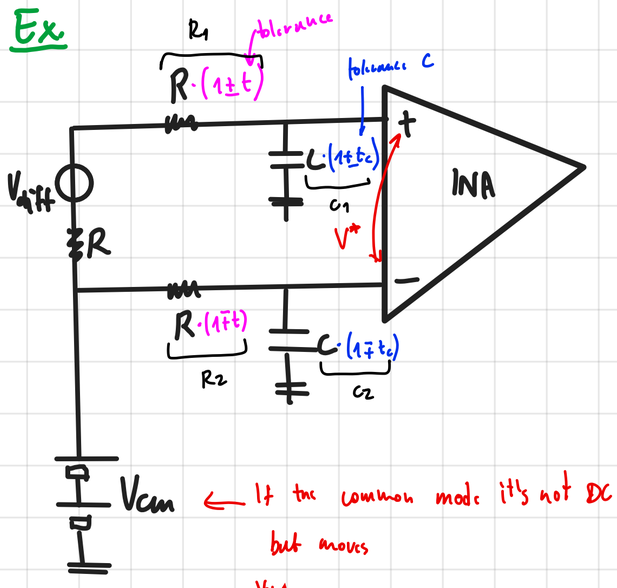


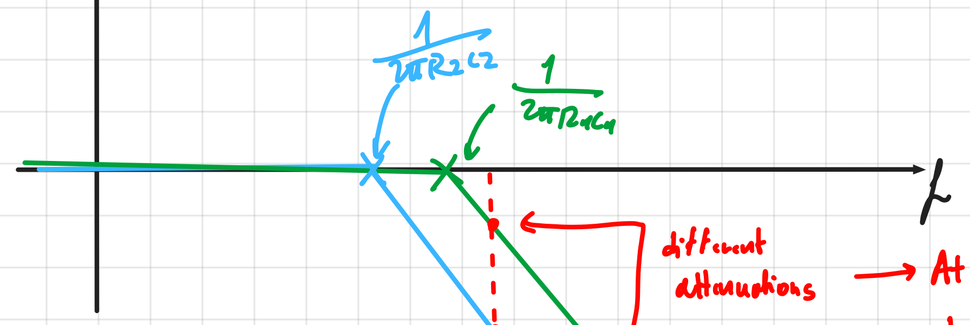
Fig. 4.8: Differential low-pass (on the left) and common mode (on the right) filtering.

Ex



← It the common mode it's not DC but moves

← we want it to be rejected by the INA → but now it's AC so it its freq. is here



→ At the input of the INA there will be a differential V+ related to the common mode signal

Ex.

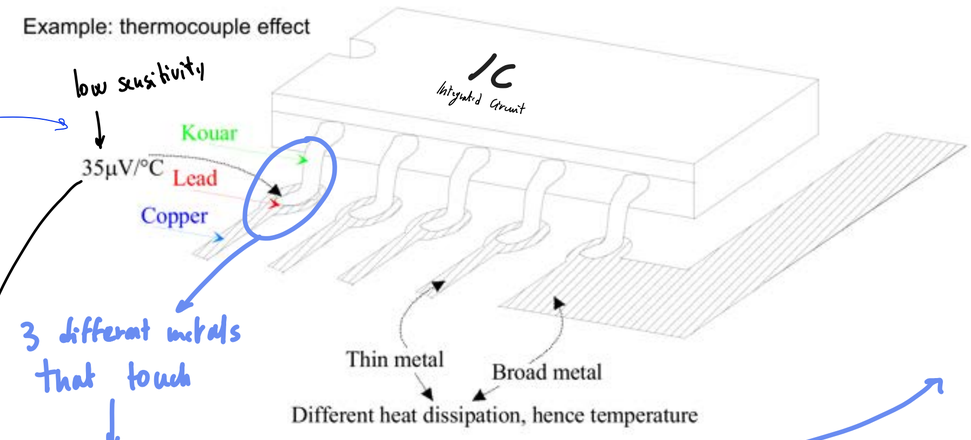
You can now consider another practical aspect of the use of these devices. The pins of the IC are an excellent "sensor" of the internal temperature of the chip, i.e. if you change the internal temperature, the output will also change and vice versa. Making precise measurements, you can see when blowing out on the INA pins that the output varies of μV or more, depending on the gain.

The welding with the track makes a thermocouple, as shown in Fig. 4.12. The worst connection is the kovlar-copper one. For this reason, some manufacturers produce copper IC pins. Nonetheless, the problem remains because the welding is made of tin. The solution is to maintain isothermal thermo-coupling to nullify the induced thermal voltage.

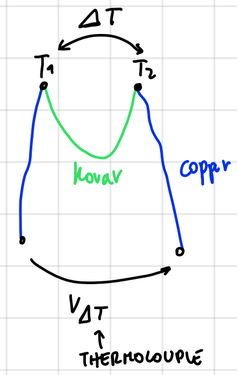
It is more important to take care of the track width, connecting all unused pins to ground or, at least, connect them to large tracks that operate as efficient heat exchangers.

Time-varying mismatches (flicker noise, drifts, offsets) sometimes due to unexpected causes:

Example: thermocouple effect

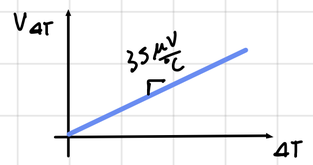


3 different metals that touch → It's a THERMOCOUPLE

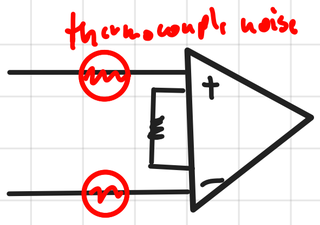


↳ given a ΔT we a VΔT (THERMOCOUPLE / Seabeck effect)

depending on the metals of the thermocouple



(W) A possible explanation of the flicker noise $\frac{1}{f}$ comes from the thermocouples formed at the pins that usually have a noise trend of $\frac{1}{f}$



Data-sheets examples

BURR-BROWN®

INA116

Ultra Low Input Bias Current INSTRUMENTATION AMPLIFIER

FEATURES

- LOW INPUT BIAS CURRENT: 3fA typ
- BUFFERED GUARD DRIVE PINS
- LOW OFFSET VOLTAGE: 2mV max
- HIGH COMMON-MODE REJECTION: 84dB (G = 10)
- LOW QUIESCENT CURRENT: 1mA
- INPUT OVER-VOLTAGE PROTECTION: ±40V

APPLICATIONS

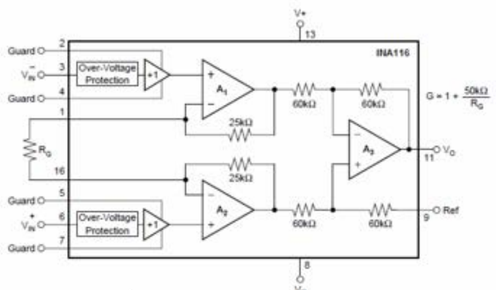
- LABORATORY INSTRUMENTATION
- pH MEASUREMENT
- ION–SPECIFIC PROBES
- LEAKAGE CURRENT MEASUREMENT

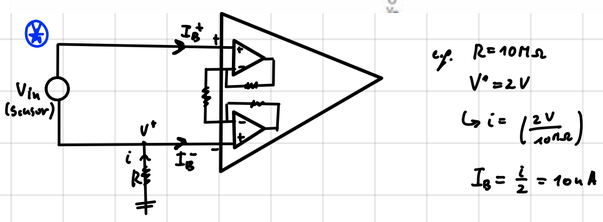
DESCRIPTION

The INA116 is a complete monolithic FET-input instrumentation amplifier with extremely low input bias current. *Difet®* inputs and special guarding techniques yield input bias currents of 3fA at 25°C, and only 25fA at 85°C. Its 3-op amp topology allows gains to be set from 1 to 1000 by connecting a single external resistor.

Guard pins adjacent to both input connections can be used to drive circuit board and input cable guards to maintain extremely low input bias current.

The INA116 is available in 16-pin plastic DIP and SOL-16 surface-mount packages, specified for the –40°C to +85°C temperature range.





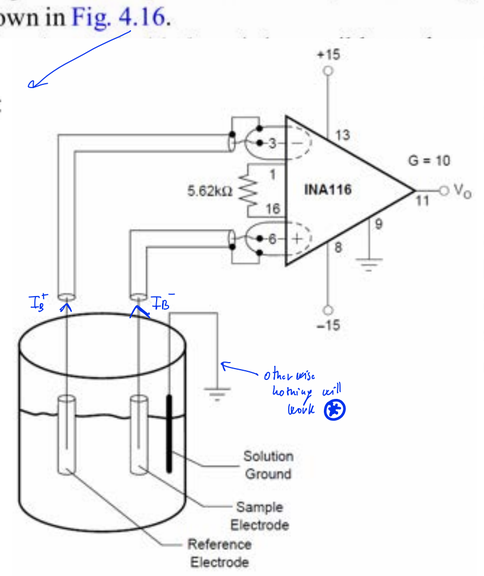
(Book p. 283)

Many INAs are provided with some internal resistors to set the gain, simply shorting the pins to the ground. Some INAs have circuitry to protect the front-end Op-Amps from overvoltage. Furthermore, there are buffers that can be used as Guard Outputs, which are connected to the coaxial cable that brings the signal (Fig. 4.13).

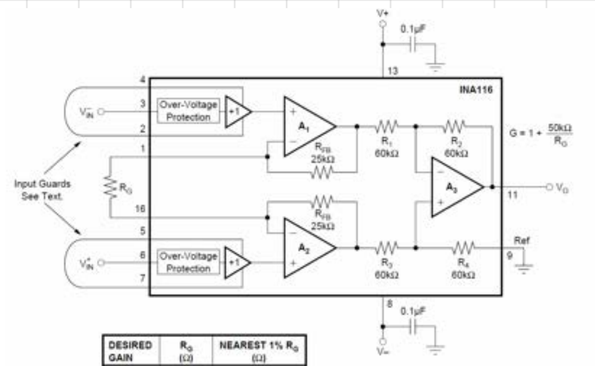
To protect the INA pins, it is a good idea to make *protection rings* on the PCB layout, as shown in Fig. 4.14. The 1MΩ resistance ensures the conductive path towards the ground for the INA input supply currents. The high value introduces a great voltage offset (due to I_B) and a great voltage noise. Parenthetically, this is a common mode input voltage and should thus be rejected by the INA.

To improve the response of the Guard with fast input signals, it is possible to add an external Op-Amp that acts as a buffer on the cable to be protected (Fig. 4.15). To ensure the loop for the input bias currents, it is always necessary to take actions, as shown in Fig. 4.16.

Always remember to provide I_B:



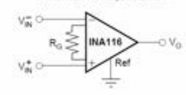
Input Protections:



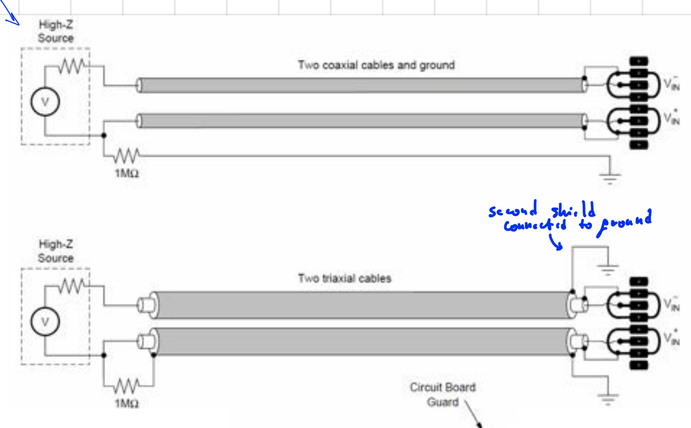
DESIRED GAIN	R_G (Ω)	NEAREST 1% R_G (Ω)
1	NC	NC
2	50.00k	49.9k
5	12.50k	12.4k
10	5.556k	5.62k
20	2.632k	2.61k
50	1.02k	1.02k
100	505.1	511
200	251.3	249
500	100.2	100
1000	50.05	49.9
2000	25.01	24.9
5000	10.00	10
10000	5.001	4.99

NC: No Connection.

Also drawn in simplified form:



Shielding:



BURR-BROWN® from Texas Instruments

INA110

SBOS147A – SEPTEMBER 1986 – JULY 2005

Fast-Settling FET-Input INSTRUMENTATION AMPLIFIER

FEATURES

- LOW BIAS CURRENT: 50pA max
- FAST SETTLING: 4µs to 0.01%
- HIGH CMR: 106dB min; 90dB at 10kHz
- INTERNAL GAINS: 1, 10, 100, 200, 500
- VERY LOW GAIN DRIFT: 10 to 50ppm/°C
- LOW OFFSET DRIFT: 2µV/°C
- LOW COST
- PINOUT SIMILAR TO AD524 AND AD624

APPLICATIONS

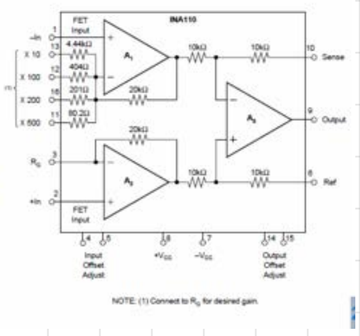
- MULTIPLEXED INPUT DATA ACQUISITION SYSTEM
- FAST DIFFERENTIAL PULSE AMPLIFIER
- HIGH SPEED GAIN BLOCK
- AMPLIFICATION OF HIGH IMPEDANCE SOURCES

DESCRIPTION

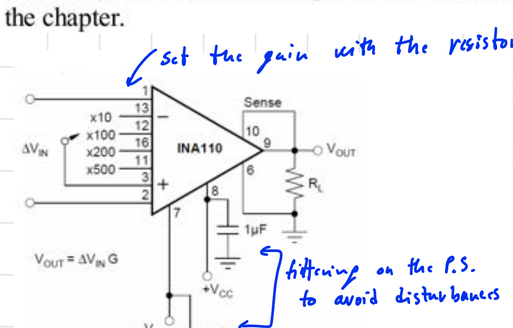
The INA110 is a versatile monolithic FET-input instrumentation amplifier. Its current-feedback circuit topology and laser trimmed input stage provide excellent dynamic performance and accuracy. The INA110 settles in 4µs to 0.01%, making it ideal for high speed or multiplexed-input data acquisition systems.

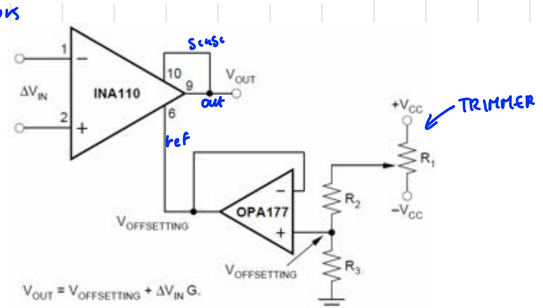
Internal gain-set resistors are provided for gains of 1, 10, 100, 200, and 500V/V. Inputs are protected for differential and common-mode voltages up to ±V_{CC}. Its very high input impedance and low input bias current make the INA110 ideal for applications requiring input filters or input protection circuitry.

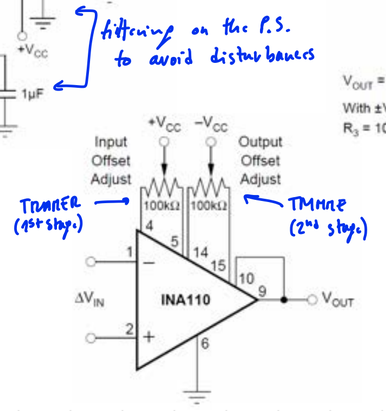
The INA110 is available in 16-pin plastic and ceramic DIPs, and in the SOL-16 surface-mount package. Military, industrial and commercial temperature range grades are available.

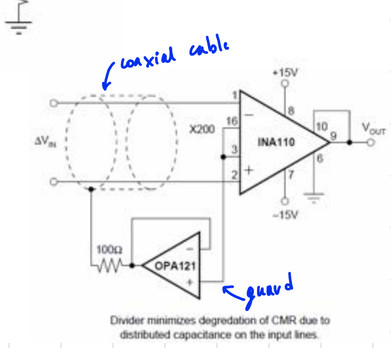


In the INA110, whose datasheet is attached below, it is possible to choose between x10, x100, x200, and x500 gains. The Guard connection (to drive the signal coaxial cable) is not present in the INA110. It is possible to put it externally, as shown in the figure of the data-sheet of the INA110 at the end of the chapter.









BURR-BROWN®

PGA206
PGA207

High-Speed Programmable Gain INSTRUMENTATION AMPLIFIER

FEATURES

- DIGITALLY PROGRAMMABLE GAINS:
  PGA206: G=1, 2, 4, 8V/V
  PGA207: G=1, 2, 5, 10V/V
- TRUE INSTRUMENTATION AMP INPUT
- FAST SETTLING: 3.5µs to 0.01%
- FET INPUT: I_B = 100pA max
- INPUT PROTECTION: ±40V
- LOW OFFSET VOLTAGE: 1.5mV max
- 16-PIN DIP, SOL-16 SOIC PACKAGES

APPLICATIONS

- MULTIPLE-CHANNEL DATA ACQUISITION
- MEDICAL, PHYSIOLOGICAL AMPLIFIER
- PC-CONTROLLED ANALOG INPUT BOARDS

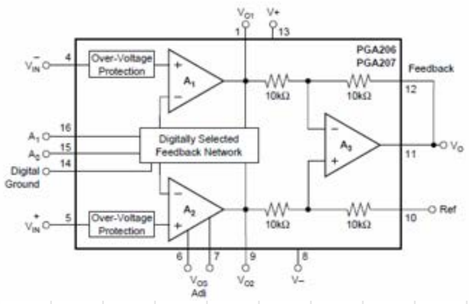
DESCRIPTION

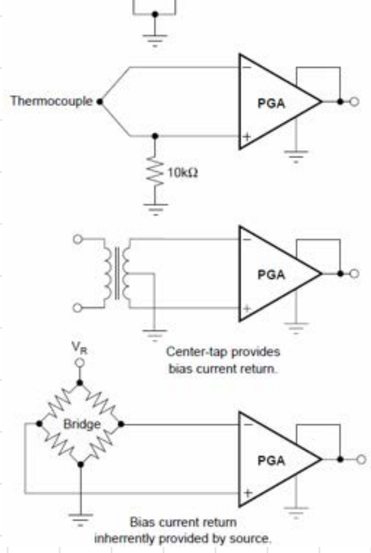
The PGA206 and PGA207 are digitally programmable gain instrumentation amplifiers that are ideally suited for data acquisition systems.

The PGA206 and PGA207's fast settling time allows multiplexed input channels for excellent system efficiency. FET inputs eliminate I_B errors due to analog multiplexer series resistance.

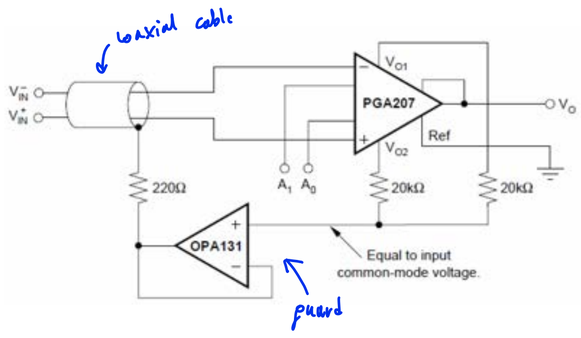
Gains are selected by two CMOS/TTL-compatible address lines. Analog inputs are internally protected for overloads up to ±40V, even with the power supplies off. The PGA206 and PGA207 are laser-trimmed for low offset voltage and low drift.

The PGA206 and PGA207 are available in 16-pin plastic DIP and SOL-16 surface-mount packages. Both are specified for –40°C to +85°C operation.

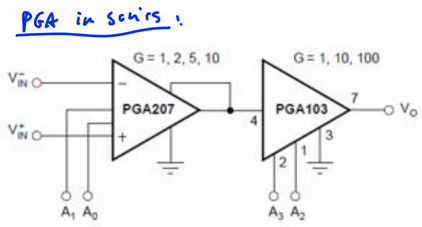




Some INAs also have offset nulling control and overvoltage protection. Moreover, other INAs have an internal network to select the gain. Often, this network can be driven through simple digital pins. Such ICs are named PROGAMMABLE GAIN AMPLIFIER, PGA. At the end of this chapter is attached a data-sheet of one of these PGAs (PGA206/207). Cascading more PGAs, it is possible to obtain different gains. Incidentally, it is necessary to pay attention to noise and bandwidth, which change with varying set gain.



PGA in series:



GAIN (V/V)	A_1	A_0	A_3	A_2
1	0	0	0	0
2	0	1	0	0
5	1	0	0	0
10	1	1	0	0
20	0	1	0	1
50	1	0	0	1
100	1	1	0	1
200	0	1	1	0
500	1	0	1	0
1000	1	1	1	0

Intro

(Book p. 298)

In this section, we will examine a particular type of OpAmp: the so-called Current Feedback Amplifiers (CFA). As we can see, these amplifiers have exceptional characteristics regarding the bandwidth and the SR. We could try to explain how these values are justifiable by looking for the causes at microelectronics device level. We will see, however, that this extraordinary performance is counterbalanced by some worsening factors with respect to the classical VOAs seen until now.

Current Feedback Amplifier (Starting from the Voltage Mode Amplifier (VOA) analysis)

(Book p.302)

Consider again the amplifier in Fig. 4.31. We can make some general considerations: it is possible to say that, since the stage has high input impedance both on the inverting and non-inverting terminals, with any feedback network, we created a "voltage feedback". It makes sense to speak of the input voltage Vdiff.

We see in the preceding chapters that because of the compensation capacitance C and the maximum current 2I of the differential stage, the limitations are attributable to the Slew Rate and the open loop bandwidth seen in the preceding chapters.

This implies that, when we implement a voltage amplifier with feedback, as depicted in Fig. 4.34, we find out strong limitations on performance. In fact, feedback imposes a relationship between Bandwidth and Gain: their product (GBWP) must be constant, i.e.:

$$\text{Gain}_{\text{closed loop}} \cdot \text{Bandwidth}_{\text{closed loop}} = \text{Bandwidth}_{\text{open loop}}$$

For these reasons, we understand that if we desire our feedback amplifier to have a high gain and wide bandwidth at the same time, we must design an operational amplifier having a high gain (to have a high G_{loop}) and wide bandwidth as well.

There are limits (related to the technology and the architecture circuit) at maximum gain and maximum GBWP, which a traditional OpAmp can have. However, since the late '80s, a new architecture of monolithic operational amplifiers known as Current Feedback Amplifiers was brought to markets.

CFA principle

(Book p.305)

Observe now the scheme of Fig. 4.37. The closed loop gain is as before:

$$G_c = 1 + \frac{R_F}{R_S}.$$

The internal structure is the same as depicted in Fig. 4.38. We find an input buffer followed by a high impedance gain node and finally an output buffer.

With a closed loop, a current comparison is made, and it generates a current error signal. As done in the VOA OpAmp, here too we must take the error signal and amplify it. In this block diagram, this signal is somewhat picked up (practically through current mirrors), then amplified, and sent on the resistance R_{OL} where a current-voltage conversion is made. Finally, it is sent out through another buffer.

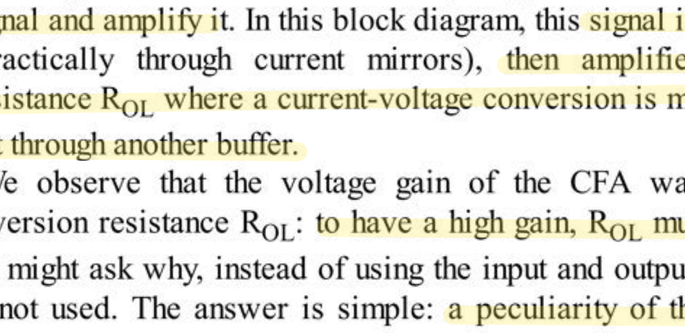
We observe that the voltage gain of the CFA was solely due to the conversion resistance R_{OL}: to have a high gain, R_{OL} must have a high value. One might ask why, instead of using the input and output buffers, gain blocks are not used. The answer is simple: a peculiarity of this CFA is to have a broadband; therefore, the internal components must be fast, and the buffers are the faster elements (think of the follower, which ideally has infinite bandwidth).

Also the use of high conversion resistance is not only due to the reasons related to the gain, but also for the reasons related to compensation. In these circuits, the Miller effect (with all the problems it entails) is not used. Here, we use a direct compensation. To obtain a high time constant, we must maximize $R_{OL} \cdot C_{COMP}$ and show how R_{OL} will be made through the collector resistor (very high resistance).

Compute the closed loop bandwidth for the circuit in Fig. 4.38. At the node 1 we have:

Fig. 4.37

$$\frac{V_1}{R_G} + \frac{V_1 - V_O}{R_F} + \frac{V_1 - V_{in}}{R_B} = 0$$



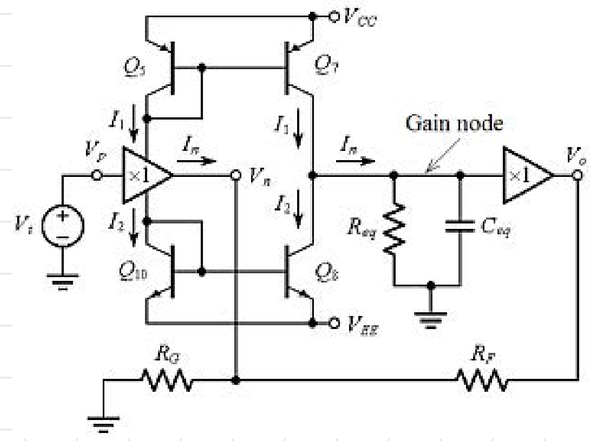
Requirements:

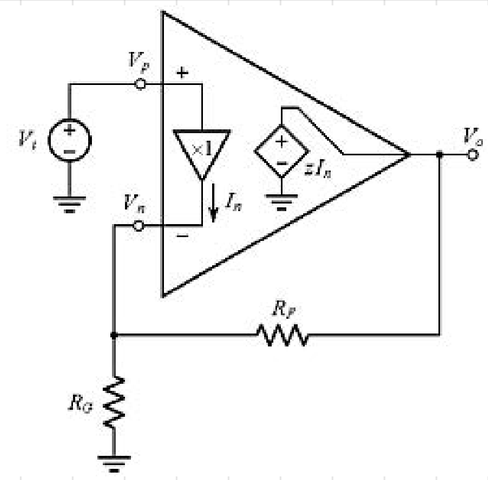
one high-impedance input	(hence voltage-driven input)
one low-impedance input	(hence input-current, i.e. Acting as an output)
voltage output	(proportional to the input current)

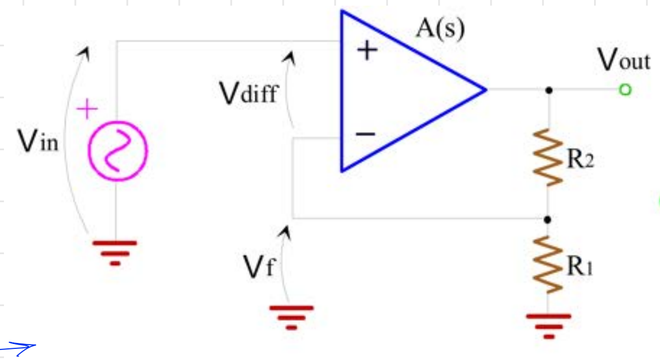
$$G_c = 1 + \frac{R_F}{R_S}$$

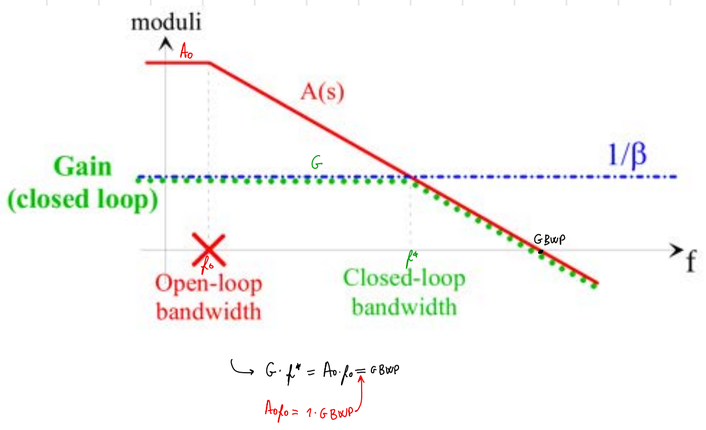
It looks identical to VOA, but completely different feedback action

(other possible scheme for following explanations)









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)

$$V_1 = \frac{\frac{V_{in}}{R_B} - \frac{V_O}{R_F}}{\frac{1}{R_F} + \frac{1}{R_G} + \frac{1}{R_B}}$$

At node 2, it results:

$$V_2 = I_2 \cdot \frac{R_{OL}}{1 + sR_{OL} \cdot C_{comp}}$$

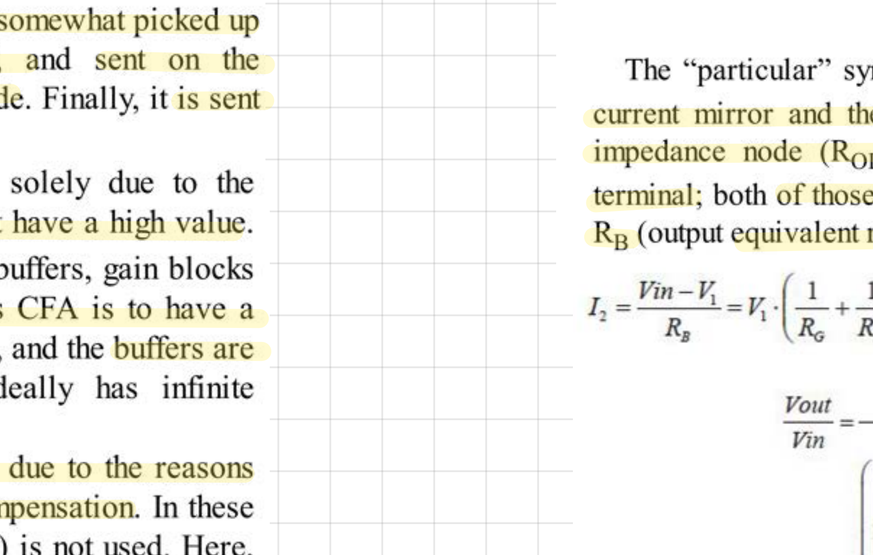
The "particular" symbolism highlights that I_2 is originated from the input current mirror and then the two branches, one of which goes in the high impedance node (R_{OL}), and the other goes out of the OpAmp negative terminal; both of those have the same current. The current that flows through R_B (output equivalent resistance of the input buffer) is

$$I_2 = \frac{Vin - V_1}{R_B} = V_1 \cdot \left(\frac{1}{R_G} + \frac{1}{R_F}\right) - \frac{Vout}{R_F} \quad \text{and because} \quad V_{out} = A_{out} \cdot V_2 \cong V_2 \quad \text{we can extract:}$$

$$\frac{Vout}{Vin} = \frac{1 + \frac{R_F}{R_G}}{\left(1 + \frac{R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B}{R_{OL} \cdot A_{out}}\right) \cdot \left[1 + s\frac{\left[R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B\right] \cdot G_{loop}}{A_{out} + \frac{R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B}{R_{OL}}}\right]}$$

The gain tends to the ideal value as much as $R_{OL} \cdot Aout \to \infty$, as for the "voltage mode" OpAmp. The expression of the time constant seems more complex to analyze. However, considering the case with great R_{OL}, we have:

Fig. 4.38



$$I_2 = \frac{Vin - V_1}{R_B} = V_1 \cdot \left(\frac{1}{R_G} + \frac{1}{R_F}\right) - \frac{Vout}{R_F} \qquad V_2 = I_2 \cdot \frac{R_{OL}}{1 + sR_{OL} \cdot C_{comp}} \qquad \frac{V_1}{R_G} + \frac{V_1 - V_O}{R_F} + \frac{V_1 - V_{in}}{R_B} = 0$$

With open-loop: $pole_{openloop} \approx \dfrac{-1}{C_{comp} \cdot R_{OL}}$ $\qquad G_{loop} \approx -R_{OL} \cdot A_{out} \cdot \dfrac{1}{R_B + R_F} \approx -\dfrac{R_{OL}}{R_F}$

… hence, when loop is closed:

$$pole_{closedloop} = pole_{openloop} \cdot (1 - G_{loop}) \approx \frac{-1}{C_{comp} \cdot R_{OL}} \cdot \frac{R_{OL}}{R_F} = \frac{-1}{C_{comp} \cdot R_F}$$

$$f_{pole} \cong \frac{A_{out}}{2\pi \left[R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B\right] \cdot C_{comp}}$$

- For low gain ($R_B \ll \frac{R_F \cdot R_G}{R_F + R_G} = R_F \parallel R_G$) we get $f_{pole} \cong \dfrac{1}{2\pi \cdot R_F \cdot C_{comp}}$

Bandwidth depends only on R_F, not on R_G, hence NOT ON GAIN !

depending only on the feedback resistance and on C_{comp} while it is independent of R_G and thus of the closed loop gain! Therefore, the choice of R_F is more important in this case than in the "voltage mode" OpAmp case.

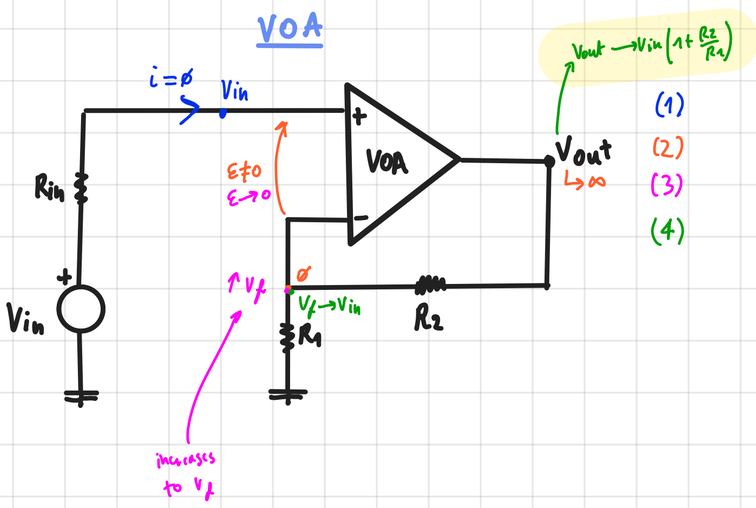
- Instead for high gain (>50) we get $GBWP = \dfrac{A_{out}}{2\pi R_B \cdot C_{comp}}$

As for VOA, again trade-off Gain &Bandwidth (i.e. constant GBWP)

↳ **Recap:** CFA improvement wrt VOA

same gain

VOA



(1)
(2)
(3)
(4)

- $G = 1 + \frac{R_2}{R_1}$
- $\beta = \frac{R_1}{R_1+R_2} = \frac{1}{1+\frac{R_2}{R_1}}$ G

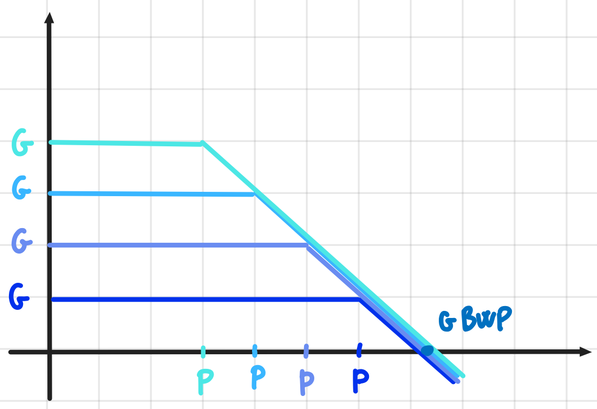
β related to the gain

↳ change the gain (R_1 or R_2 change)

↳ change β

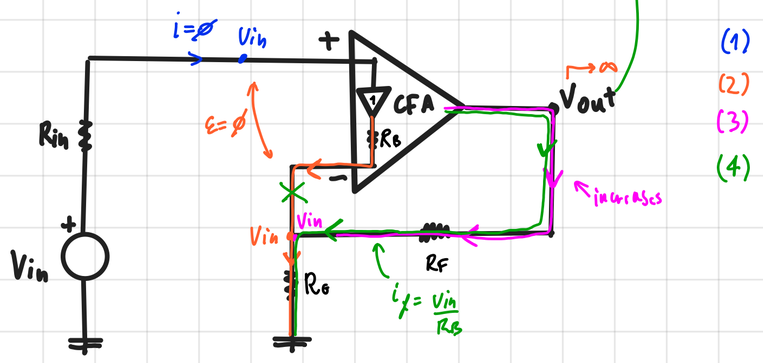
↳ change the pole

↳ $GBWP = G \cdot pole = const$



$G \cdot P = const = GBWP$

CFA



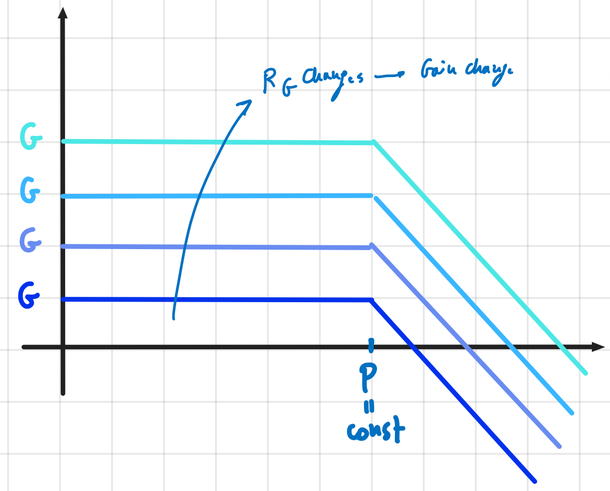
(1)
(2)
(3)
(4)

- $G = 1 + \frac{R_F}{R_G}$
- $\beta \propto R_F$ ← does NOT depend on R_G

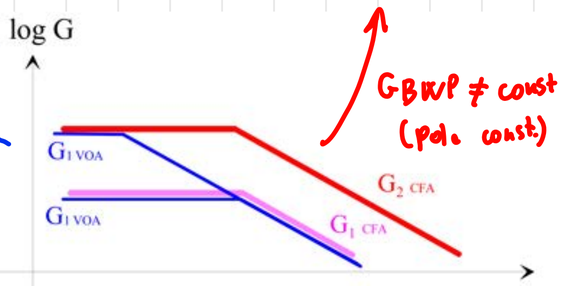
↳ change the gain (by changing R_G)

↳ G_{loop} does NOT change

↳ pole does NOT change



GBWP = const.



$GBWP \neq const$ (pole const.)

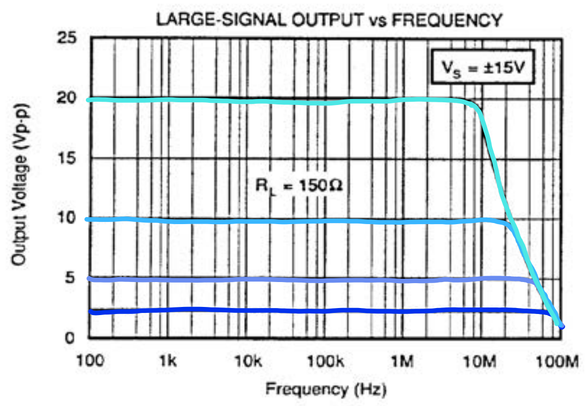
- In conclusion, at low gains...

(CFA ≠ VOA)

- Instead at high gains... ← if we increase Rf too much the pole start to reduce

(CFA = VOA)

(GBWP = const)



BJT : RECAP

BJT Structure

- The BJT is constructed with three doped semiconductor regions (**emitter**, **base**, and **collector**) separated by two pn junctions.
- One type consists of two n regions separated by a p region (npn), and the other type consists of two p regions separated by an n region (pnp).
- The term bipolar refers to the use of both holes and electrons as current carriers in the transistor structure.

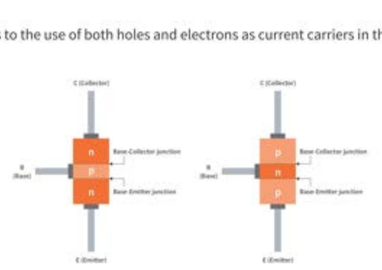


Figure 1: Basic BJT structure

- The pn junction joining the base region and the emitter region is called the *base-emitter junction.*
- The pn junction joining the base region and the collector region is called the *base-collector junction.*
- A wire lead connects to each of the three regions.
- The leads are labeled E, B, and C for emitter, base, and collector, respectively.
- The base region is lightly doped and very thin compared to the heavily doped emitter and the moderately doped collector regions.

Operation: Inside the npn structure

- The heavily doped n-type emitter region has a very high density of conduction-band (free) electrons.
- These free electrons easily diffuse through the forward-based BE junction into the lightly doped and very thin p-type base region.
- The base has a low density of holes, which are the majority carriers.
- A small percentage of the total number of free electrons injected into the base region recombine with holes and move as valence electrons through the base region and into the emitter region as hole current.

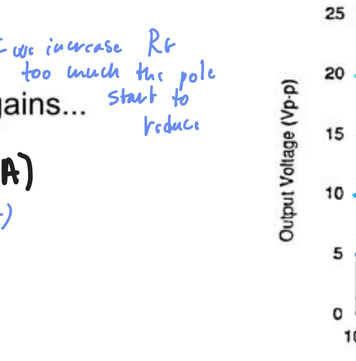


Figure 4: BJT operation showing electron flow

- When the electrons that have recombined with holes leave the crystalline structure of the base, they become free electrons in the metallic base lead and produce the external base current.
- As the free electrons move toward the reverse-biased BC junction, they are swept across into the collector region by the attraction of the positive collector supply voltage.
- The free electrons move through the collector region, into the external circuit, then return into the emitter region along with the base current.
- The emitter current is slightly greater than the collector current because of the small base current that splits off from the total current injected into the base region from the emitter.
- The operation of the pnp is the same as for the npn except that the roles of the electrons and holes, the bias voltage polarities, and the current directions are all reversed.
- I_B is through the base-emitter junction because of the low impedance path to ground and, therefore, I_C is zero

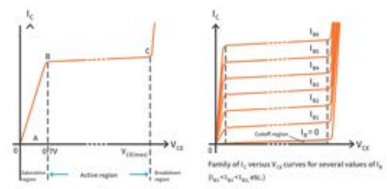


Figure 9: Collector characteristic curves

- When both junctions are forward-biased, the transistor is in the *saturation* region of operation. **Saturation** is the state of a BJT in which I_C has reached a maximum and is independent of I_B.

1. As V_{CC} is increased, V_{CE} increases as I_C increases. This is the portion between points A and B in Fig. 9. I_C increases as V_{CC} is increased because V_{CE} remains less than 0.7 V due to the forward-biased base-collector junction.

- When V_{CE} exceeds 0.7 V, the base-collector junction becomes reverse-biased and the transistor goes into the *active*, or **linear**, region of operation.

1. I_C increases very slightly for a given I_B as V_{CE} increases due to widening of the base-collector depletion region. This causes a slight increase in β_{DC}.
2. This is the portion between points B and C in Fig. 9. I_C in this portion is determined only by $I_C = \beta_{DC} I_B$.

- When V_{CE} reaches a sufficiently high voltage, the base-collector junction goes into *breakdown*; and I_C increases rapidly, shown by the portion to the right of point C. A transistor should never be operated in this region.
- A family of curves is produced when I_C versus V_{CE} is plotted for values of I_B. When $I_B = 0$, the transistor is in the *cutoff* region. **Cutoff** is the nonconducting state of a transistor.

- **BJT with multi-emitter**

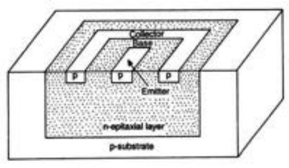
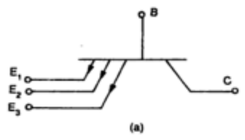


Fig. 1.17 A *pnp* lateral transistor

Transistor with multiple emitters: The applications such as transistor- transistor logic (TTL) require multiple emitters. The below figure shows the circuit sectional view of three N-emitter regions diffused in three places inside the P-type base. This arrangement saves the chip area and enhances the component density of the IC.



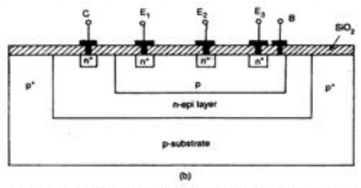
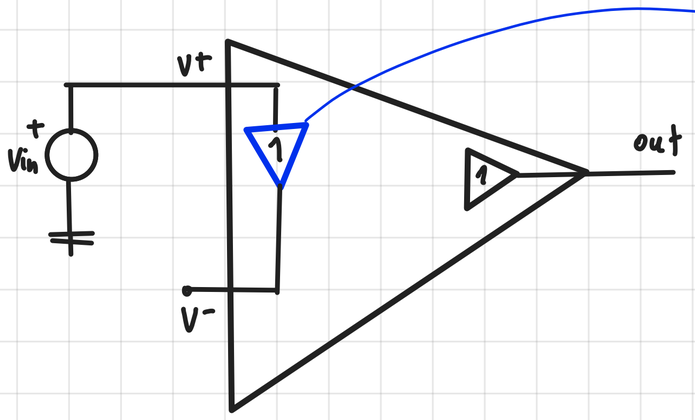


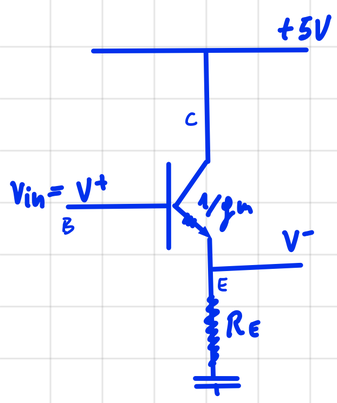
Fig. 1.18 (a) Multi-emitter transistor, (b) Cross-sectional view of a multi-emitter transistor

REAL ARCHITECTURE

- BUFFER COMPONENT:



Possible real architecture:



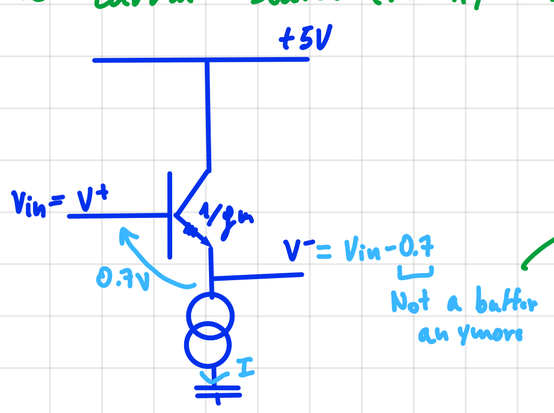
→ for $R_E >> \frac{1}{g_m}$

↳ $V^- = V_{in} \cdot \frac{R_E}{R_E + \frac{1}{g_m}} \simeq V_{in}$ (BUFFER)

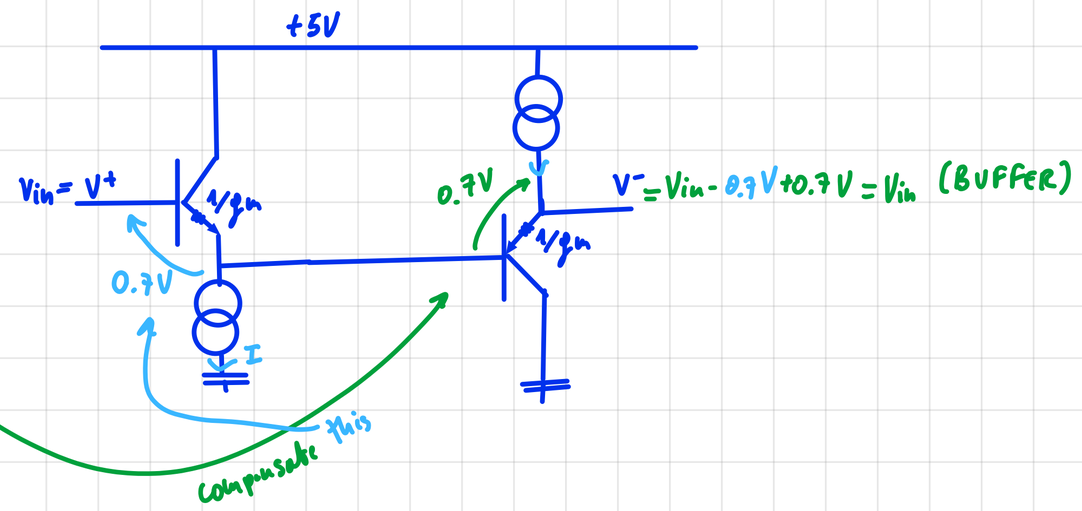
↳ but since we have current passing through R_E it will mean that we have a huge voltage on it

↳ not possible

We can consider instead of a very big R_E a current source (ideally its resistance → ∞)

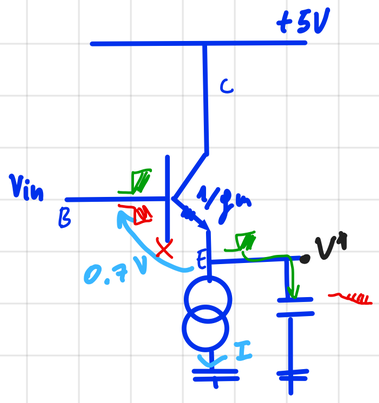


We can add this part



↳ But now we have to analyze the transistors' behaviour wrt the Vin trend, indeed we know that transistors allows current in one direction and for certain input conditions, while they block the behaviours in opposite conditions

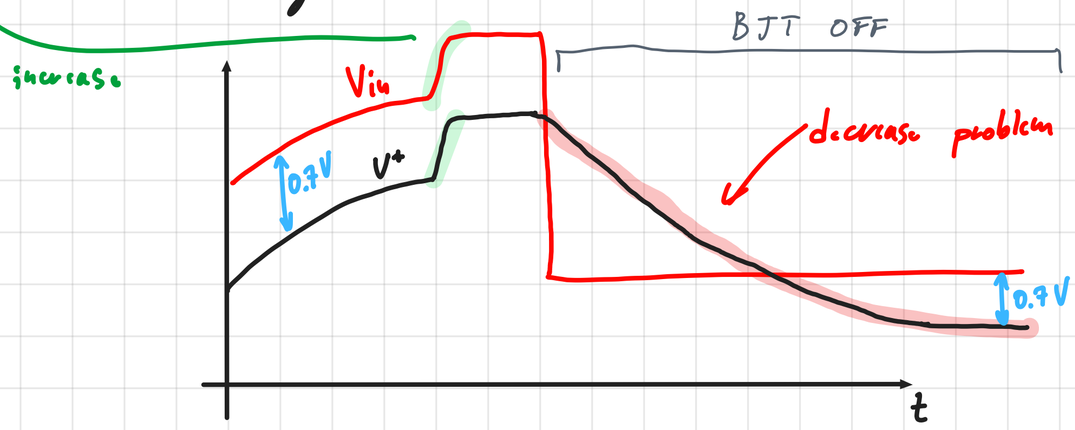
↳ Indeed we have that the npn BJT works well when Vin is increasing, indeed if we put a capacitor C:



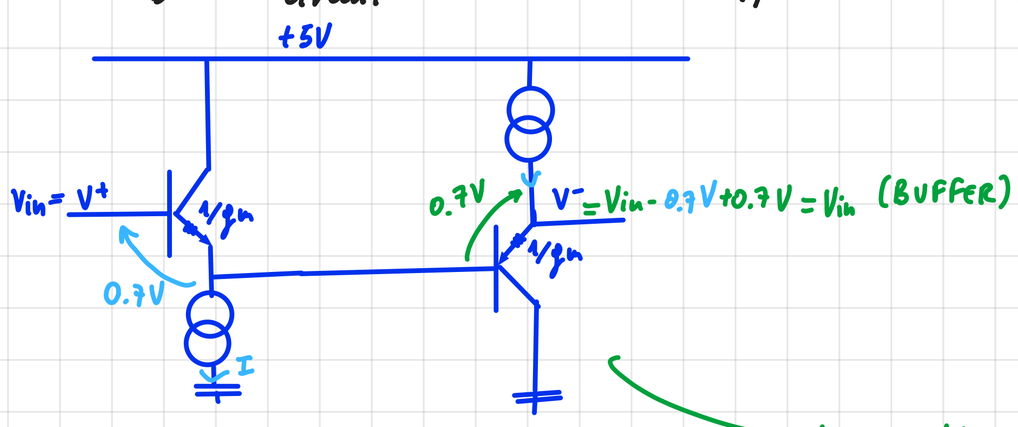
→ The capacitor is quickly charged by the transistor → PULL-UP stage

→ If we instead decrease Vin → we get to the point where $V_B - V_E < 0.7$ → so the transistor opens

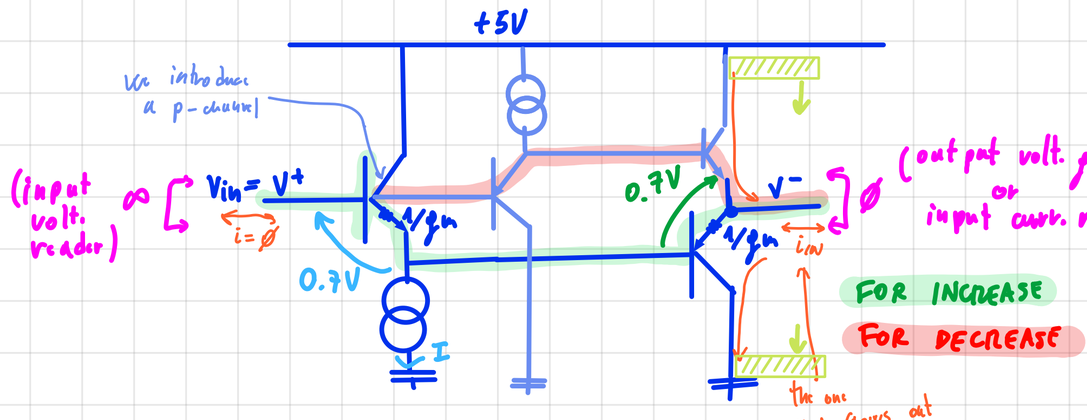
↳ on the capacitor side it stays V_E an slowly discharge, while maybe on the $V_{in} = V_B$ is decreasing much faster



↳ So this buffer circuit behaves well only for increasis:



We add simmetry to the circuit (darlington config.)

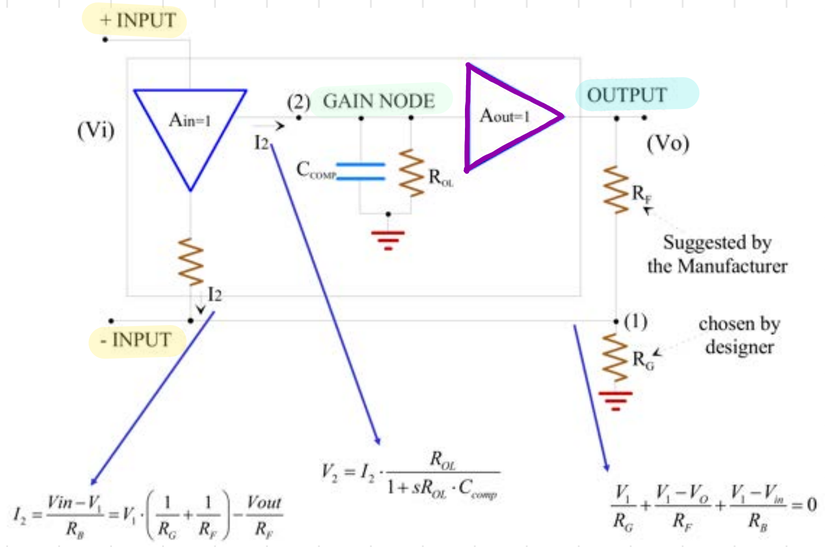


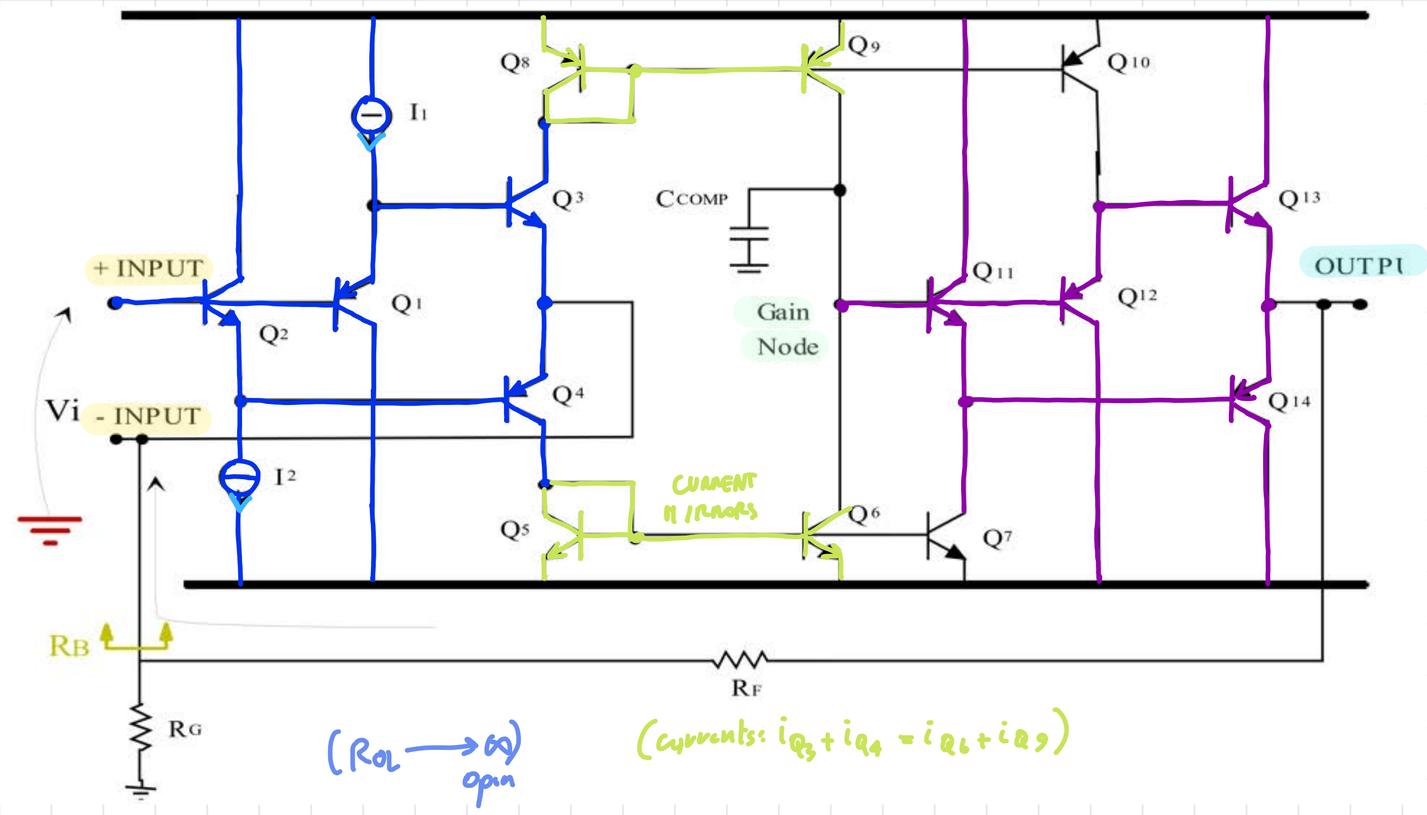
in this way the current gets copied on the other mirror

We can put current mirrors

to use this current that:
- can come from the up-side
- can come from the down-side
- can come from both sides

on the real structure





- Comparison with VOA

(Book p. 298)

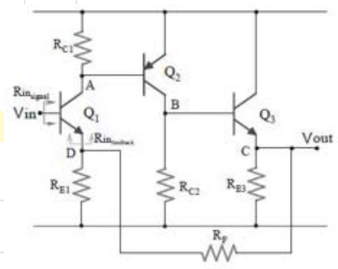
The simplified structure of a classical voltage-mode amplifier is shown in Fig. 4.31. Now instead let us study a different feedback network, as the one shown in Fig. 4.32. The first and the second stages are *common emitters* while the third is an *emitter follower*. The first stage output resistance is about R_{C1} because we suppose that r_o of Q_1 is reasonably higher ($Rout_1 = R_{C1}$)

Let us think **in terms of currents**. The first stage produces a current proportional to the input voltage signal (*transconductance stage*). This second stage output current can follow three paths: re-flows into r_o of Q_1; flows through R_{C1}; enters into the base of Q_2. According to the hypothesis, the first path will be followed by a current negligible with respect to the currents flowing through the remaining paths (reasonably, the bipolar Early resistance will be greater than discrete resistance).

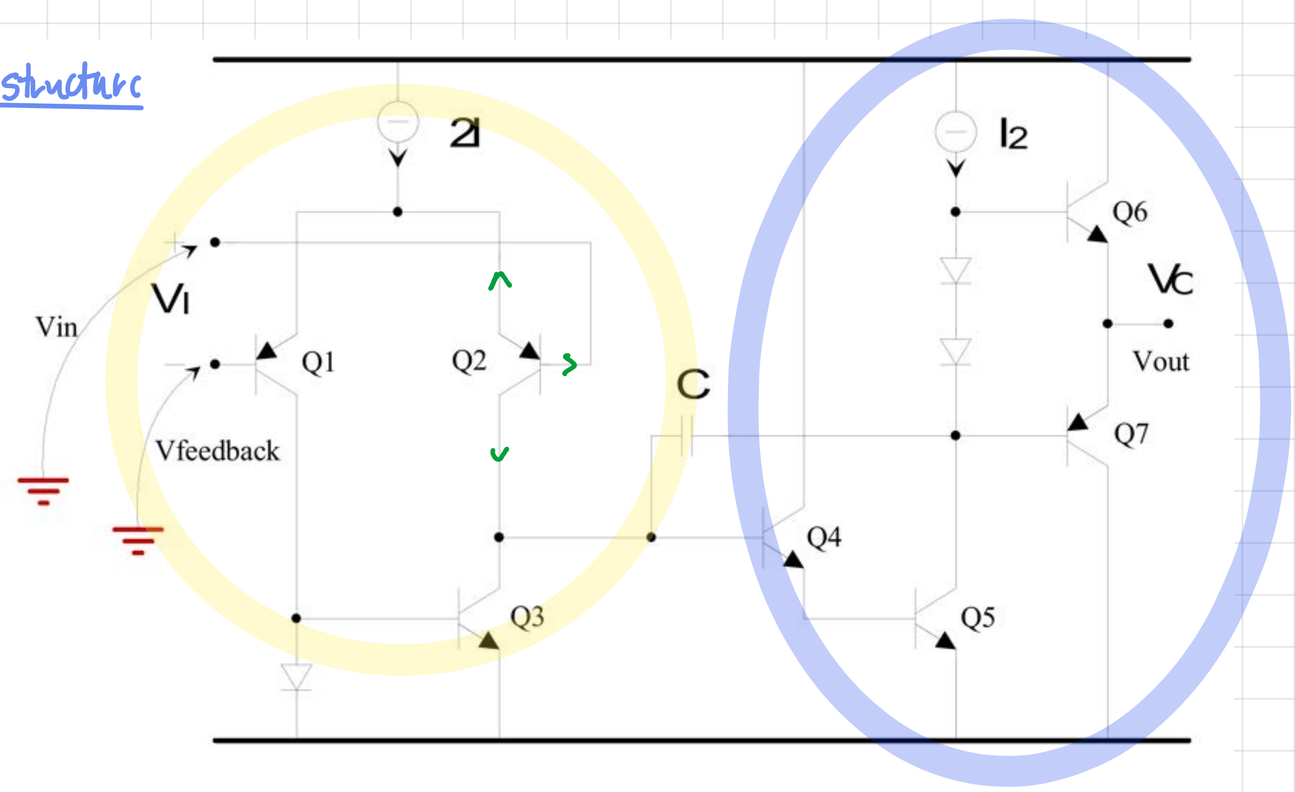
At this step, we should ask which alternative this current should follow. Naturally, the answer is obvious: the second stage works well if it receives the input current. The more is the current flowing into the input, the more is the current produced as the output; this current should be sent to a load to produce a voltage across the load. To send more current as the input of the second stage, we must have $\mathbf{RIN_{2°STAGE} < Rout_1}$.

The different reasoning modes give contrasting conclusions. But then, to maximize the gain, is it convenient to choose VOLTAGE MODE or CURRENT MODE approach? If we make some calculations on the studied structure (neglecting the emitter degeneration for the common emitter stages), we can find that we have the optimal condition when $Rout_1$ is much greater than $RIN_{2°STAGE}$.

As has just been said, it is wrong to proceed with voltage mode. When you design with voltage mode, people tend to turn the current into voltage soon; in this mode, already at the first stage output, we would like a high-impedance node to immediately transform the input transistor collector current into a voltage signal. However, that takes us away from the optimal solution in terms of gain.



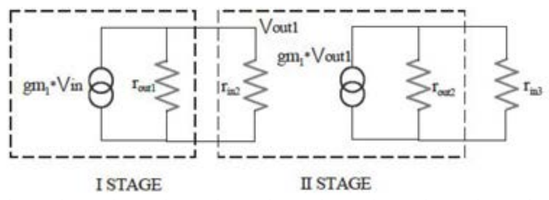
VOA structure



Requirements:

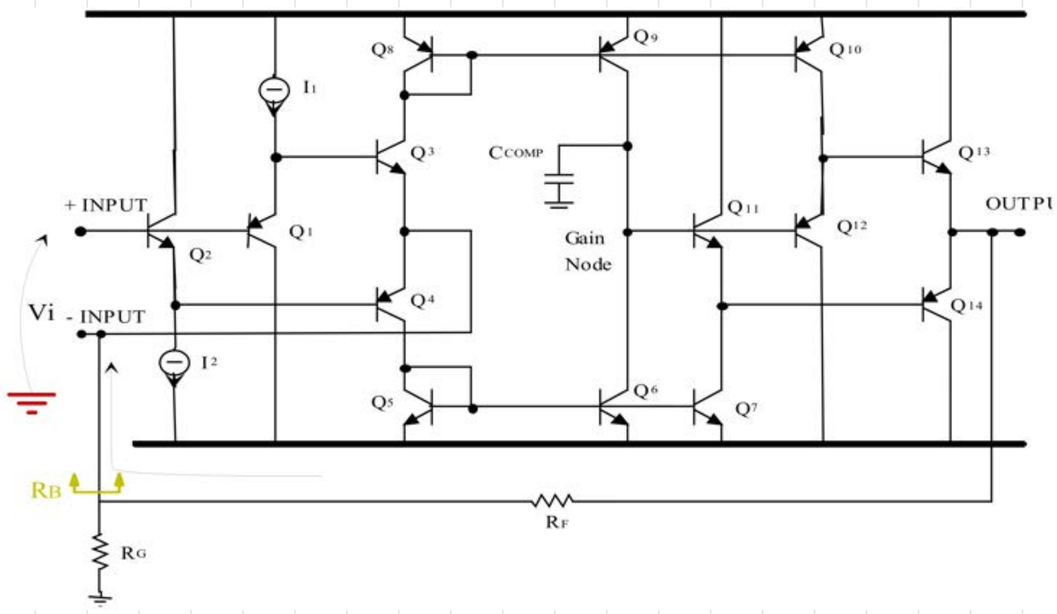
voltage-driven inputs → high impedance inputs

voltage output → low-impedance output



(Book p.311)

CFA structure



We could find in fact:

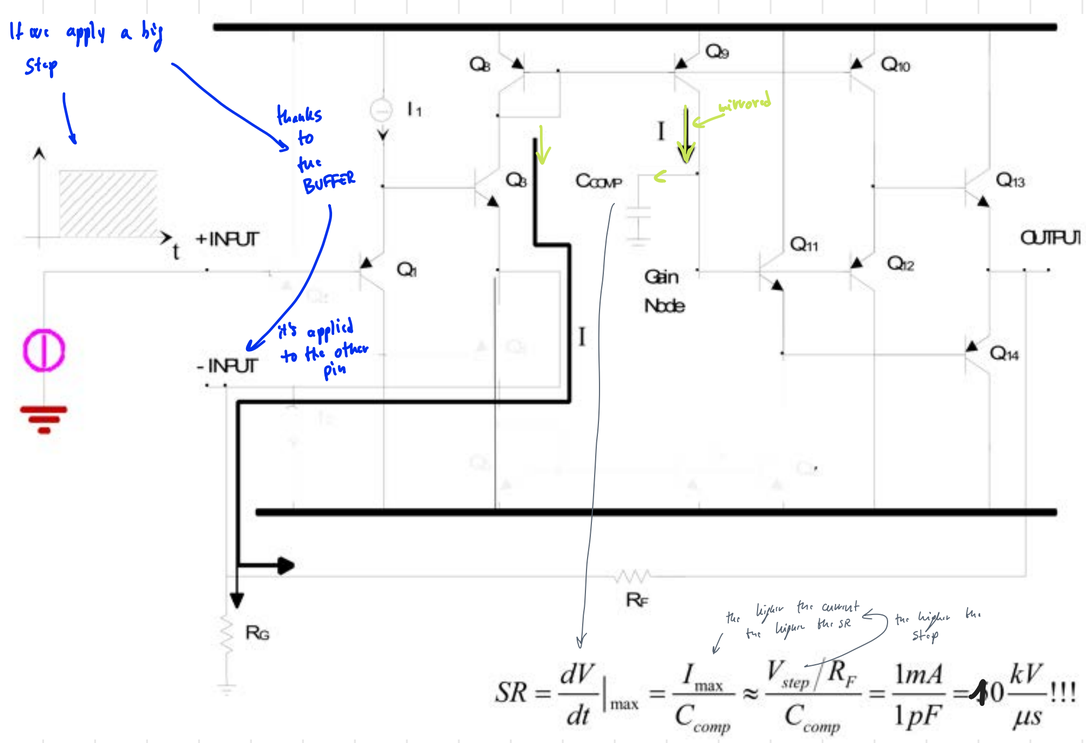
- a buffer-like structure as an input;
- the current mirror scheme to read the error signal;
- inverting and non-inverting inputs;
- gain node with compensation capacitance;
- output buffer, like the input buffer.

As an input circuit, we find a peculiar circuit that we can identify as a kind of complementary Darlington Push-Pull. There are at least two good reasons for its presence: to improve the **OFFSET** voltage and improve the **SLEW RATE**. Let us see how it works. The input stage acts as a buffer. Suppose that we apply a voltage step to the non-inverting input (Fig. 4.42). The input transistor Q_2 acts as a follower and quickly copies the voltage V_{in} at the inverting node. (Recalling the properties of the follower, we immediately realize that during the transient if the signal increases, also the gm of the transistor increases, and consequently the performance of the buffer improves.) Since the output V_{out} has not enough time to react, the buffer should provide a current to R_G and R_F. This current is drawn from the emitter of Q_3 (or Q_4 if the step input is decreasing) and brought to the current mirror through the collector of that transistor and goes on to the high impedance gain node.

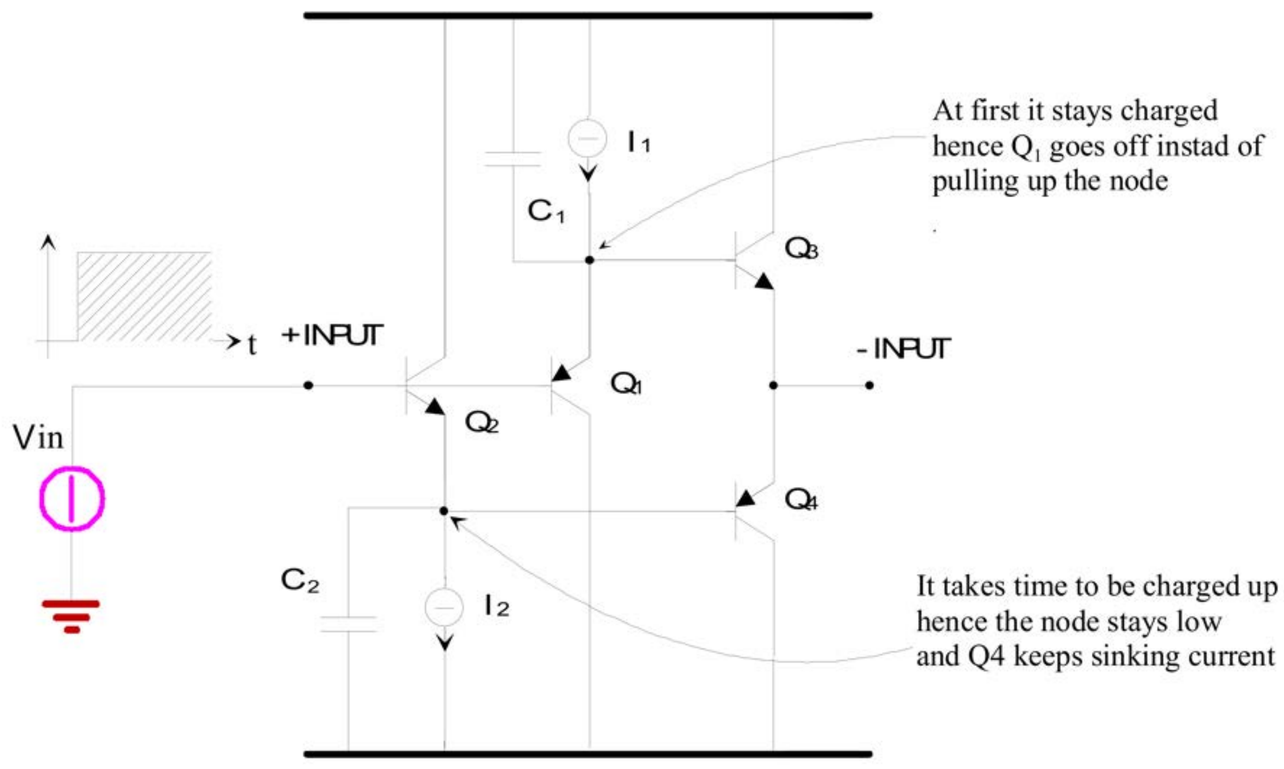
If, for example, we suppose to give a 1V step as an input and that R_F is 1KΩ (typical value for most practical cases), such a current has a value of 1mA. In first approximation, doubling the step doubles the current that flows to charge C_{comp}. If C=1pF, in the case of 1V step, we should have a slew rate equal to:

$$SR = \frac{dV}{dt}\Big|_{max} = \frac{I_{max}}{C_{comp}} = \frac{1mA}{1pF} = 10\frac{kV}{\mu s}!!!$$

Actually, there are various second order effects which limit the Slew Rate, such as the parasitic capacitors connected to the input buffer nodes, shown in Fig. 4.43.



Indeed there are second-order effects, which degrade SR



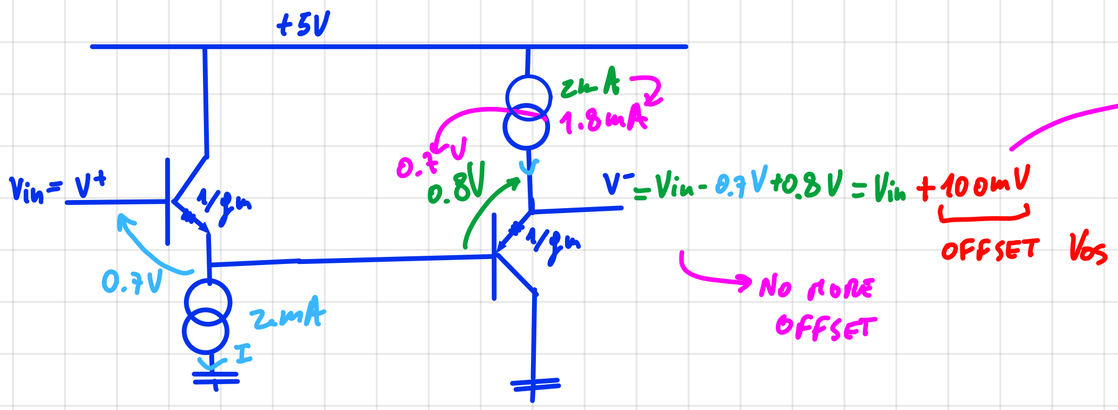
A positive step applied to the positive terminal determines the charge of the parasitic capacitance C_2 and the discharge of C_1. The voltage of this capacitance is copied on the other input (the output of the input buffer). To reduce this effect, CFAs are designed with "high" current (hundreds of μA or even mA) in the input stage. Even the current mirrors have parasitic capacitors and can reach the saturation due to current crowding (for high currents caused by a large step input).

Suppose that the additional path provided by Q_1 is NOT present. If we suppose to give a negative step input to turn off the current of the Q_2 *npn*, its transconductance decreases, and the system struggles to respond to a large input signal. We cannot speak of time constants because the system is operating in the non-linear regime, but the delay is equivalent to that due to the increasing signal. In the extreme case, if I give a great step, the base of Q_2 suddenly falls while the emitter sees parasitic capacitors and could not instantaneously follow the base (however small the emitter resistance may be). Eventually transistor Q_2 enters interdiction region, i.e. its transconductance decreased down to zero. It means that the transistor emitter load network is free to evolve with the time constants given by the same capacitors as before, but no longer sees $1/g_m$ because the transistor is off, but the parasitic resistance is large, so time constants become very long. This is a very high non-linearity. To avoid this, we could use a push-pull at the output for the same reason.

The big advantages of the CFA in terms of bandwidth and slew rate are evident in the table below. You can also notice the larger offset voltage.

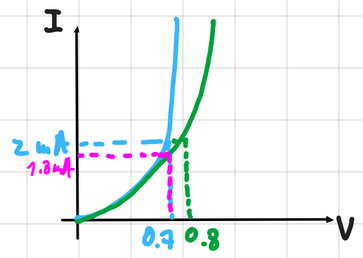
- **Offsets**

Offsets may come from the different thresholds of the transistors



we can try to reduce it considering a diff. current on the second transistor (so small V_{BE})

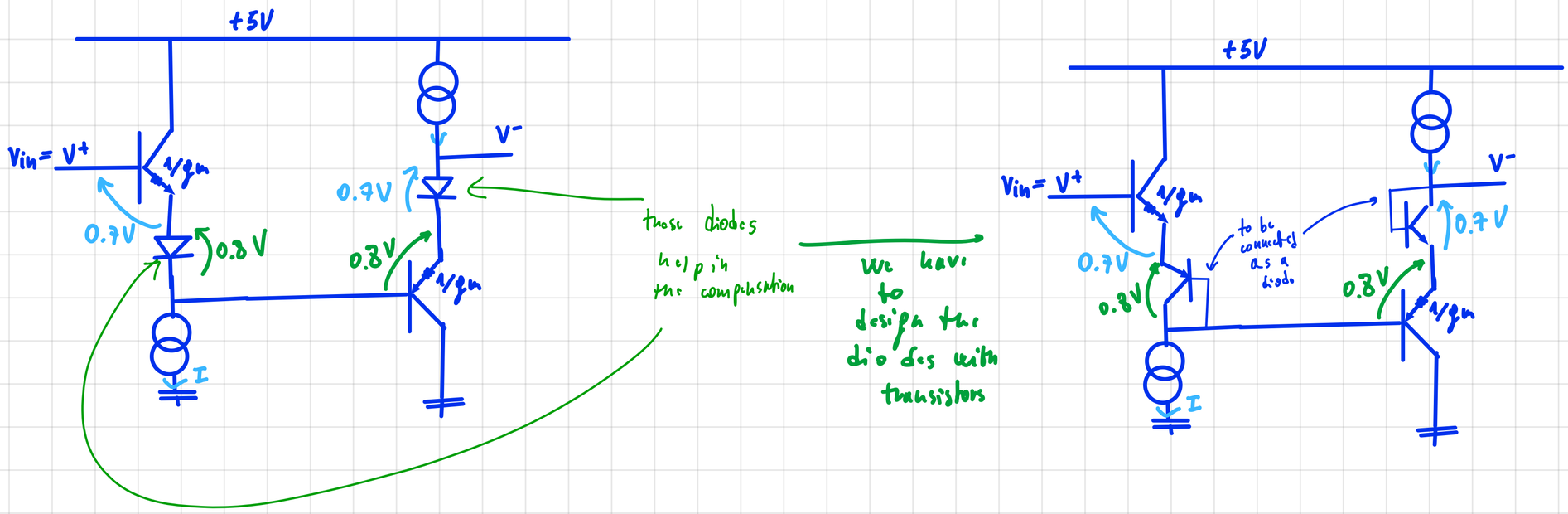
e.g. instead of 2mA - 0.8V
1.8mA - 0.7V



↳ But in reality the transistor characteristics change also for temperature changes → so we still have offsets!

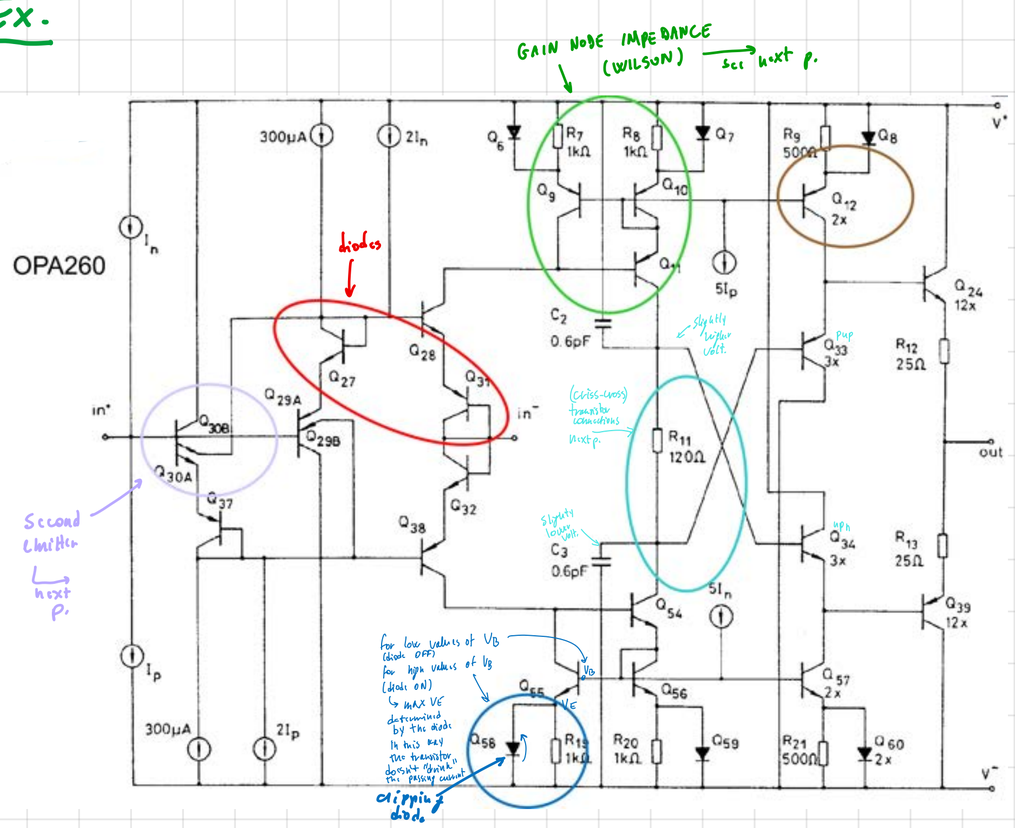
→ So we know that these offset are due to the mismatches b/w the npn and the pnp transistors

↳ A possible solution could be to use more transistors:



↳ the offset cannot be removed BUT REDULED

Ex.

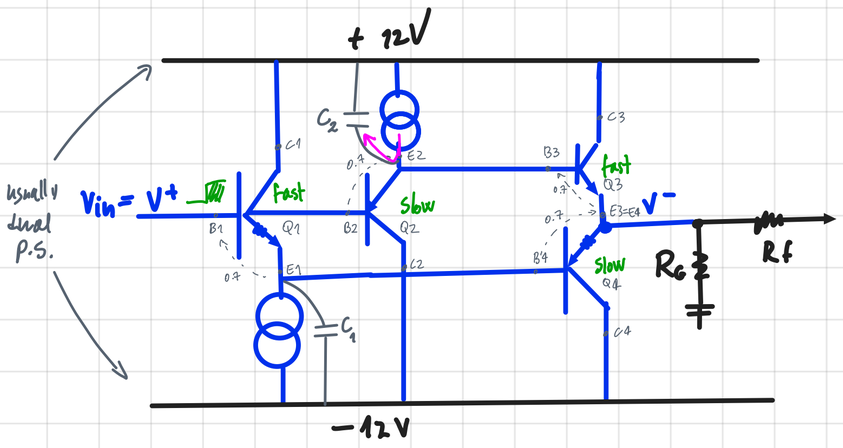


An improved scheme is that of the CFA OP260 shown in Fig. 4.46. At the input, there are four supplementary diodes (Q37, Q27, Q31, and Q32) to counterbalance the drops of VBE_{pnp} and VBE_{npn}: the voltage offset is reduced, but the output impedance is doubled (i.e. the bandwidth is constant up to a gain of about half of that obtainable without additional diodes). The auxiliary emitters improve the CFA behavior for great signals as an input because they directly drive the current mirrors. The transresistance R_{OL} is equal to 7MΩ and it is possible to use RF up to 2.5kΩ

From what we have seen so far, it can be concluded that the CFA is advantageous from the AC point of view (bandwidth, speed, and slew rate), but not from the DC point of view (accuracy, offset, and precision). Thus, in the case of ADC conversion at high frequencies, the CFA is not suitable because it is inaccurate. Typical CFA offsets vary between 2÷15mV, so it is hard to have systems with more than 8÷10 bit.

- **Other issues**

- Consider:

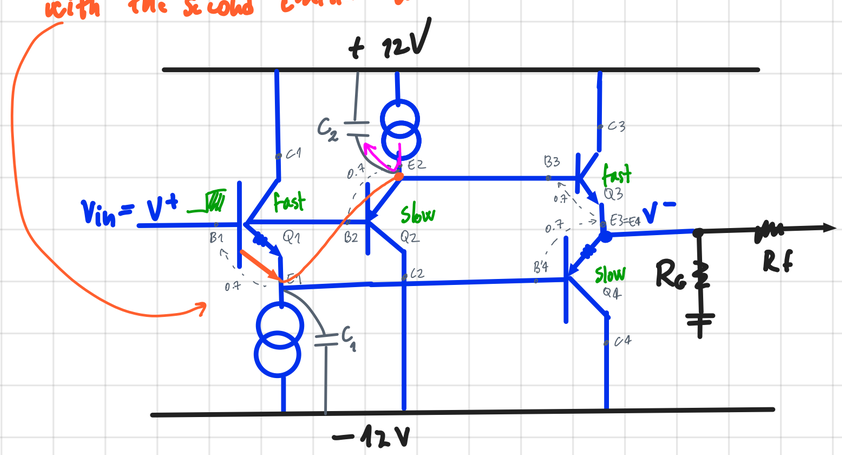


→ for an increase of the npn are fast the pnp are slow

↳ when $V_{B2} - V_{E2} < 0.7$ → Q2 OFF → the current discharge the parasitic C_2

↳ To speed up the circuit we can use a double emitter transistor

with the second emitter we touch the node E

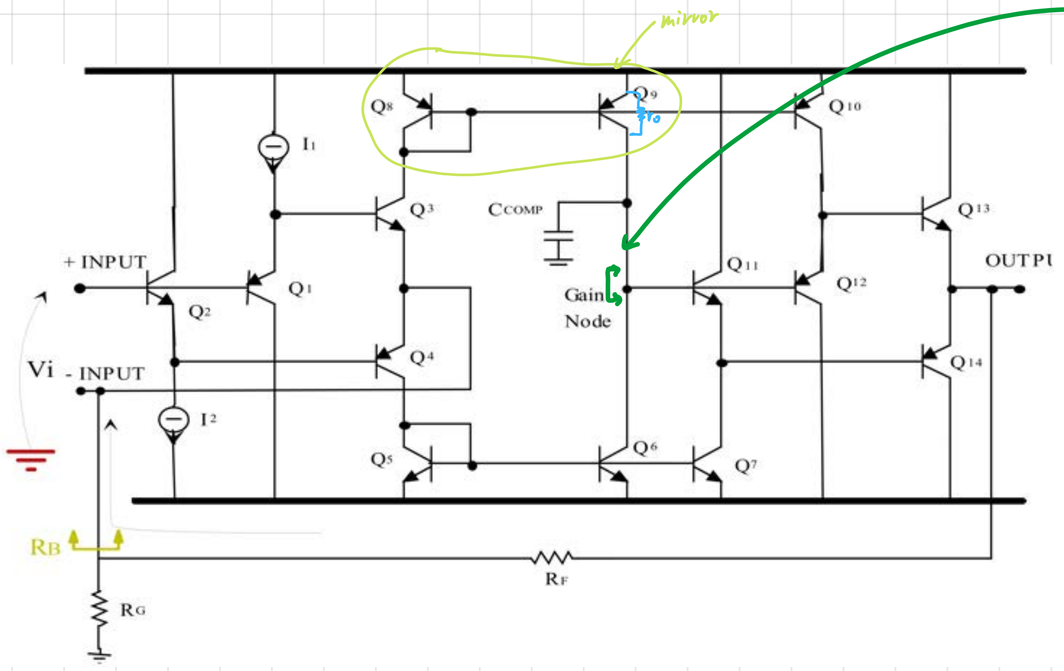


→ Now we QUICKLY pump current into that node

letting E2 quickly increase

(we increase also the SR)

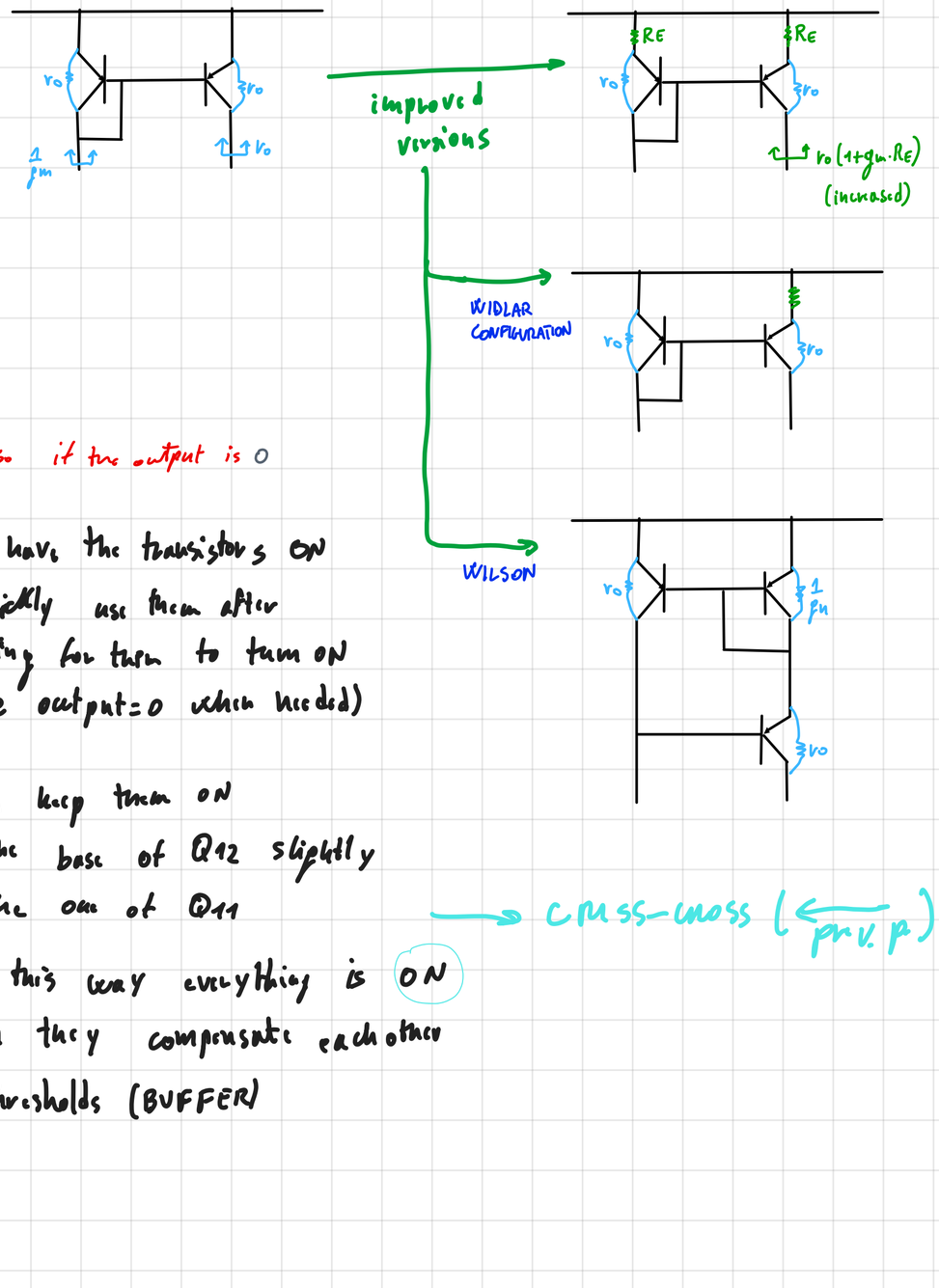
- Consider:



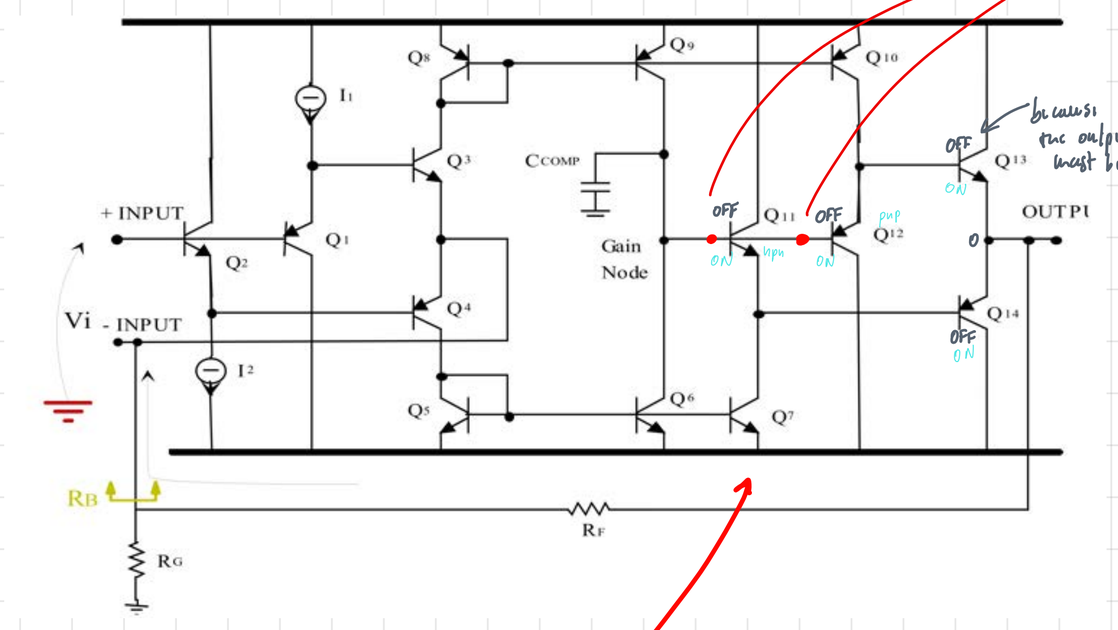
We want the impedance at the gain node to be very high

↳ We want ro to be high

↳ To improve the current mirror



- Consider



↳ It's better to have the transistors ON so we can quickly use them after without waiting for turn to turn ON (but still give output=0 when needed)

↳ We can keep them ON pulling the base of Q12 slightly below the one of Q11

↳ in this way everything is ON an they compensate each other thresholds (BUFFER)

→ cross-cross (← prev. p)

instead of using the same buffer of the input we can introduce an improvement (criss-cross)

Ex. - Datasheet

Data-sheet LH4117:

Some manufacturers put R_F directly inside, because we don't change it anyway (const. pole property of CFA for low gain)

↳ 'Cause if we want to change the gain we change R_G

Application-note:

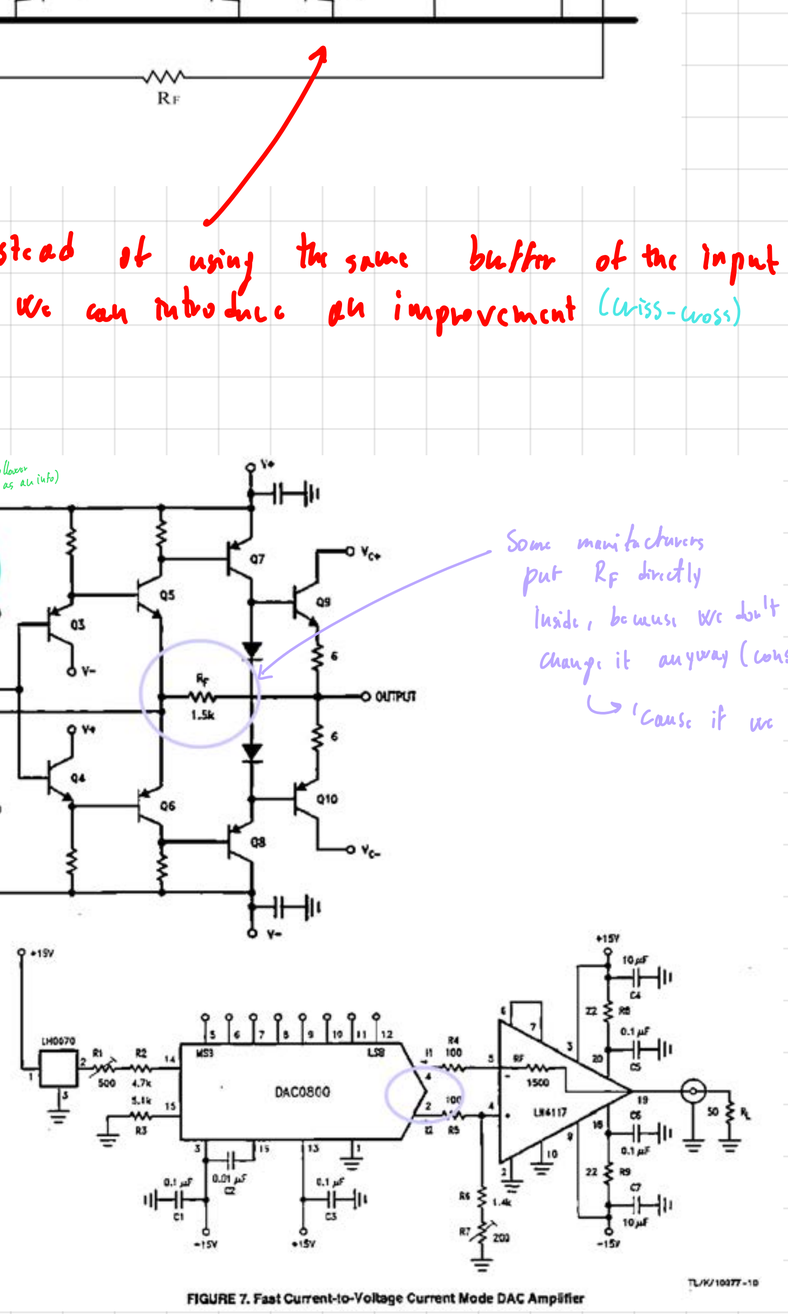


FIGURE 7. Fast Current-to-Voltage Current Mode DAC Amplifier

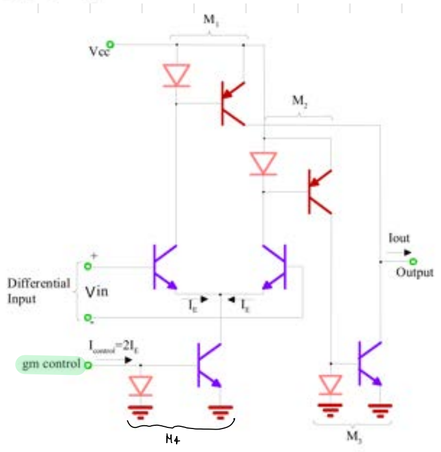
OTA

Intro

These amplifiers produce an output current proportional to the differential signal applied as the input signal. The gain is therefore a transconductance, defined as follows:

$$gm = \frac{dI_{OUT}}{dV_{diff}} = \frac{dI_{OUT}}{d(V_+ - V_-)}$$

Consider now the classical OTA simplified structure shown in Fig. 4.64. You can notice four current mirrors: it is interesting to observe that three of these mirrors translate a differential signal to a single-ended signal without any amplification of the signal (we suppose a mirroring ratio equal to one). The transconductance of the overall stage is equal to that of the input BJT, i.e. $g_m=I_E/V_{th}$.



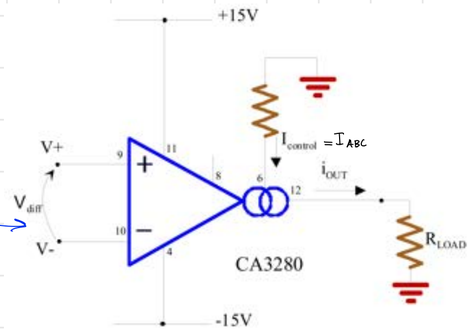
$$g_m = \frac{dI_{OUT}}{dV_{diff}} = \frac{dI_{OUT}}{d(V_+ - V_-)} = \frac{I_E}{kT/q}$$

both inputs be high impedance	(like VOA)
output be current	(like Norton)
transconductance gain	(I_{out}/V_{diff}) adjustable through a pin!

The transconductance expression shows that varying the current I_E, for example by means of an auxiliary input, it is possible to control the stage transconductance. This adjustment is linear on a wide range. Varying the control current, it is possible to modify the gain and other parameters related, such as the bandwidth, the power consumption, the input bias current, and others.

The capability of the transconductance to be controllable by the tail current is symbolized in the schematic shown in Fig. 4.65, but every producer uses its own symbolism. The current I_{ABC} in the schematic of the Harris Semiconductor CA3280 OpAmp plays the same role as the current I_E of the circuit shown in Fig. 4.64. The internal circuit of the CA3228 is reported in the attached data-sheet. Practically, choosing the external resistance, we can set the desired I_{ABC} for the desired performance, as shown in the curves depicted in Fig. 4.66. In this OpAmp, there is another control pin, I_D, which allows acting on the input-output characteristics linearization (Fig. 4.67).

We must observe that, during the open-loop operation, the input voltage V_{diff} could be smaller than the thermal voltage kT/q. Thus, the amplifier characteristic could be non-linear, in fact, a hyperbolic tangent. To improve the linearity, it is possible to use two techniques:emitting degeneration; diode linearization. In the first case, we add two resistors R_E in series with the emitters of the BJT differential input couple (Fig. 4.68) to increase the linearity range up to $I_{ABC}\cdot R_E$. If we use a current I_{ABC}=1mA, and a resistor R_E=1kΩ, we will obtain a linearity region up to $|V_{diff}|\leq$1V instead of $|V_{diff}|\leq V_{th}$=25mV. In this way, however, the stage gain is penalized because it is reduced. The second solution is better and consists of placing two diodes as an input, and the voltage in these diodes "counterbalances" the exponential characteristic of the base-emitter junction of the input BJT couple (Fig. 4.69).



$$i_{out} = G_m(I_{control})\cdot v_{diff} = \frac{I_{control}}{kT/q}\cdot v_{diff} = \frac{I_{control}}{25mV}\cdot v_{diff}$$

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

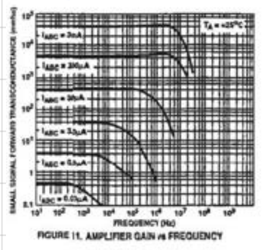


FIGURE 11. AMPLIFIER GAIN vs FREQUENCY

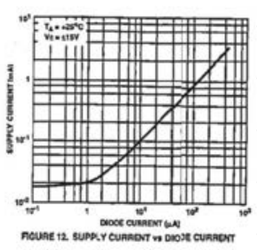
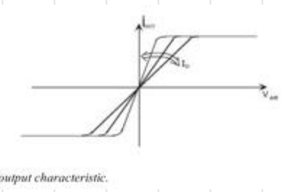
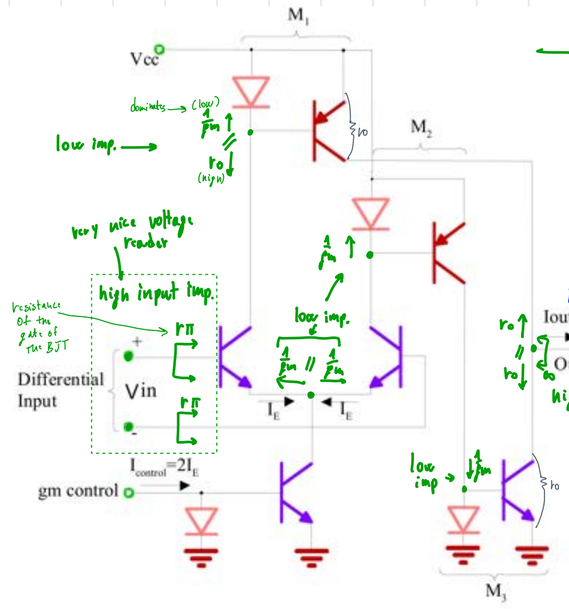


FIGURE 12. SUPPLY CURRENT vs DIODE CURRENT


Input-output characteristic.

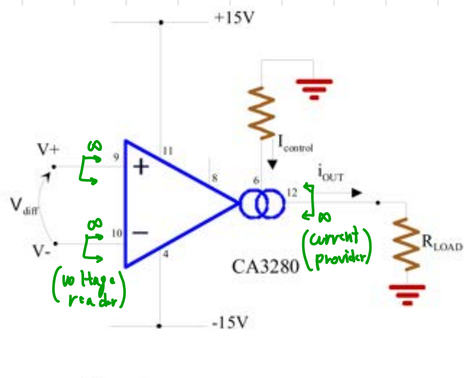
Working Principle



→ all internal nodes of the OTA are low impedance

the output is a CURRENT (no current-voltage conversion)

$$g_m = \frac{dI_{OUT}}{dV_{diff}} = \frac{dI_{OUT}}{d(V_+ - V_-)} = \frac{I_E}{kT/q}$$



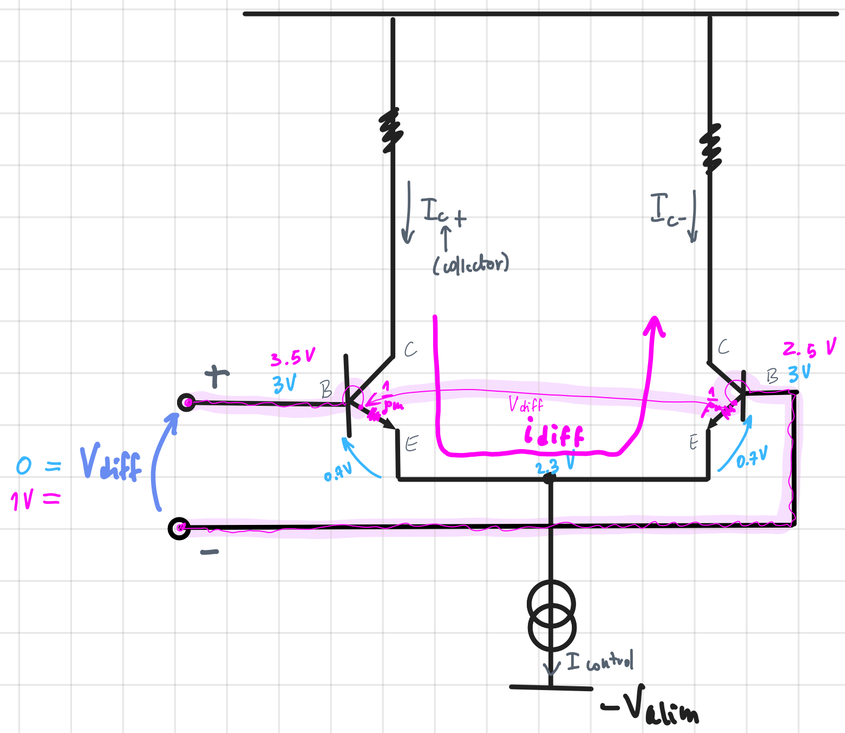
$$i_{out} = G_m(I_{control})\cdot v_{diff} = \frac{I_{control}}{kT/q}\cdot v_{diff} = \frac{I_{control}}{25mV}\cdot v_{diff}$$

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

both inputs be high impedance	(like VOA)
output be current	(like Norton)
transconductance gain	(I_{out}/V_{diff}) adjustable through a pin!

- Let's consider the simpler case (just resistor loads)



- Case 1: $I_{control}$ splits in 2 $\longrightarrow I_c = \frac{I_{control}}{2}$ ($V_{diff}=0$)
- case 2: V_{diff} is applied bw the transistor also in this way ($V_{diff}\neq 0$)
  - $\frac{1}{g_m} = \frac{KT/q}{I_c} = \frac{25mV\ @RT}{I_{control}/2}$
  - $i_{diff} = \frac{V_{diff}}{\frac{1}{g_m}+\frac{1}{g_m}} = \frac{V_{diff}}{2\cdot\frac{25mV}{I_{control}/2}}$

- Now consider again connected the current mirrors instead of the resistors

$$\Rightarrow i_{OUT} = 2\,i_{diff} = 2\,\frac{V_{diff}}{2\cdot\frac{1}{g_m}} = \underbrace{G_m}_{g_m} V_{diff}$$

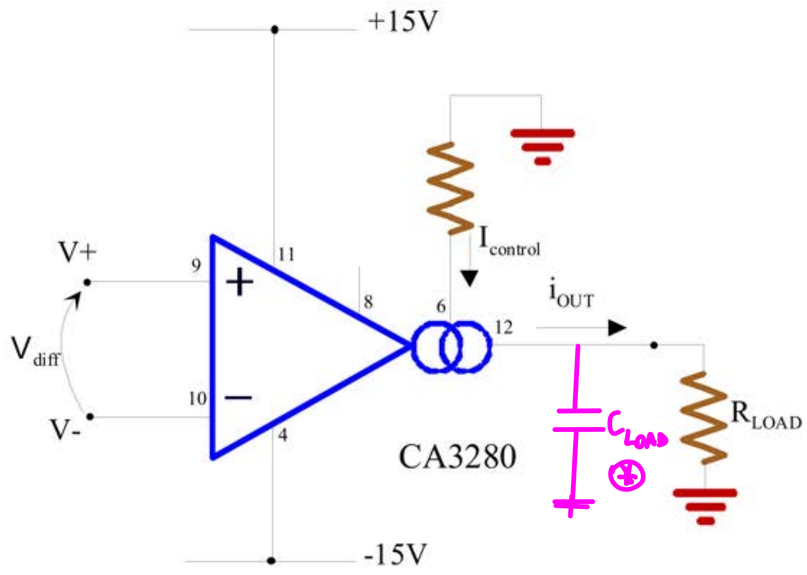
$$G_m = g_m = \frac{I_{control}/2}{25mV}$$

depends on the architecture we have

e.g. for different mirroring factors

$$g_m = \frac{dI_{OUT}}{dV_{diff}} = \frac{dI_{OUT}}{d(V_+ - V_-)} = \frac{I_E}{kT/q} = \frac{2I_E}{I_{control}/2 \,/\, kT/q}$$

→ It can be then represented with this scheme:

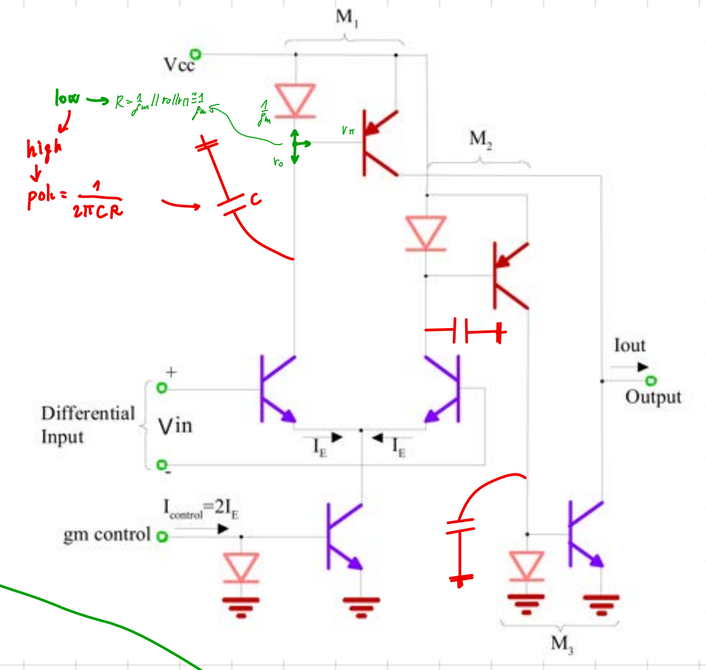


$$i_{out} = G_m(I_{control}) \cdot v_{diff} = \frac{I_{control}}{kT/q} \cdot v_{diff} = \frac{I_{control}}{25mV} \cdot v_{diff}$$

Obs. Low impedance on the internal nodes means that: if we have some stray capacitances, they will introduce high freq. poles

↓

HIGH BANDWIDTH



Advantages: (seen before)

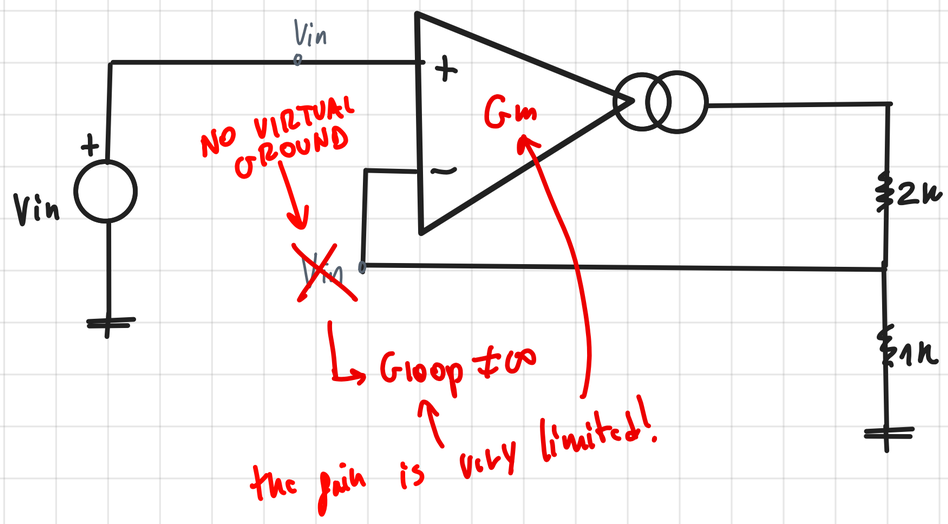
- all low-impedance nodes
- wide bandwidth
- $f_{pole} = 1/2\pi C_{LOAD} R_{LOAD}$ ← set just by the output capacitor

Disadvantages:

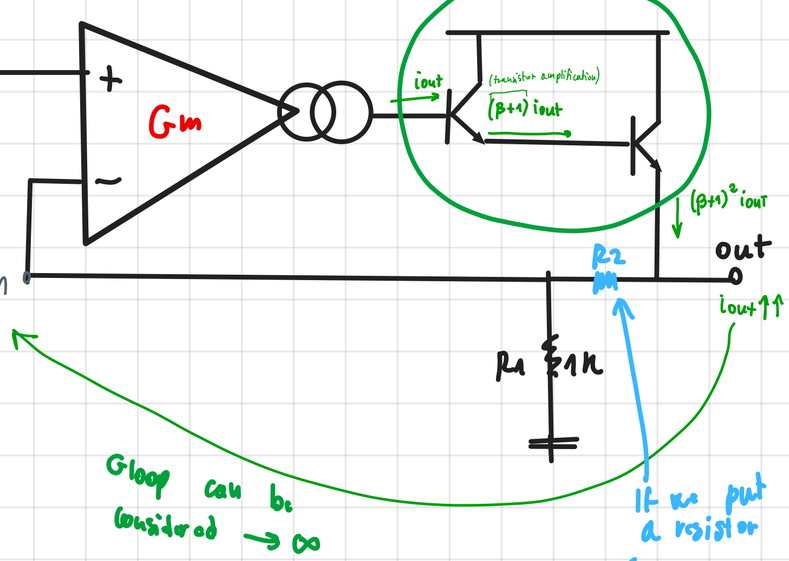
- no infinite gain
- no "VIRTUAL GROUND"
- it is used open-loop or together with amplifying components
- + STRONG NONLINEARITY

It only remains the output pole ← we have that the internal parasitic capacitors are at very high freq. (↳ negligible) ← so if we connect a C_{LOAD} (+)

↳ **No infinite gain** Indeed if we consider:



We can try to improve it in this way

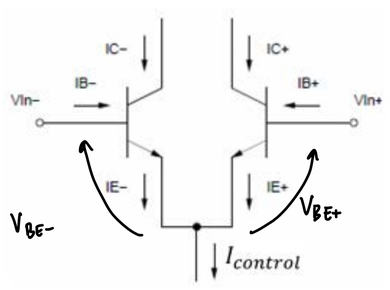


Darlington couple to amplify the gain

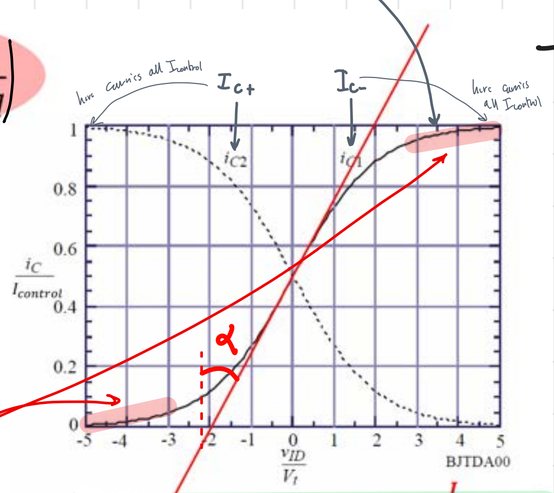
↳ **Nonlinearity**

$$I_{out} = I_{C+} - I_{C-} = I_{control} \cdot \tanh\left(\frac{V_{diff}}{2 \cdot kT/q}\right)$$

$\left(\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}\right)$ hyperbolic tangent



$\frac{I_C}{I_{control}}$



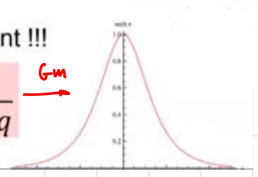
Linearization $g_m = \frac{I_{control}}{2?\,kT/q} = tg\,\alpha$ (depending on architecture)

→ Using BJT transistor equation:

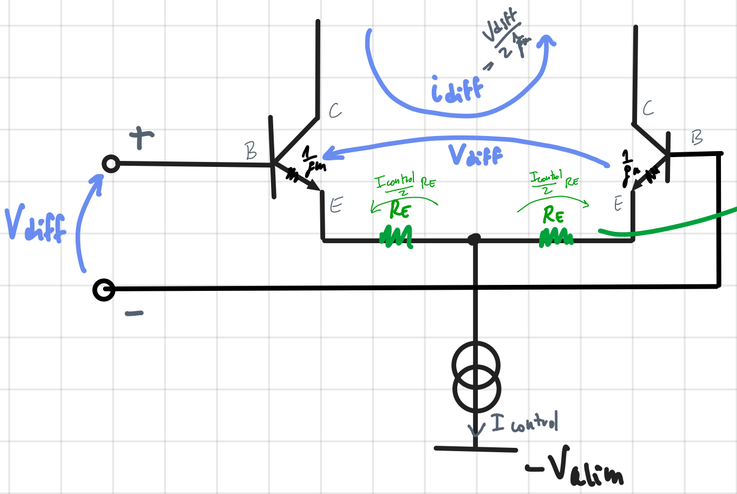
- $I_c = I_s\, e^{\frac{V_{BE}}{kT/q}}$ (collector)
- $V_{diff} = V_{BE-} - V_{BE+}$
- $I_{control} = I_{C+} - I_{C-}$

The transconductance is anything but constant !!!

$$g_m = \frac{I_{control}}{2 \cdot kT/q} \cdot sech^2 \frac{V_{diff}}{2 \cdot kT/q}$$ → Gm



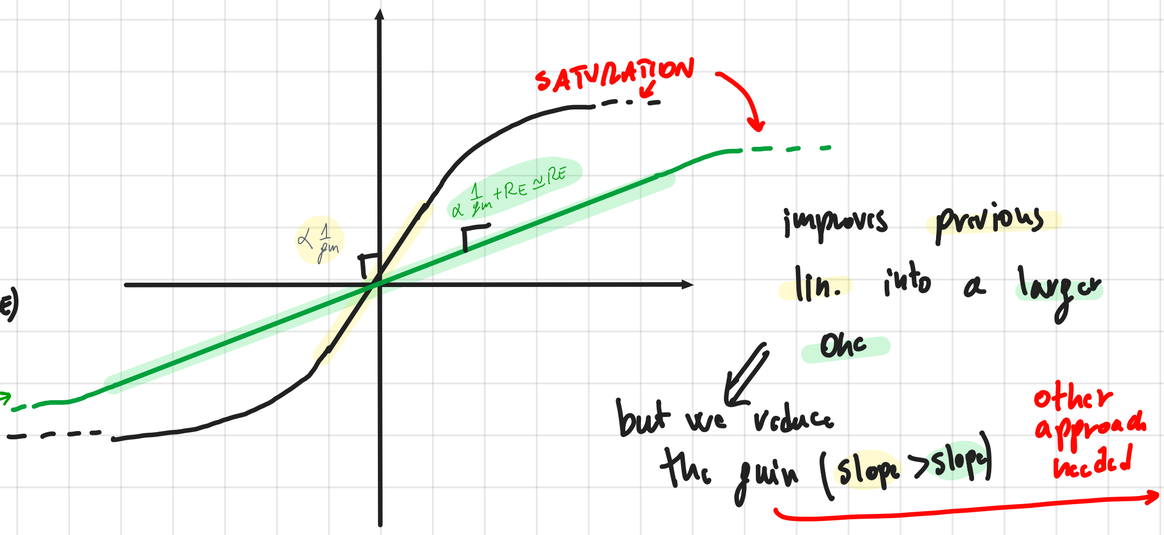
↳ to improve it and increase the linearity region we can



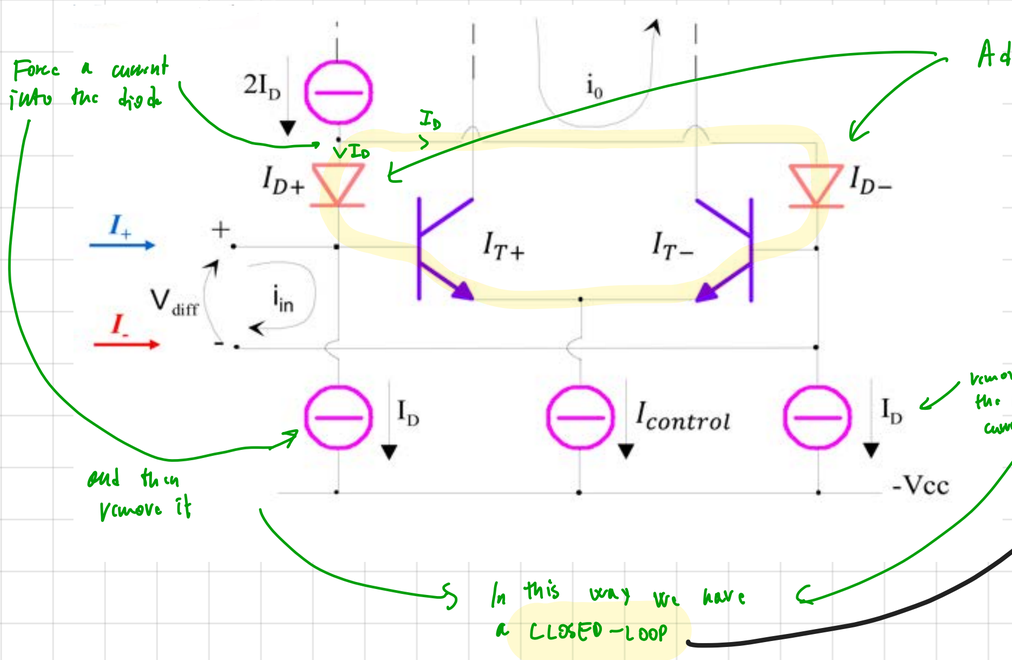
We can add these generation resistors

↳ i_{diff} (with RE) $= \frac{V_{diff}}{\frac{1}{g_m} + R_E + R_E + \frac{1}{g_m}} = \frac{V_{diff}}{2\left(\frac{1}{g_m} + R_E\right)}$

drastically reduces



→ A better approach to improve the linearity is to consider:



↳ Every time we have a loop with diodes/transistors we can apply the translinear principle

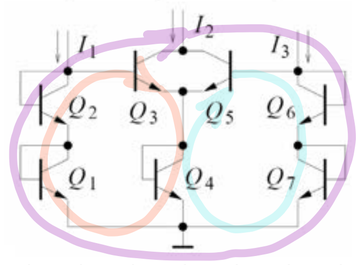
Given a loop of closed-connected branches, Kirchoff's law says: $0 = \sum_{j\in\{CW\}} V_{ej} - \sum_{l\in\{CCW\}} V_{el}$

$$V_e = K \ln\left(\frac{I}{I_d}\right)$$

$$\prod_{j\in\{CW\}} I_j = \prod_{l\in\{CCW\}} I_l$$

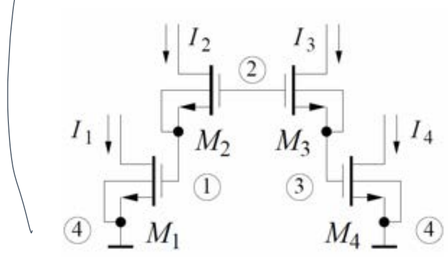
(we can also create a similar principle but there are different thresholds)

For loops with just BJTs, K=kT/q

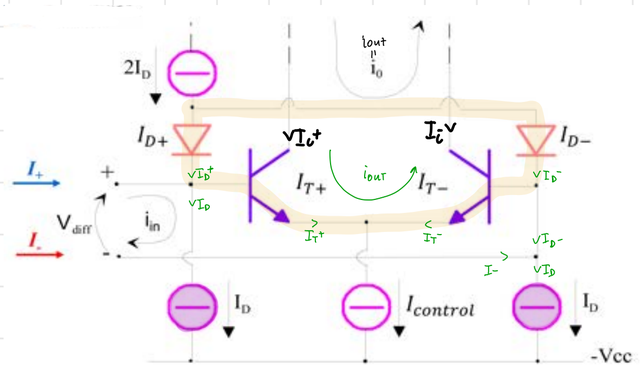


loop: $Q_1 - Q_2 - Q_6 - Q_7$
loop: $Q_4 - Q_5 - Q_6 - Q_7$
loop: $Q_1 - Q_2 - Q_3 - Q_4$

For loops with just MOSFETs, $K = nV_T$



↳ So coming back to the solution:



$\rightarrow I_+ = I_D - I_{D+}$ and $\rightarrow I_- = I_D - I_{D-}$

They can differ, so an input differential current exists!

$$I_{in} = I_+ - I_- = I_{D-} - I_{D+}$$

But the out/in relationship is linear!!!

$$I_{out} = I_{C+} - I_{C-} = \frac{I_{control}}{I_d} \cdot I_{in} \quad \text{when } |I_{in}| < I_d$$

(gain G)

Translinear principle: $I_{D+} \cdot I_{T+} = I_{D-} \cdot I_{T-}$ (ccw = cw)

Node currents: $\rightarrow I_{D+} = I_D - I_{in}$ $\rightarrow I_{D-} = I_D + I_{in}$

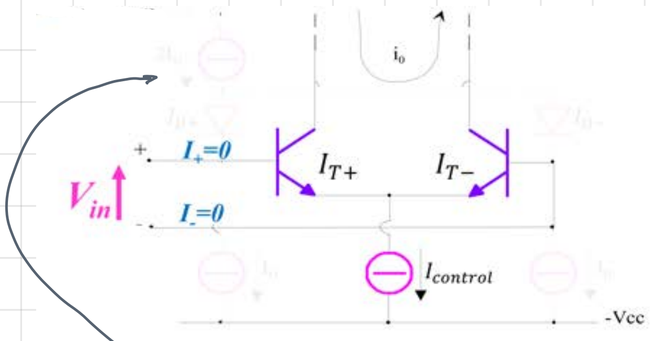
Output signal current: $\rightarrow I_{out} = I_{T+} - I_{control} = I_{control} - I_{T-} = I_{T+} - I_{T-}$

($I_{out} = I_D$; $I_{T+} = I_{C+}$; $I_{T-} = I_{C-}$)

Case 1 (No I_D): $I_D^+ = I_D$ (removed); $I_D^- = I_D$ (removed); $I^+ = 0$; $I^- = 0$ $\rightarrow I_{IN} = 0$

Case 2 (I_D): $I_D^+ \neq I_D^- \neq I_D \rightarrow \dots \rightarrow i_{out}$ is created

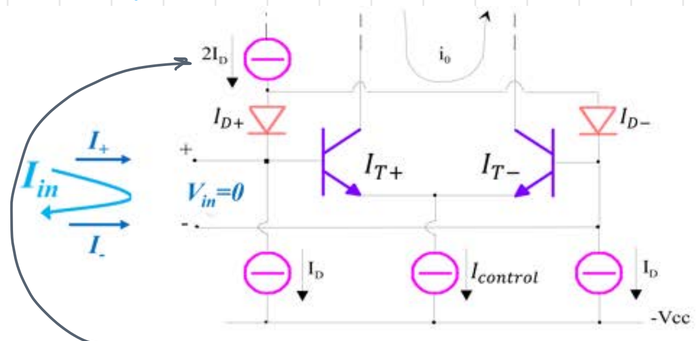
Operating modes



When I_D is not used, the linearization is **off** and:

$$I_{out} = \frac{I_{control}}{kT/q} \cdot V_{in} \quad when\ |V_{in}| < kT/q$$

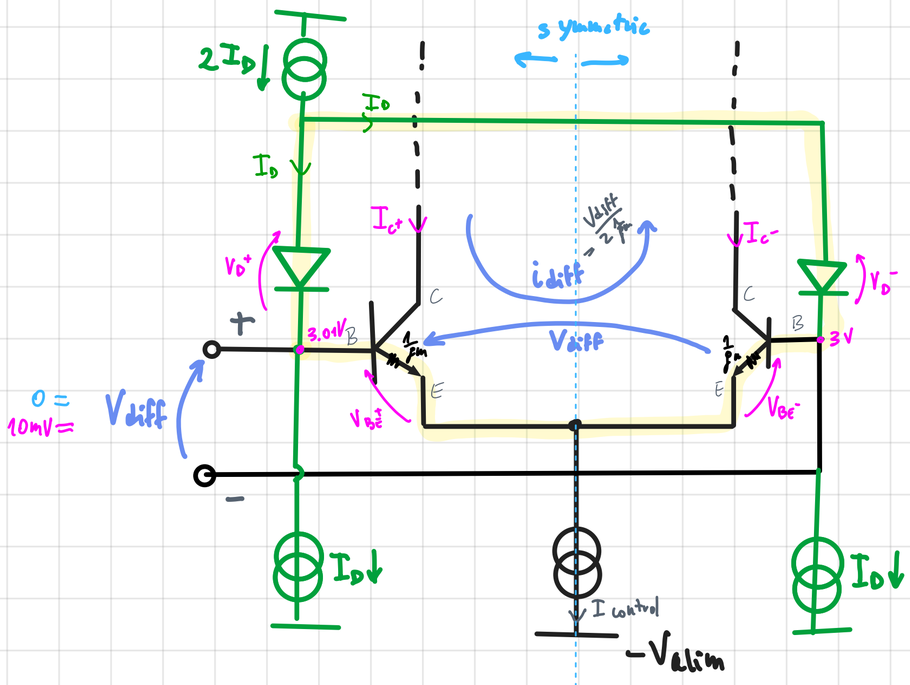
Transconductance amplifier



When I_D is set, the linearyzation is **on** and:

$$I_{out} = \frac{I_{control}}{I_d} \cdot I_{in} \quad when\ |I_{in}| < I_d$$

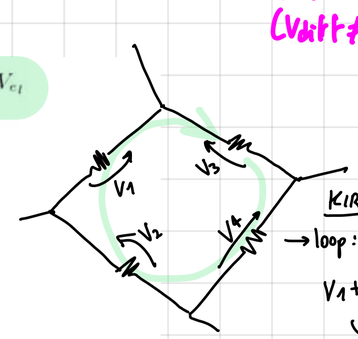
Current amplifier



• Case 1: The stage is fully symmetric ($V_{diff} = 0$)

• Case 2: ($V_{diff} \neq 0$) $V_{D+} + V_{BE+} = V_{D-} + V_{BE-}$

causes I_{C+} causes I_{C-}



KIRCHOFF: → loop: $\sum V_i = 0$

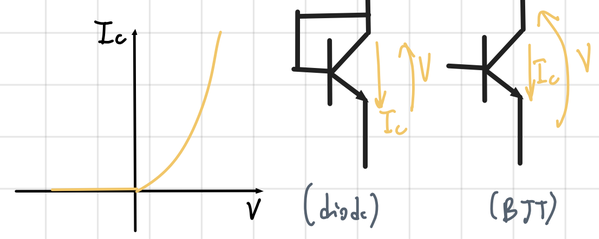
$V_1 + V_2 - V_3 - V_4 = 0$ ← just for resistors

$\hookrightarrow V_1 + V_2 = V_3 + V_4$ (cw = ccw)

↳ now we consider a similar principle for transistors

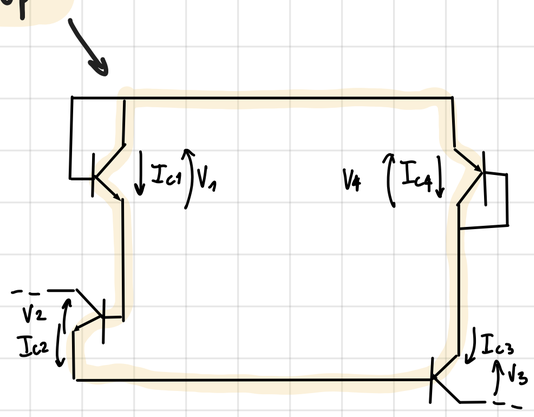
$V = \frac{nT}{K} \ln\left(\frac{I_C}{I_S}\right)$ ← $I_C \approx I_{collector} = I_S e^{\frac{V}{nT}}$

$\hookrightarrow V = K \ln\left(\frac{I_C}{I_S}\right)$



applying KIRCHOFF for multiple transistors e.g.

loop



KV: $V_1 + V_2 = V_3 + V_4$

$\hookrightarrow K \cdot \ln\left(\frac{I_{C1}}{I_{S1}}\right) + K \ln\left(\frac{I_{C2}}{I_{S2}}\right) = K \ln\left(\frac{I_{C3}}{I_{S3}}\right) + K\left(\frac{I_{C4}}{I_{S4}}\right)$

↳ If the transistors are all equal $I_{S1} = I_{S2} = I_{S3} = I_{S4}$

$\hookrightarrow \ln(I_{C1}) + \ln(I_{C2}) = \ln(I_{C3}) + \ln(I_{C4})$

$\hookrightarrow \ln(I_{C1} \cdot I_{C2}) = \ln(I_{C3} \cdot I_{C4})$

$\Rightarrow \underbrace{I_{C1} \cdot I_{C2}}_{ccw} = \underbrace{I_{C3} \cdot I_{C4}}_{cw}$

Translinear principle: $\underbrace{(I_D - I_{in})}_{I_D^+} I_{C+} = (I_D + I_{in}) I_{C-} \rightarrow I_{C+} I_D - I_{C+} I_{in} - I_{C-} I_D - I_{C-} I_{in} = 0$

$I_C = \frac{I_{control}}{2}$ → $I_D (I_{C+} - I_{C-}) = (I_{C+} + I_{C-}) I_{in}$

I_{out} $\frac{I_{control}}{2}$

$\hookrightarrow I_{out} = \frac{I_{control}}{I_D} I_{in}$

linear

$$I_{out} = \frac{I_{control}}{I_D} \cdot I_{in} = \frac{650\mu A}{200\mu A} \cdot \frac{V_{diff}}{20k\Omega} = \frac{V_{diff}}{6153\Omega}$$

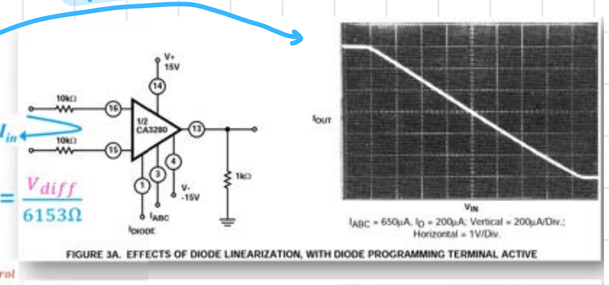
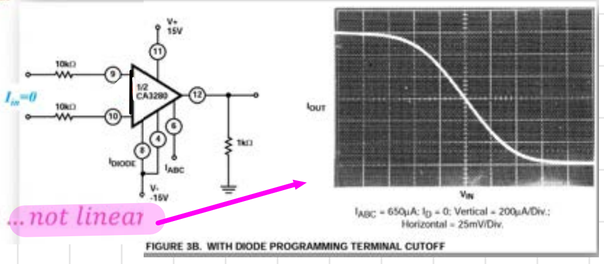


FIGURE 3A. EFFECTS OF DIODE LINEARIZATION, WITH DIODE PROGRAMMING TERMINAL ACTIVE

$I_{control}$

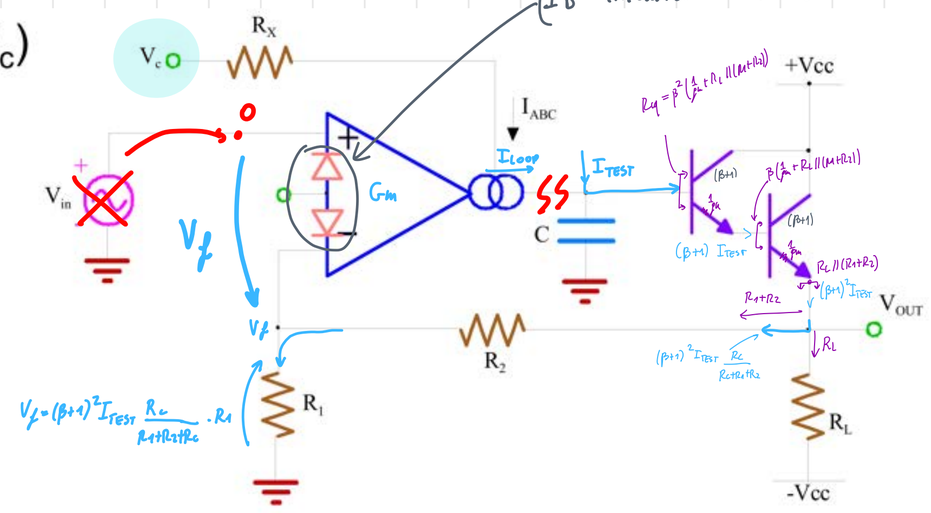


$I_{out} \approx \frac{I_{control}}{2\,kT/q} \cdot V_{diff} \approx \frac{V_{diff}}{77\Omega}$ but ... not linear

FIGURE 3B. WITH DIODE PROGRAMMING TERMINAL CUTOFF

OTA: example of application

i.e. $f_{pole}(V_c)$



↳ Gloop computation

1. cut the loop ⚡ and turn off ~~Vin~~
2. Identify I_{loop} and apply I_{test} $\longrightarrow$ $G_{loop} = \frac{I_{loop}}{I_{test}}$

↳ $I_{loop} = G_m \cdot V_f \longrightarrow G_{loop} = (\beta+1)^2 \frac{R_c}{R_1+R_2+R_c} \cdot R_1 \cdot G_m$

↳ $G_{loop} = \frac{I_{control}}{\frac{kT}{q}} (\beta+1)^2 \frac{R_c \cdot R_1}{R_1+R_2+R_c} \cong 10^5$ ← G_{loop} is high (ideal)

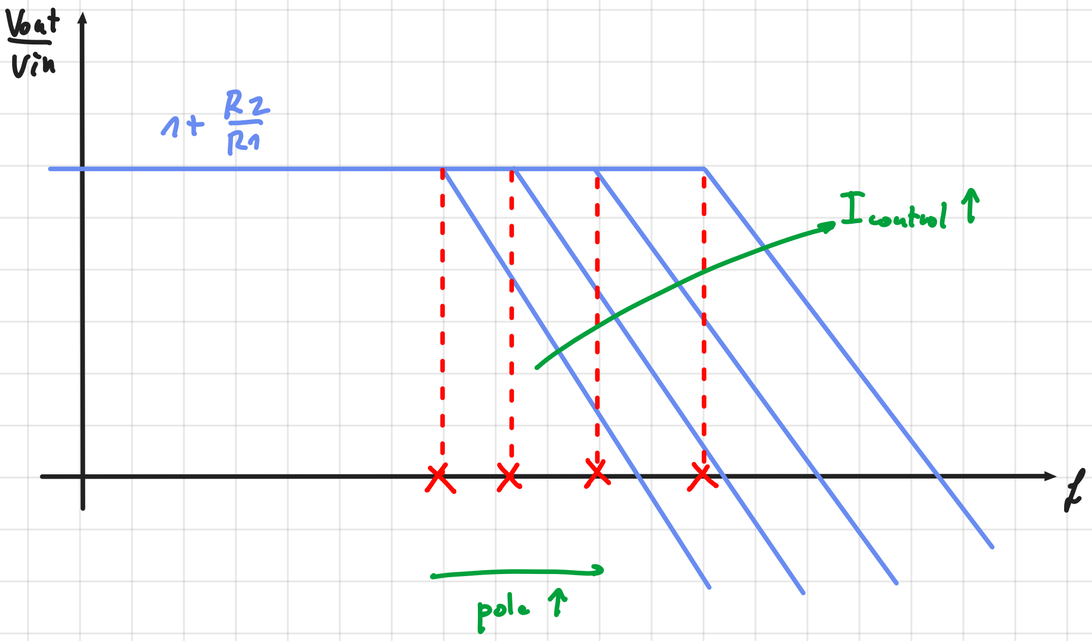
e.g. $\frac{10mA}{25mV}$; $(\beta+1)^2 \approx 10^4$; $\frac{1M \cdot 1M}{1M+1M+100k}$

$\Downarrow$

$G_{real} \cong G$

Gain $G = \frac{V_{out}}{V_{in}} = 1 + \frac{R_2}{R_1}$

Bode $\frac{V_{out}}{V_{in}}$



A secondary effect of the differential stage is that when we increase $I_{control}$ the pole increases, this is due to the parasitic capacitances inside the OTA amplifier.

The capacitances change with $\frac{1}{g_m}$, $\frac{1}{g_m}$ change with the current

↳ $I_{control} \uparrow \Rightarrow g_m \uparrow \Rightarrow$ parasitism ↓ (pole ↑ improves)

$I_{control} = I_{ABC}$

Data-sheet CA3280:

intersil — CA3280, CA3280A

Data Sheet — May 2002 — FN1174.6

Dual, 9MHz, Operational Transconductance Amplifier (OTA)

The CA3280 and CA3280A types consist of two variable operational amplifiers that are designed to substantially reduce the initial input offset voltage and the offset voltage variation with respect to changes in programming current. This design results in reduced "AGC thump," an objectionable characteristic of many AGC systems. Interdigitation, or crosscoupling, of critical portions of the circuit reduces the amplifier dependence upon thermal and processing variables.

The CA3280 has all the generic characteristics of an operational voltage amplifier except that the forward transfer characteristics is best described by transconductance rather than voltage gain, and the output is current, not voltage. The magnitude of the output current is equal to the product of transconductance and the input voltage. This type of operational transconductance amplifier was first introduced in 1969, and it has since gained wide acceptance as a gateable, gain controlled building block for instrumentation and audio applications, such as linearization of transducer outputs, standardization of widely changing signals for data processing, multiplexing, instrumentation amplifiers operating from the nanopower range to high current and high speed comparators.

For additional application information on this device and on OTAs in general, please refer to Application Notes: AN6818, AN6668, and AN6077.

Features

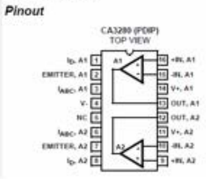
- Low Initial Input Offset Voltage: 500µV (Max) (CA3280A)
- Low Offset Voltage Change vs I_{ABC}: <500µV (Typ) for All Types
- Low Offset Voltage Drift: 5µV/°C (Max) (CA3280A)
- Excellent Matching of the Two Amplifiers for All Characteristics
- Internal Current-Driven Linearizing Diodes Reduce the External Input Current to an Offset Component
- Flexible Supply Voltage Range ..... ±2V to ±15V

Applications

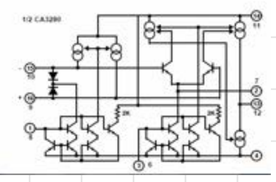
- Voltage Controlled Amplifiers
- Voltage Controlled Oscillators
- Multipliers
- Demodulators
- Sample and Hold
- Instrumentation Amplifiers
- Function Generators
- Triangle Wave-to-Sine Wave Converters
- Comparators
- Audio Preamplifier

Pinout

CA3280 (PDIP) TOP VIEW



Functional Diagram



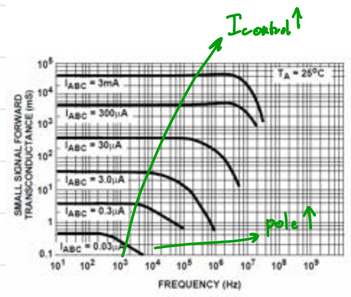


FIGURE 11. AMPLIFIER GAIN vs FREQUENCY

• $pole_{open\,loop} = \frac{1}{2\pi C \cdot \underbrace{\beta^2 \left(\frac{1}{g_m} + R_c // (R_1+R_2)\right)}_{R_{eq} \cong \beta^2 R_c}} = \frac{1}{2\pi C \beta^2 R_c}$

↳ $pole_{closed\,loop} = pole_{open\,loop} \cdot (1 - G_{loop}) \cong \frac{\frac{I_{control}}{25mV} \cdot \beta^2 \frac{R_c \cdot R_1}{R_1+R_2+R_c}}{2\pi C \beta^2 R_c} \propto I_{control}$

the pole varies with $I_{control}$

Note It's called voltage-controlled LP filter because:

if we change V_c: $V_c \Rightarrow I_{control} \Rightarrow pole \Rightarrow$ LP filter

Other examples

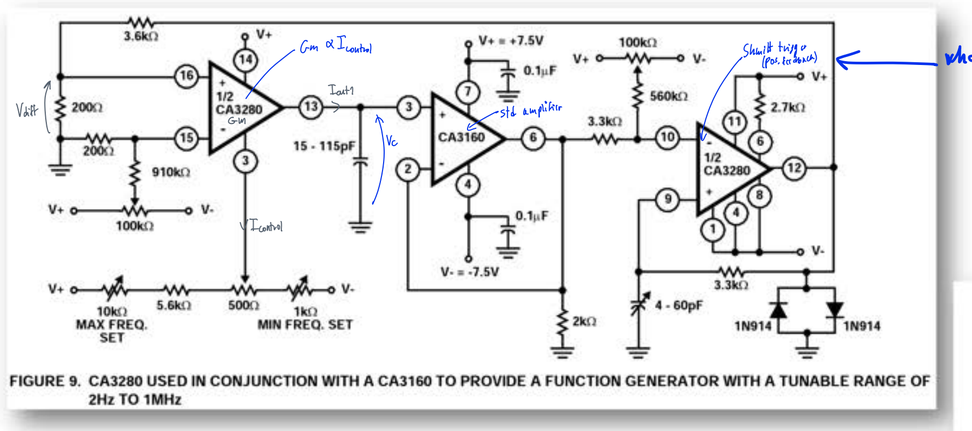
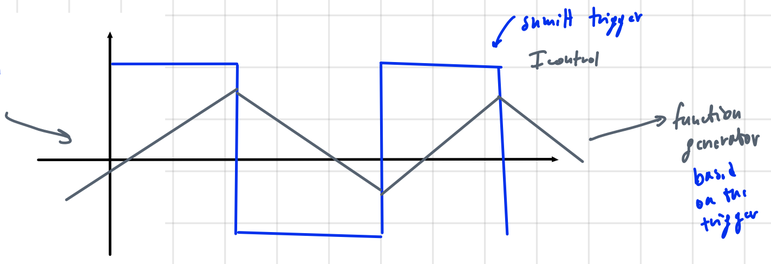


FIGURE 9. CA3280 USED IN CONJUNCTION WITH A CA3160 TO PROVIDE A FUNCTION GENERATOR WITH A TUNABLE RANGE OF 2Hz TO 1MHz

when it commutes $I_{control}$ change sign ↳ start decreasing



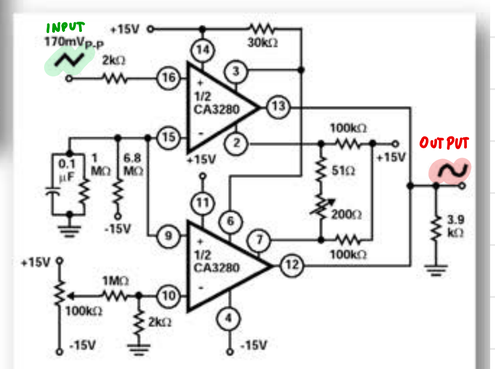
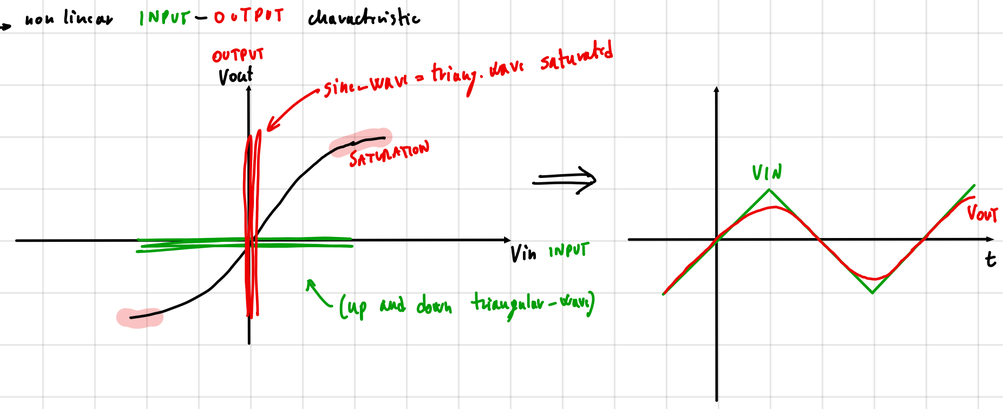


FIGURE 10. TRIANGLE WAVE-TO-SINE WAVE CONVERTER

non linear INPUT – OUTPUT characteristic



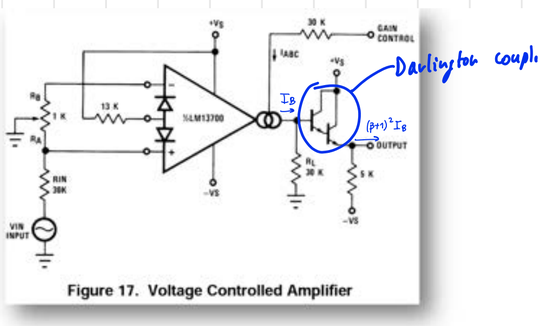


Figure 17. Voltage Controlled Amplifier

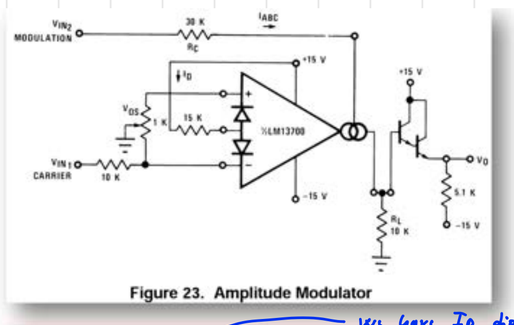


Figure 23. Amplitude Modulator

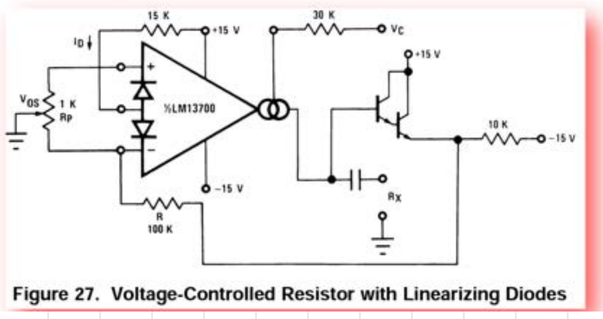


Figure 27. Voltage-Controlled Resistor with Linearizing Diodes

we have ID diodes → we can change the gain

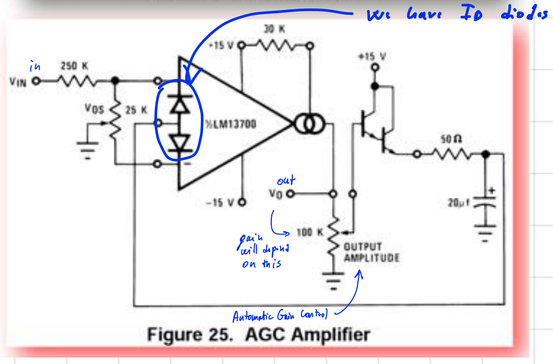


Figure 25. AGC Amplifier

ISO

(Book p. 287)

Many acquisition data systems have multiplexers to select different sources. These front-end circuits are damaged by voltages over 20-30V. To protect them from common-mode voltages and to ensure a galvanic isolation, it is necessary to use the Isolation-Amplifiers (Fig. 4.17). These OpAmps do not protect the input terminals from excessive differential voltages applied to them, but eliminate the high current which would flow from the input towards the output (and then towards the acquisition system made of MUX, ADC, etc.) if at the input, a high common-mode voltage were applied. The Isolation Mode Rejection Ratio (IMRR) is very important, defined as shown in Fig. 4.18.

With increasing amplitude and frequency of V_{iso}, the isolation techniques become more important because it gets more difficult to effectively isolate the two worlds. Typical values are between 140dB and 160dB for DC while the IMRR decreases with a slope of 20dB/dec with increasing frequency (Fig. 4.19). Also PCB can deteriorate performance.

The IMRR does not suffice to satisfactorily indicate the capability of the ISO to isolate the two worlds. Particularly, the *steep edges* can overload the amplifier, causing non-linearity, offset errors, and saturation effects. To have a response on the behavior of the ISO from this point of view, we can use the TRANSIENT IMMUNITY (TI) expressed in [V/µs]. For example, for the ISO122, we have:

TI (dV/dt of the V_{ISO} without errors)<1000V/µs

If, for a given application, it is important to have a low distortion, the V_{iso} frequency must be lower than half of the modulation frequency of the Op-Amp. As we can see, the coupling can be made with a modulated transformer or with capacitors:

fV_{ISO} (for low distortion) < $f_{mod}/2$

Information (not just digital, but also analog) can be coupled through three different ways: optical, magnetic, or capacitive. Every method has its own pros and cons, thus being essential to properly trade-off for the correct choice.

Fig. 4.17

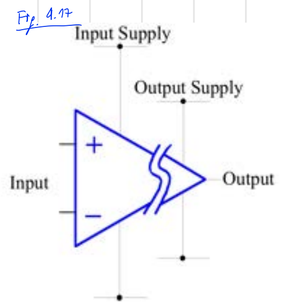


Fig. 4.18

Figure of Merit: **Isolation Mode Rejection Ratio**

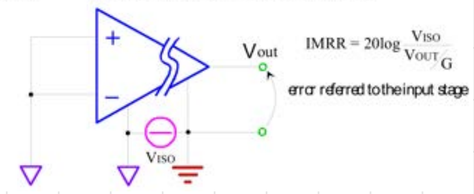
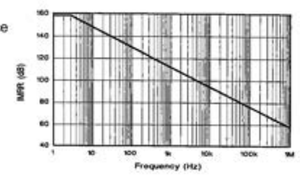


Fig. 4.19

IMRR frequency dependance



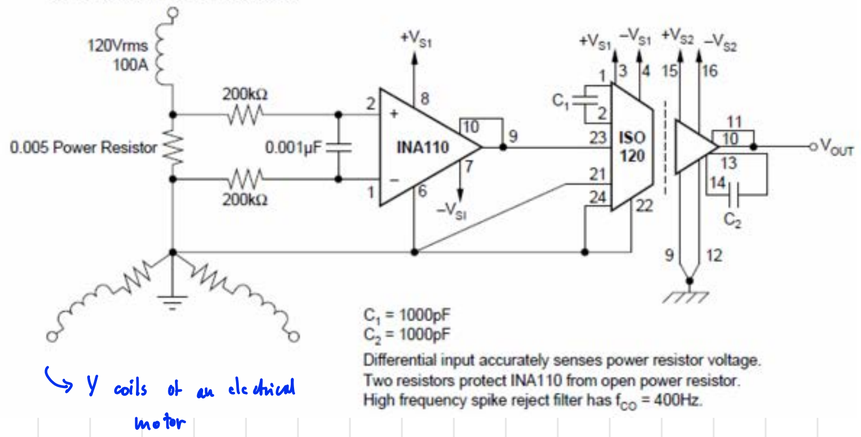
Transient Immunity: ISO122 TI < 1000V/ms

Requirements:

- isolated input and output stages
- double power supply
- galvanic isolation

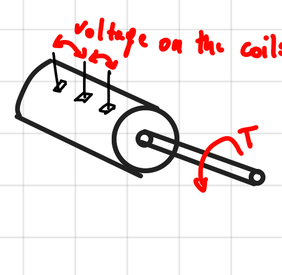
Ex. Motor powerline isolation

Isolated powerline monitor

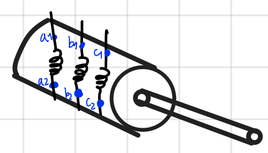


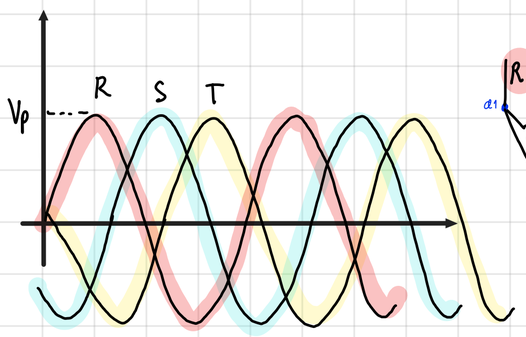
↳ Y coils of an electrical motor

motor



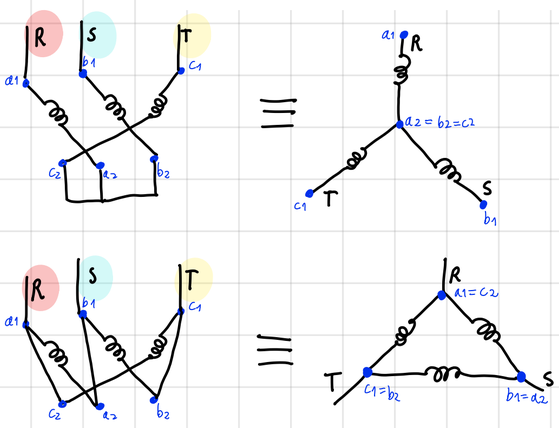
triphase coil connection





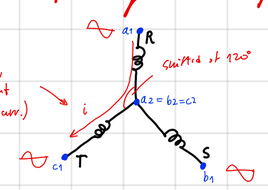
↳ at 380 V_{rms}

↳ $V_p = V_{rms} \cdot \sqrt{2} \cong 512$ V



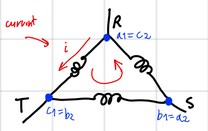
Y connection → The phases goes slowly

turns slowly
high torque ← Slow current (low curr.)

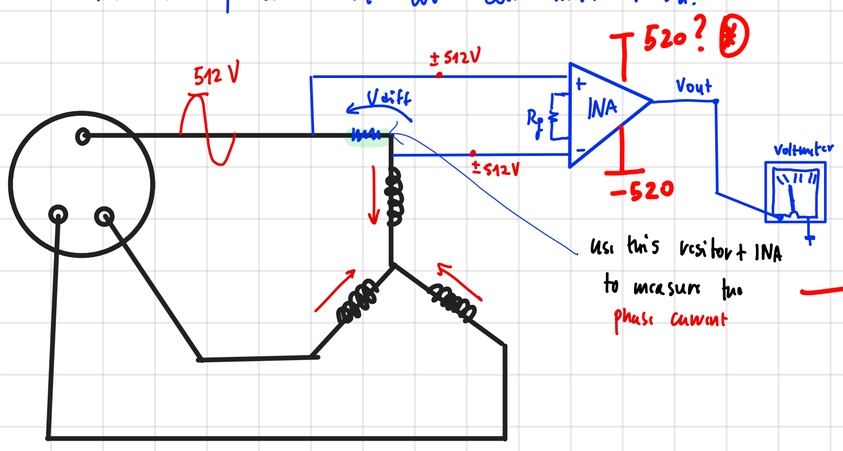


Δ connection → The phases goes faster

higher speed
lower torque ← higher current



↳ To monitor the phase current we could think to use:

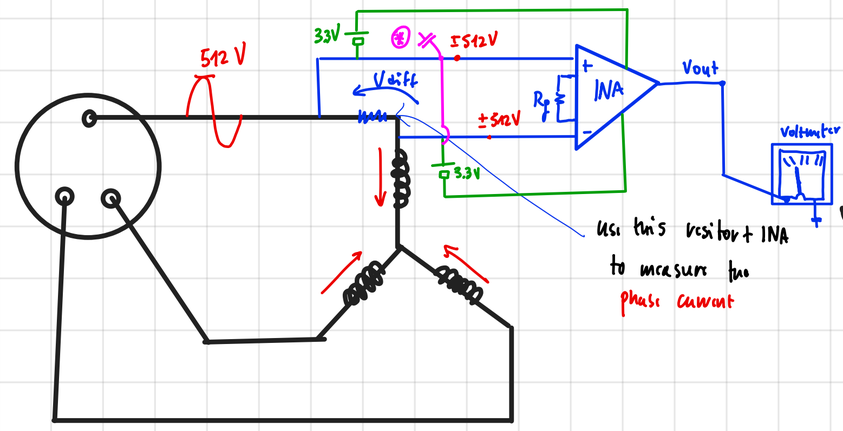


→ but the voltage phase goes up and down with Vp = 512 V

↳ we should have a power supply of (±520V) ← NOT POSSIBLE

we can only apply to an INA P.S. of e.g. ±5V

↳ A possible solution could be to bias the INA OpAmp with a battery of e.g. 3.3V



→ But now where is the ground?

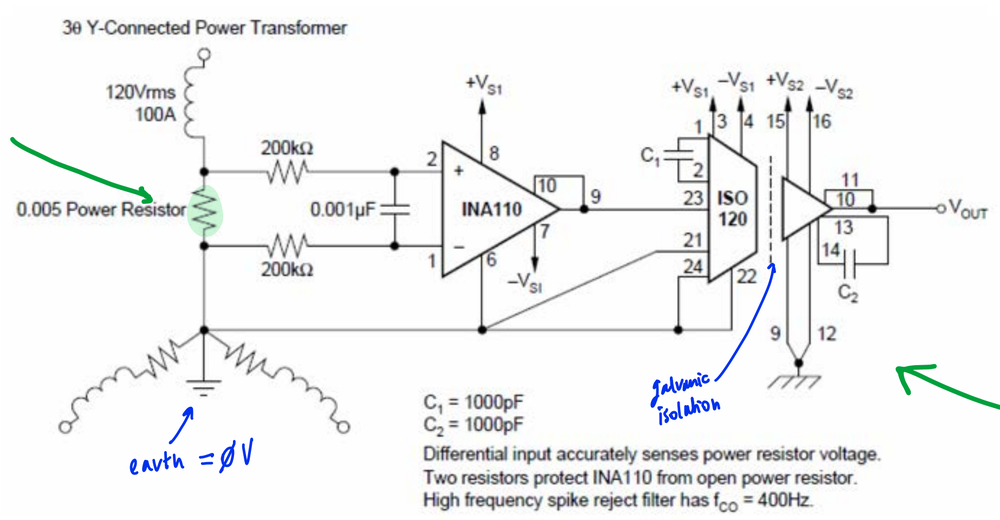
↳ we can connect it like this ⊛ and call it a ground (even though it's not ⏚)

↳ But now this ground will be at ↑↓ 512V → very dangerous

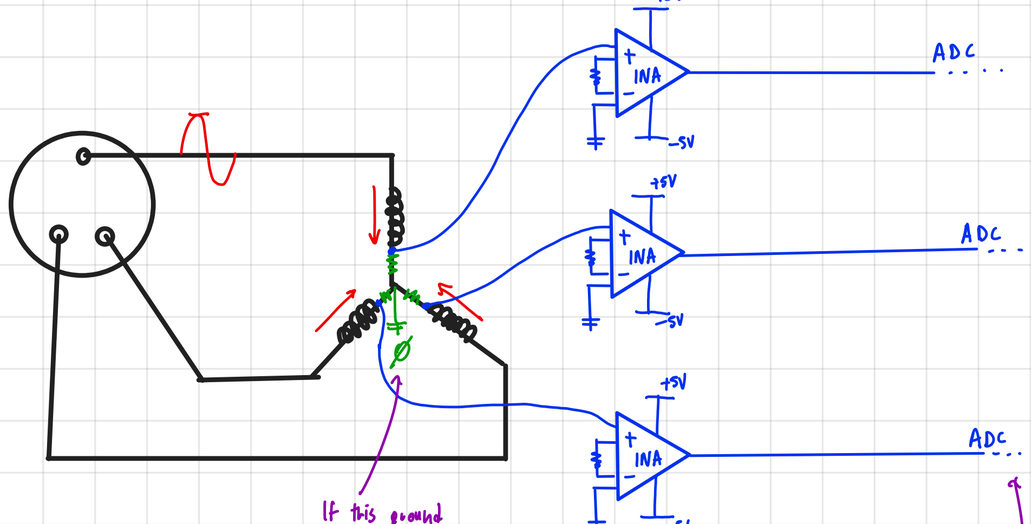
↳ Also for the different phases will have different P.S. and grounds of the Voltmeter

↳ The solution is:

Isolated powerline monitor



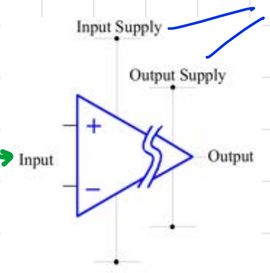
put the resistor here



We need ISOLATION

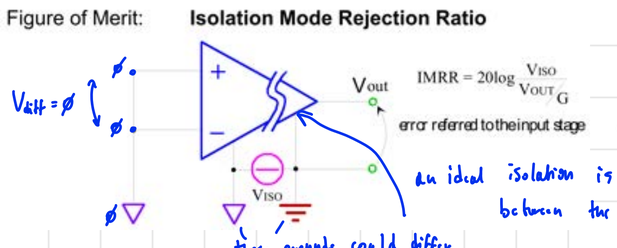
↓

← GALVANIC ISOLATION → Input / Output



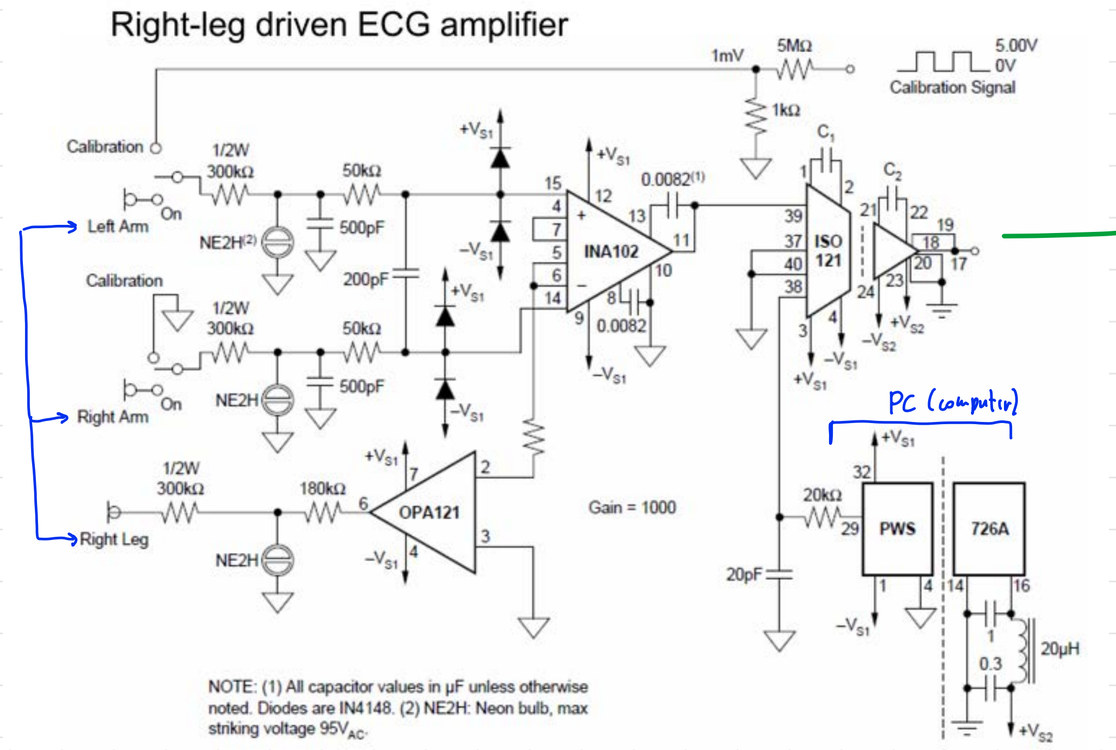
⊛ We can use different biases for the input and output stages because they will be ISOLATED

Figure of Merit: **Isolation Mode Rejection Ratio**



an ideal isolation is when, even if there's a difference between the two grounds (or biases ⊛), Vout=0

Ex. Low-voltage application — ECG (not only high voltage application like the motor phases)



Patient heartbeat (low voltage)

→ to ISOLATE patient and PC

Optical Coupling

This is the simpler and more intuitive method. With an optocoupler, it is immediate to transport digital information (on/off). To transfer analog information, it is possible to convert that information into a digital code (with an ADC within the ISO) and use "n" optocoupler and convert to the analog domain (DAC within the ISO), as done in the Capacitive Coupling. Another way consists of optically transferring analog information using a feedback configuration and trusting on the electrical and optical matching of the two detectors, as shown in Fig. 4.20.

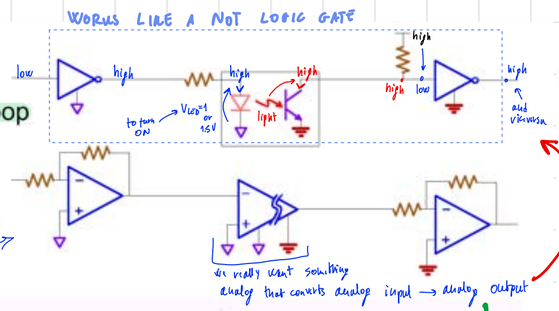
There are ISOs realized in this way, as shown in Fig. 4.21. It is also possible to use discrete components (2 Op-Amps, 2 optocouplers). The feedback compensates for the non-linearity of the photodiodes' I-V-light relationship, and the LED ends the relative ageing. It is a fast transfer method, insensitive to the external EMI interference, does not need modulation and demodulation for the signal within the ISO, and then does not have output ripple. It is vastly used for high impedance sources and sensors.

The optical insulation allows interrupting large and dangerous (EMC) ground loops and coils that can collect noise and, in turn, radiate them to the outside world.

A complete signal acquisition and preconditioning network can be realized with ISOs and PGAs, simply by adding other discrete optocouplers which separate the digital control lines (Fig. 4.22). The user can select the gain of the PGA100 (from 1V/V to 128V/V) with 8 steps, selecting one of the 8 inputs, and the PGA output correspondingly varies between +10V and –10V with a supply voltage of 15V. Through a 1MΩ resistor, there is the voltage-current translation. The ISO100 is in unity-gain configuration. The ISO supply is not shown in Fig. 4.22: a DC/DC converter must be used to provide at least ±40mA.

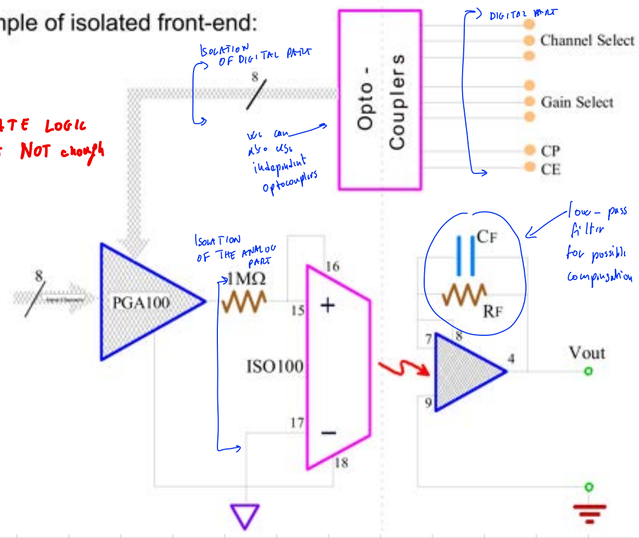
Advantages:
- "light" breaks the "electrical" loop
- high EMI
- trivial for digital signals
- problematic for analog signals

WORKS LIKE A NOT LOGIC GATE

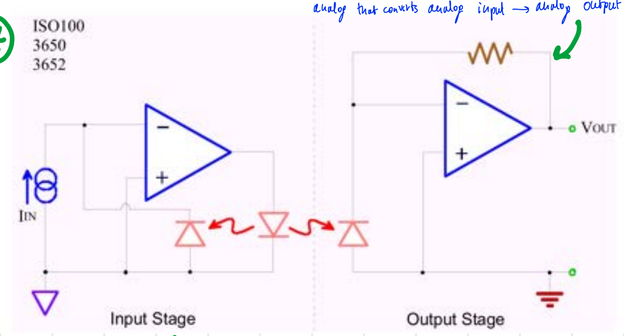


GATE LOGIC is NOT enough

Example of isolated front-end:



Example of optical coupling for analog signals:



Ex. Usually real optocouplers are done with different chips:

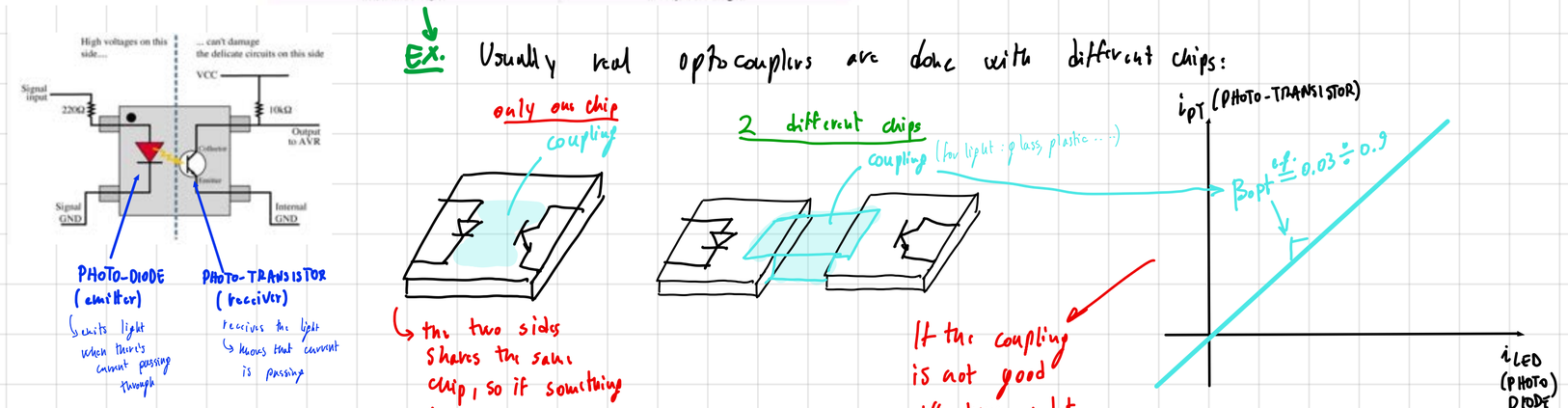
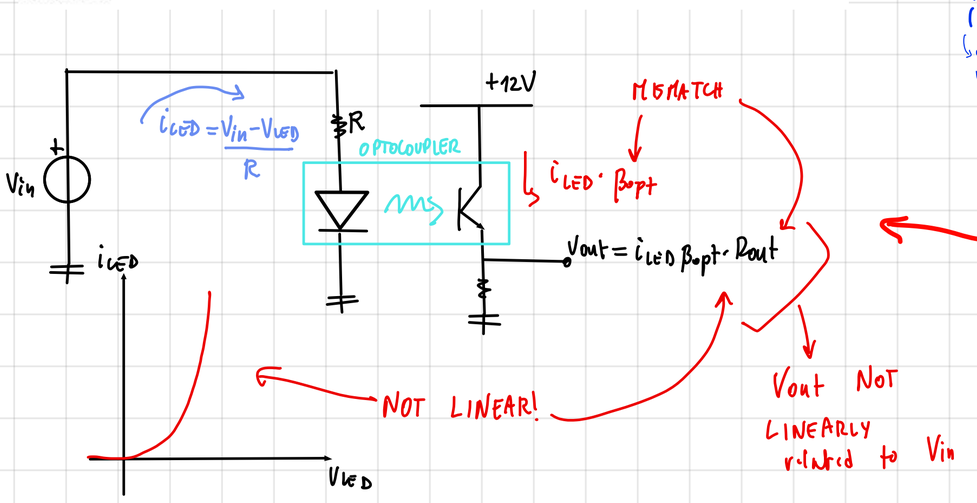


PHOTO-DIODE (emitter) → emits light when there's current passing through

PHOTO-TRANSISTOR (receiver) → receives the light → knows that current is passing

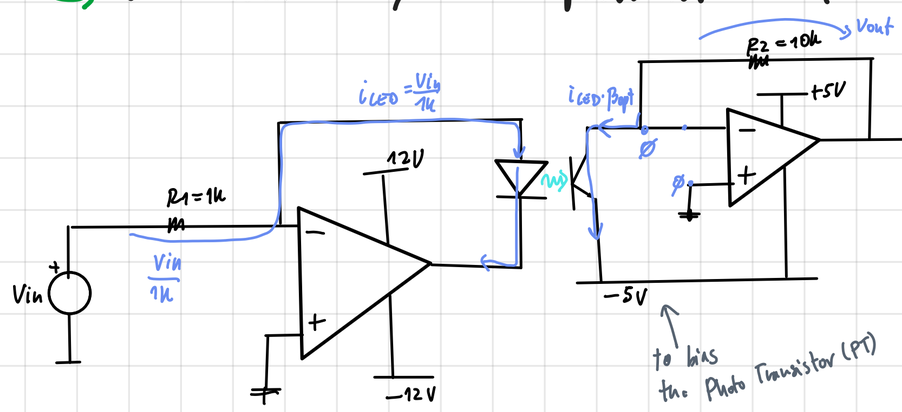


MISMATCH

Vout NOT LINEARLY related to Vin

We can NOT use an optocoupler in this way

→ We can try to improve the circuit in this way:



improvement

but this can give mismatch

lot of tolerance

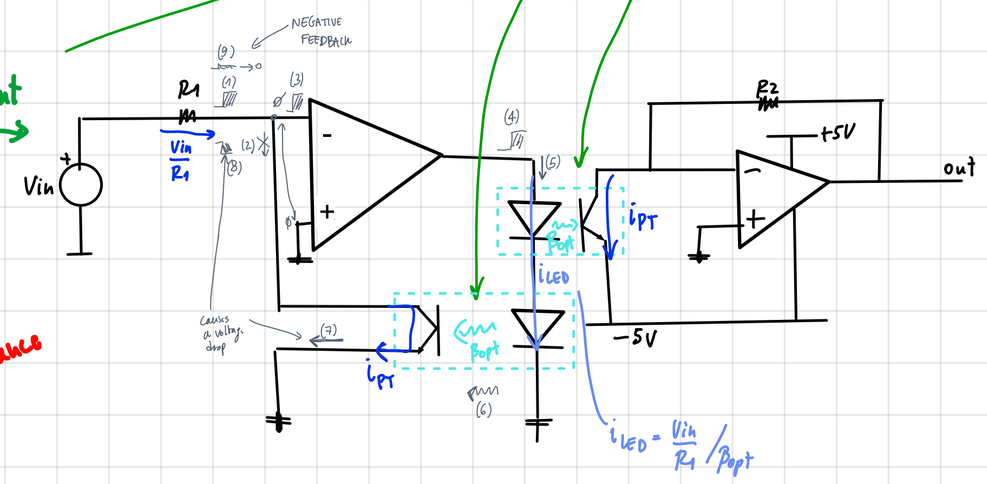
Vout = Vin / 1k · Bopt · 10k

the connection is LINEAR and is only optical

↓

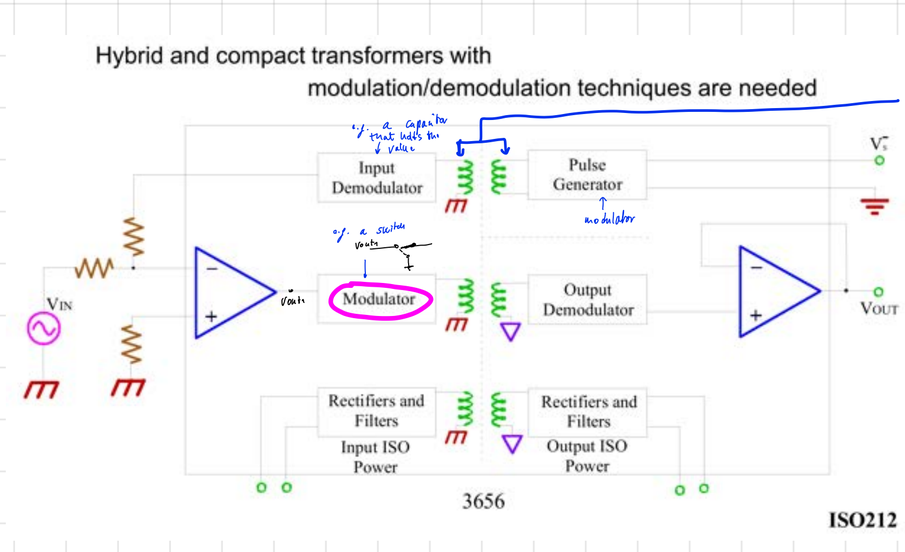
ISOLATION

USE TWICE THE COMPONENT (IDENTICAL COMPONENTS) TO COMPENSATE THE MISMATCH



→ Same solution of ✱

Magnetic coupling (Not very used)



We can isolate 2 parts through the magnetic effect of the transformers

→ TRANSFORMERS require an AC input → so to isolate DC signals we have to modulate them

→ Capacitive coupling (Skipped basically)

NORTON (Current Mode Amplifier - CMA)

(Book p. 325)

The current feedback, presented in the preceding section, is not an approach completely based on current mode because not all the nodes have low impedance. In this section, we will obtain and analyze, making considerations different from those made for the VOA, a structure completely different from the normal feedback configuration of the VOA. The preceding section has highlighted the fact that the CFA is not only an improved architecture (starting from the classical VOA), but conceptually different from the VOA itself. The analyzed positive and negative aspects of the CFA depend on the particular input impedance situation. Nonetheless, it is not correct to say that the CFA is an OpAmp totally current mode. Often we tend to confuse the *current mode* with the *current feedback*; this is due to the fact that the current amplifiers are still in an experimental phase and not present on the market. After this introduction, we can start studying these systems.

The idea is to obtain a circuit in which all the nodes have low resistance to "work with current" in the whole system, from the input to the output. We want to design an OpAmp whose both input terminals have low impedance values to read the input current, as shown in Fig. 4.55.

The internal architecture is based on the translinear principle. The translinear principle is introduced by Barrie Gilbert in a famous paper in the 1972. This principle takes its name from the fact that the BJT transconductance is linear proportional to its collector current, under the normal forward-active working conditions. The most common example of a translinear circuit is the current mirror; the classical 4 transistors output *class AB* stage can be seen in translinear terms.

Examine the Fig. 4.56 that represents a typical input stage: they are BE junctions, one opposite to the other and form closed loops (or closed meshes).

Observe that bipolar transistors are used because they ensure high transconductance values and are good current wells (1/gm as emitter). It is a system with 2 inputs and 4 outputs, which transfers the input signal towards the output. Nevertheless, more importantly, the stage equally treats the differential and common-mode signals; in other words, it has a CMRR equal to 0dB. In fact, due to the translinear principle, we have: $I_1 \cdot I_3 = I_2 \cdot I_4 = I_0^2$ and if $I_{in} \ll I_0$, we have: $I_1=I_2=I_0-I_{in}/2$ and $I_3=I_4=I_0+I_{in}/2$.

To improve the CMRR of the input stage and effectively reject the common-mode currents, we can use the stage shown in Fig. 4.57. Notice the presence of two translinear cells (T_1, T_2, T_3, T_4 and T_1', T_2', T_3', T_4') in which the second serves to eliminate the common-mode currents.

Fig. 4.55

Requirements:

- both inputs be at LOW impedance
  hence current-driven
- output be at HIGH impedance
  hence current source

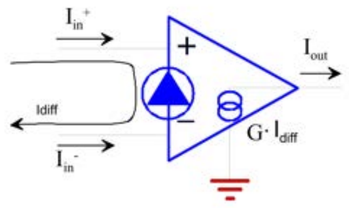
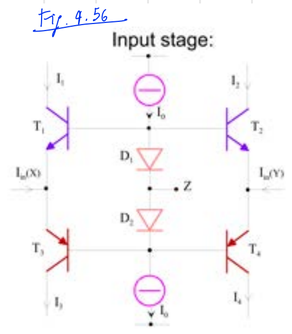
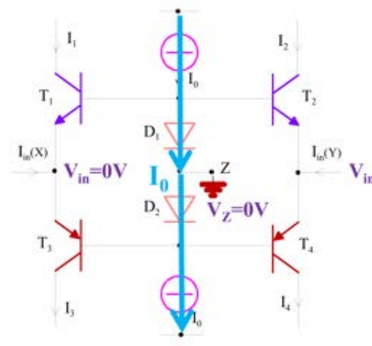


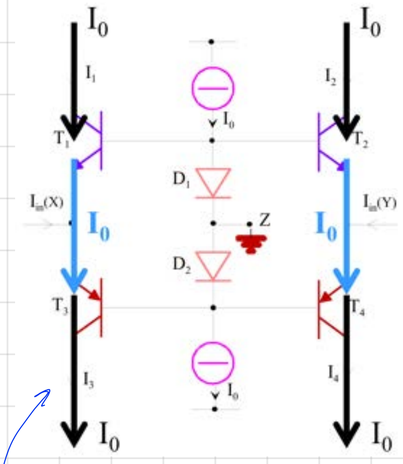
Fig. 4.56

Input stage:





Thanks to pin Z
the voltage at the inputs is set



Translinear principle:

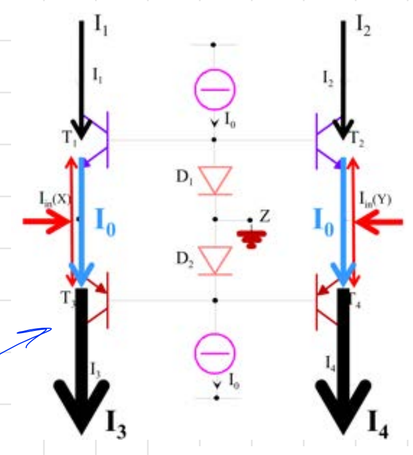
$$I_1 \cdot I_3 = I_2 \cdot I_4 = I_0^2$$

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 $I_3=I_4=I_0$



With common-mode input signal

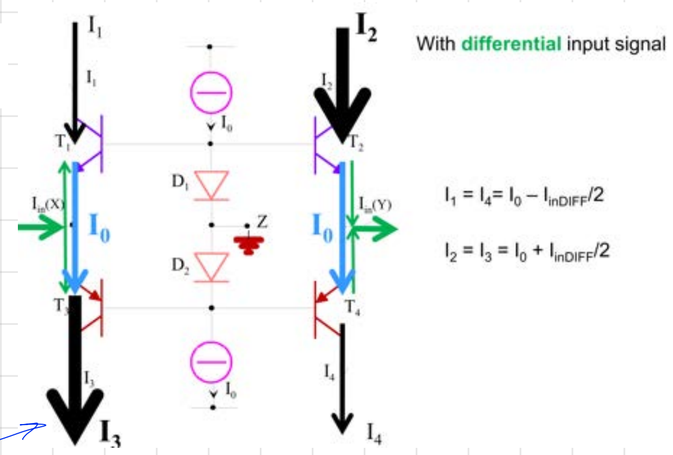
$I_1 = I_2 = I_0 - I_{inCM}/2$

$I_3 = I_4 = I_0 + I_{inCM}/2$

The stage is designed with I_E=500μA. We have obtained a CMRR of about 50dB for low input currents and of 65dB for input currents of about 2mA. Even if they are not very high values, do not forget that the stage previously seen had a CMRR of 0dB. (improved wrt odB) However, there are more optimized and modern cells which reach values up to 100dB.

We had seen the input stage, and sent the output currents I_1, I_2, I_3, and I_4 towards current mirrors to sum them up and extract the differential information.

The scheme shown in Fig. 4.58 uses 6 current mirrors. Through the use of the current mirror, at the output, we obtain: $I_{out+} = I_{in+} - I_{in-}$ and $I_{out-} = I_{in-} - I_{in+}$. Unfortunately, despite the applied improvements, the amplifier does not amplify the current signal.



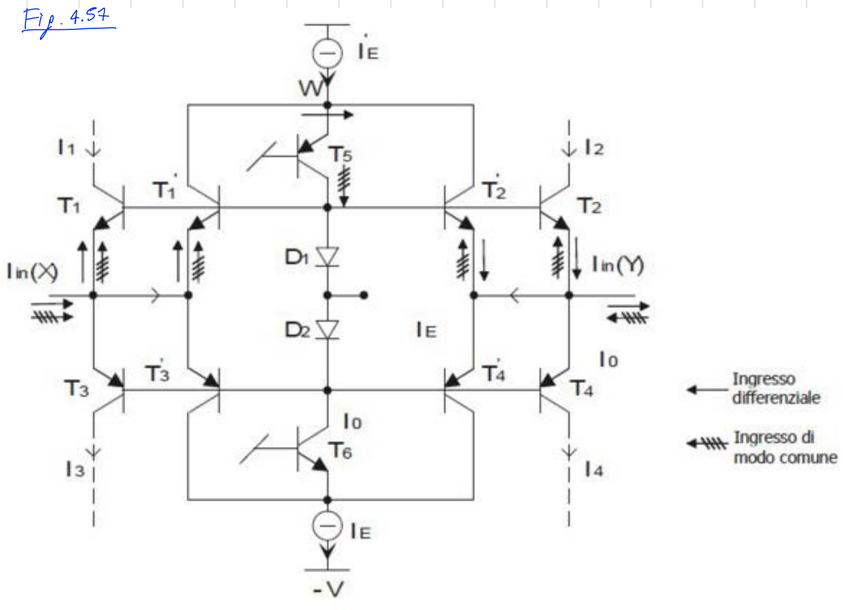
With differential input signal

$I_1 = I_4 = I_0 - I_{inDIFF}/2$

$I_2 = I_3 = I_0 + I_{inDIFF}/2$

(Better explanations in observations)

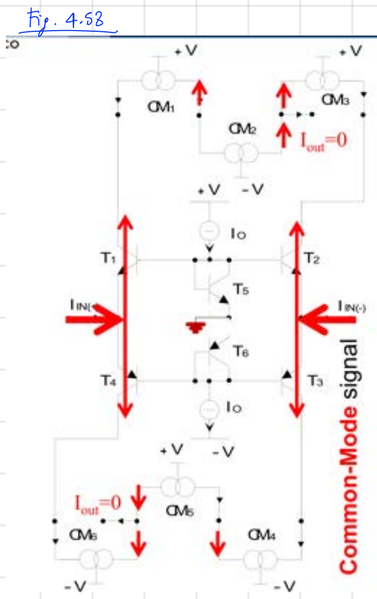
Fig. 4.54

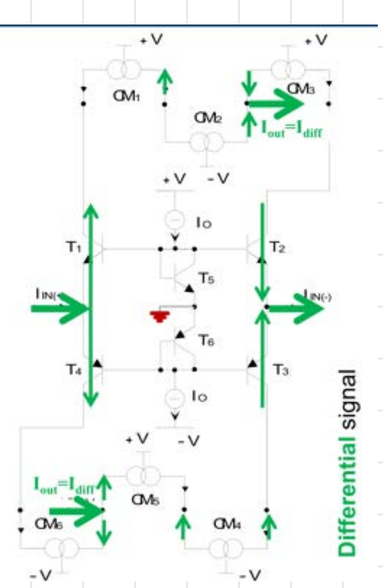


Ingresso differenziale

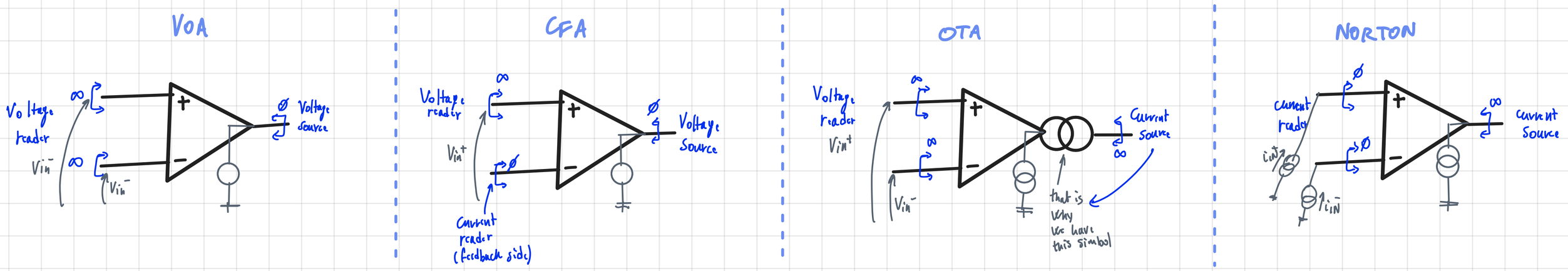
Ingresso di modo comune

Fig. 4.58



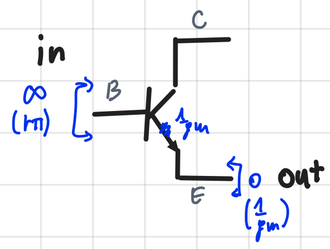


Ex. RECAP wrt other seen OpAmp.



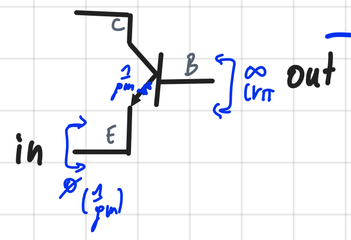
Observations

When we wanted ∞ input impedance we used this transistor configuration:

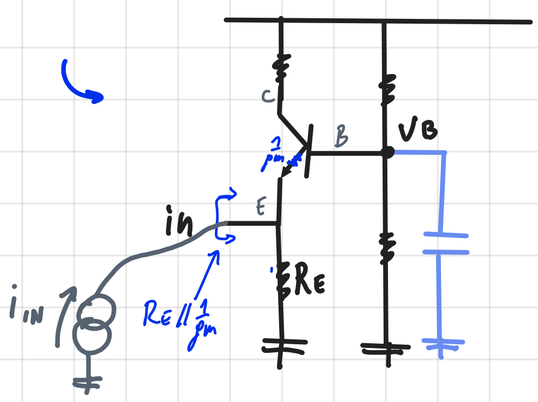


→ high input impedance

So now we want instead ∅ input impedance, therefore we can use this configuration:



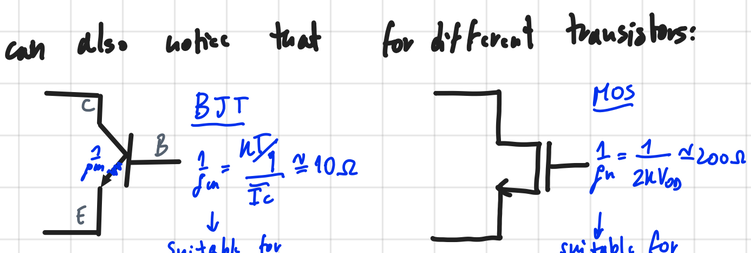
→ low input impedance



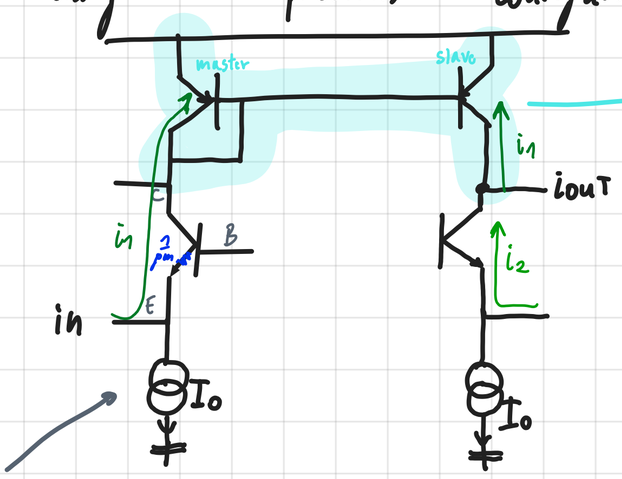
← We bias it with this configuration, V_B is const.,

If we don't want V_B to change we can add a capacitor in this way for high freq. V_B is still very const. →

We can also notice that for different transistors:



Therefore using the previous configuration for both ⊕ and ⊖ input:

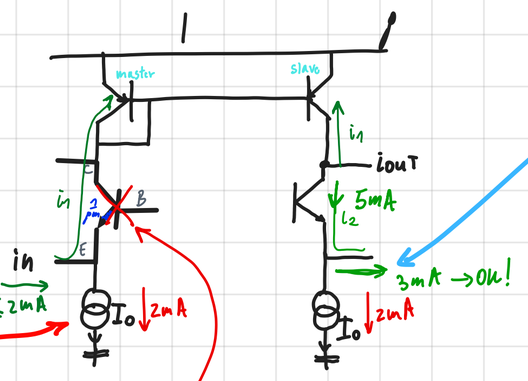


→ Current mirror → in this way the two currents subtract with eachother

i_1 i_2 → if $i_1 = i_2$ ⇒ $I_{out} = \emptyset$

It's like choosing $R_E >> \frac{1}{g_m}$ → highest resistance component → current source

→ But this configuration has an issue, indeed we cannot pump in the ⊕ and ⊖ input more current than I_0

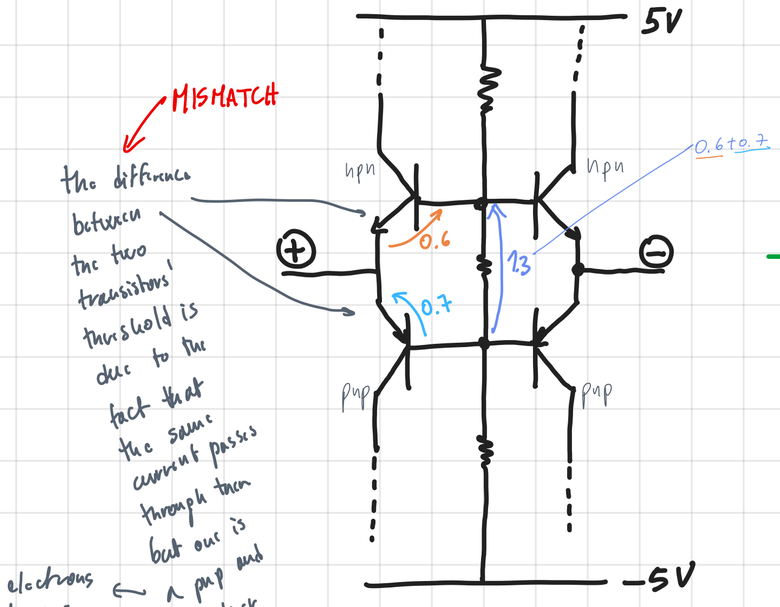


because if we pump 2mA it means we're pushing this value on the current generator meaning we leave the transistor with NO current → OFF

If we want to bring out current, there's no problem

↑ the issue is when we want to pump in current

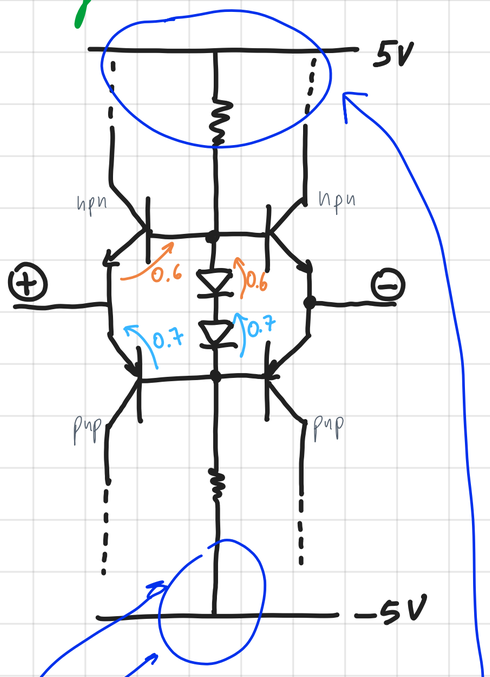
to solve this we need to introduce a double configuration.



the difference between the two transistors' threshold is due to the fact that the same current passes through them but one is a pnp and the other an npn

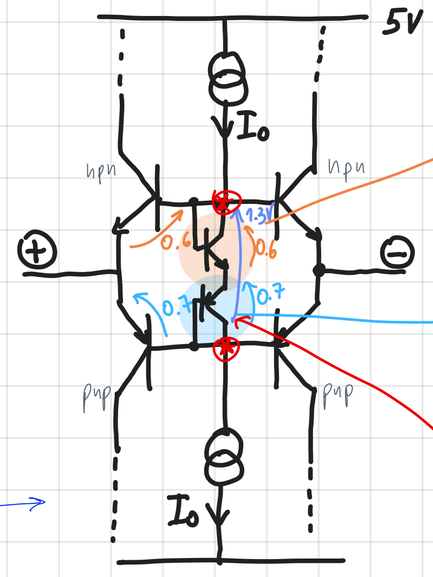
electrons have a higher mobility with respect the holes → usually for the same dimension the voltage drop across an npn will be lower, 'cause the efficiency of the holes is worst wrt the electrons

Improvement of the mismatch → add diodes



we can see it as current generator to pump current in and out

≡



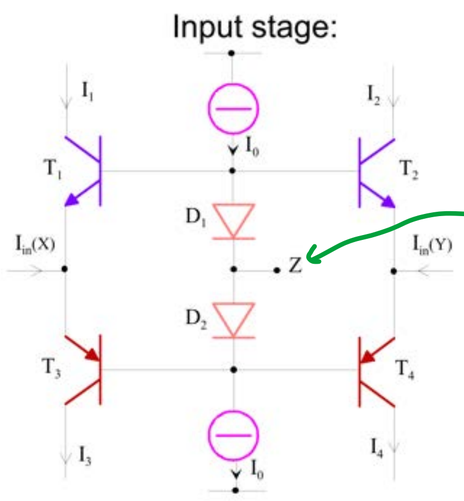
this diode is the same of the npn (connected as photo-diode)

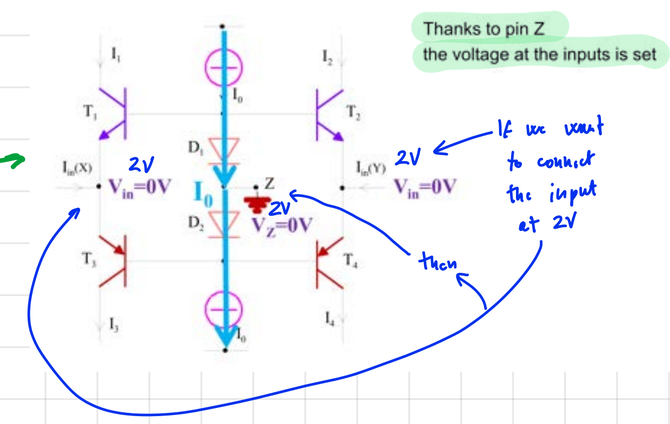
this diode is the same of the pnp (connected as photo-diode)

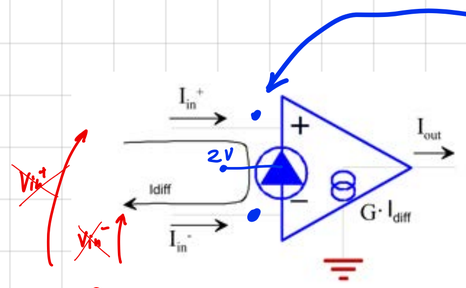
we know the voltage across the two diodes but not the voltage here ⊛

to fix these voltages we can fix at intermediate pin Z

Fig. 4.56



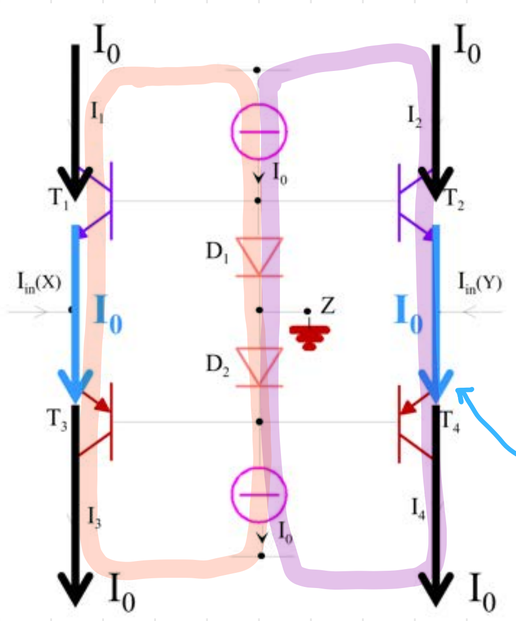




but we can set it through the currents thanks to pin Z

we cannot read voltages with our low imp.

→ Now if we apply the translinear principle:



Translinear principle:

$$\underbrace{I_1 \cdot I_3}_{CCW} = \underbrace{I_2 \cdot I_4}_{CW} = I_0^2$$

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 $I_3 = I_4 = I_0$

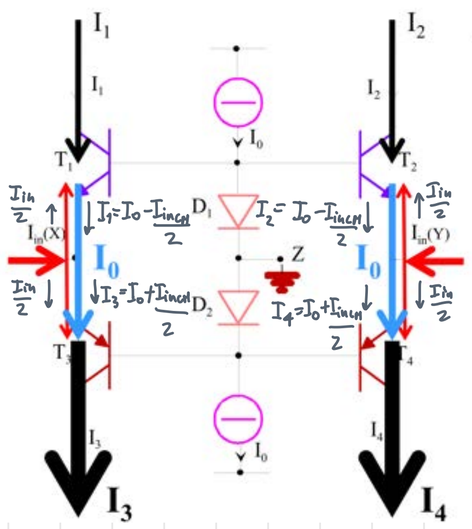
If now we apply a current at the input we have

(If all transistors are the same)

computations:

- $I_1 \cdot I_3 = I_0 \cdot I_0$
- $I_0 \cdot I_0 = I_2 I_4$

$\rightarrow I_1 I_3 = I_2 I_4 = I_0^2$



while when we a diff. input signal $I_{in}(X) = -I_{in}(Y) = I_{inDIFF}$

With common-mode input signal

→ 2 equal input currents

$I_{in}(X) = I_{in}(Y) = I_{inCM} \neq 0$

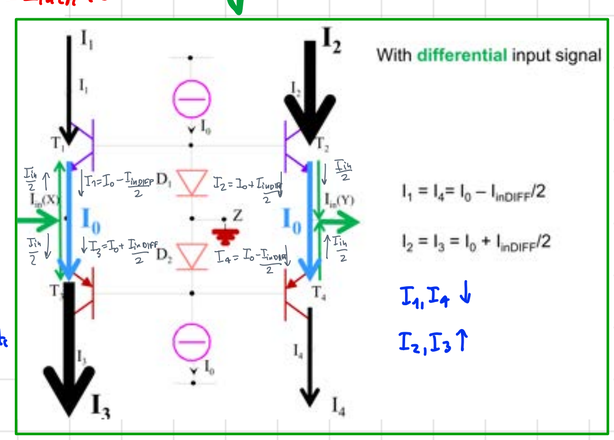
$I_1 = I_2 = I_0 - I_{inCM}/2$

$I_3 = I_4 = I_0 + I_{inCM}/2$

$I_1, I_2 \downarrow$

$I_3, I_4 \uparrow$

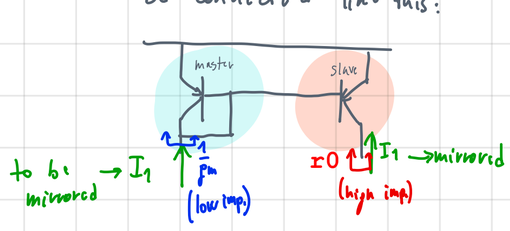
Based on their increase/decrease we can discriminate when it's CM or DIFF input signal

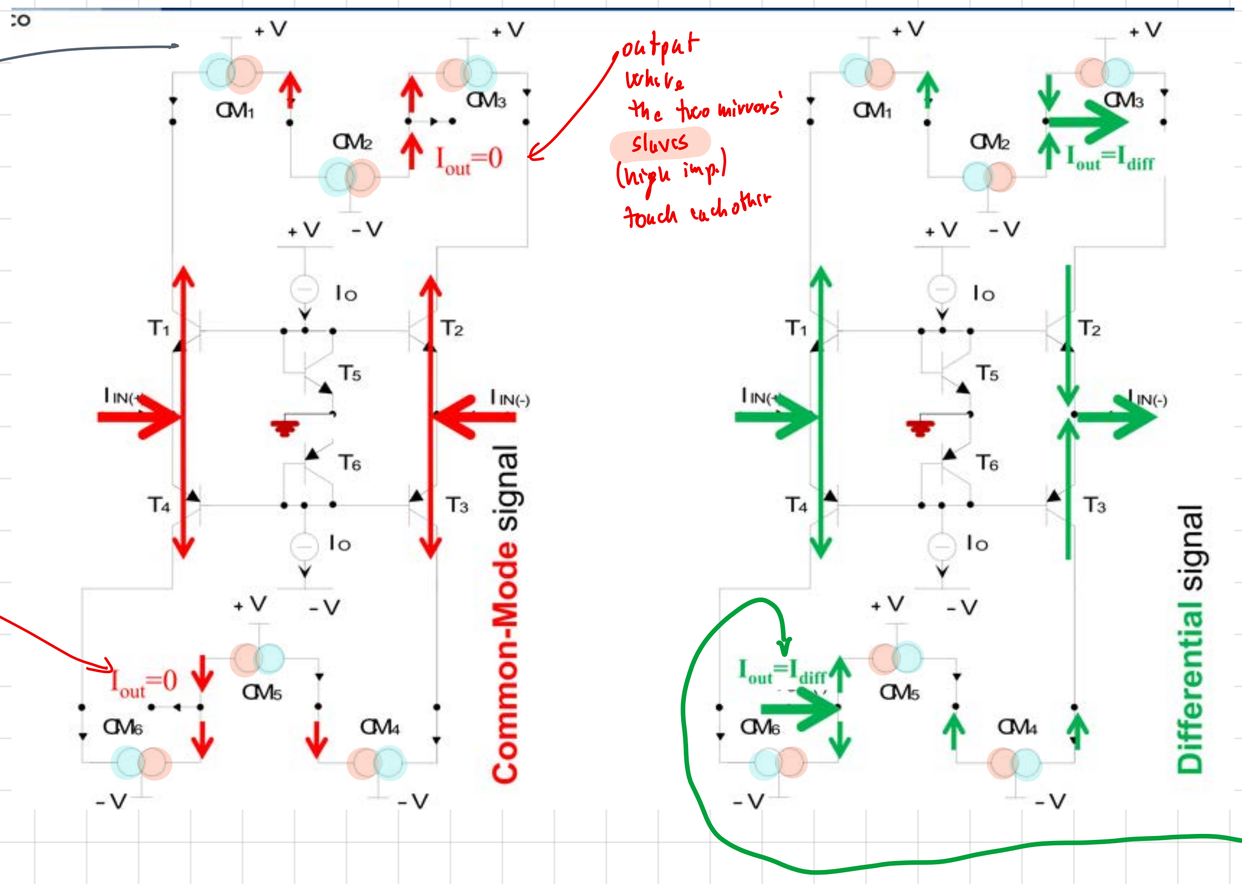


while when we a diff. input signal $I_{in}(X) = -I_{in}(Y) = I_{inDIFF}$

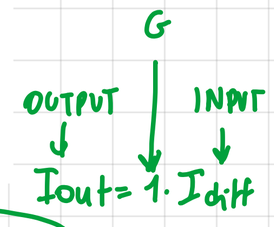
↳ We can now introduce currents mirrors in order to use the s. CM & DIFF input signal:

this current mirror should be connected like this:

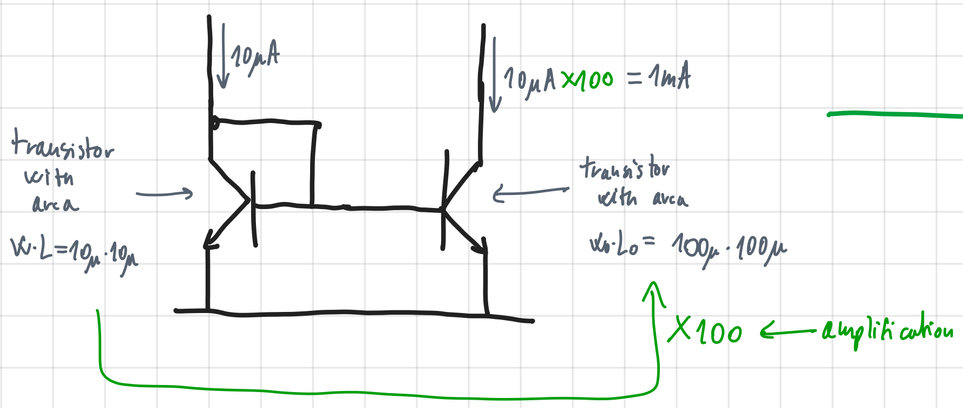


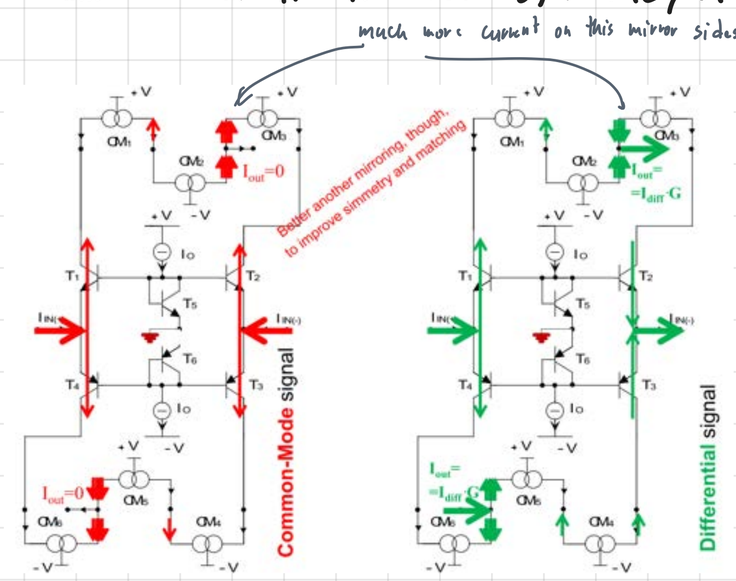


amplification → We can use two outputs to double the input current signal

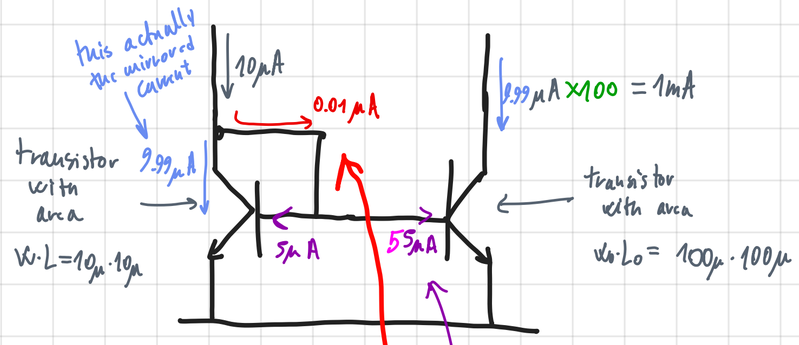


↳ But here if we use this configuration the gain of the amplifier will be 1 at most, so if we want to increase the amplification we can use a current mirror that it is not symmetric, meaning a mirror side amplifies the other:





↳ BUT we can NOT increase to much the transistor areas because we also have to consider the base current:



but limited by this

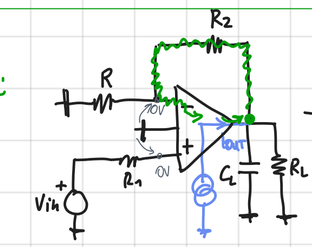
If the area of this ↑ also this current ↑ ↳ for ex. 55 μA

and also the mirroring factor will be impaired

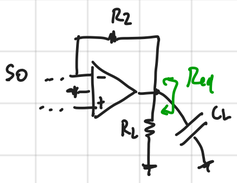
usually x10 amplification at most!

↑ the Ai amplification is limited

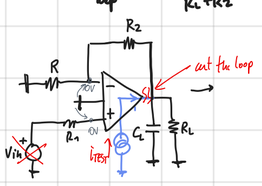
Ex.



↳ so



$G_{loop}: i_{loop} = i_{TEST} \frac{R_L}{R_L+R_2} \cdot A_i \rightarrow G_{loop} = \frac{R_L}{R_L+R_2} A_i$



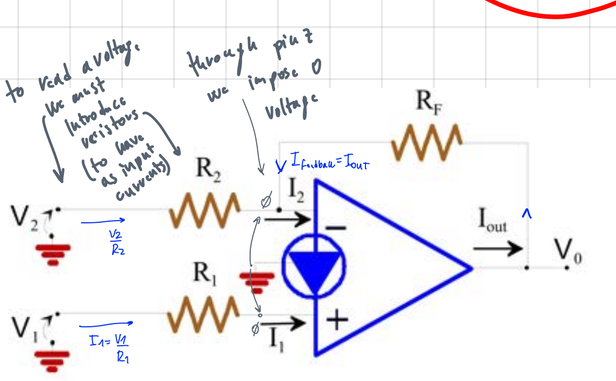
$\Rightarrow pole = \frac{\frac{R_L}{R_L+R_2} \cdot A_i}{2\pi C \frac{R_L R_2}{R_L+R_2}} = \frac{A_i}{2\pi C R_2}$

Advantages:

- all nodes have low-impedance ($1/g_m$) — hence very low parasitic time constants
- very wide bandwidth ↳ pole = $\frac{1}{2\pi C R_{loop}}$ ← big BW (R_{loop} small)
- bandwidth independent of closed-loop gain

to read voltage we must introduce resistors (to have as input currents)

through pin 2 we impose 0 voltage



$$V_O = \frac{A_i}{1+A_i} \cdot \left(V_1 \cdot \frac{R_F}{R_1} - V_2 \cdot \frac{R_F}{R_2} \right)$$

Note that A_i is NOT infinite but just few tens !

... hence **NO** "IDEAL GAIN" behaviour

$$I_1 = \frac{V_1}{R_1}$$

$$I_2 = \frac{V_2}{R_2} + I_0$$

$$I_{out} = I_0 = \frac{V_0}{R_F}$$

$$I_0 = A_i \cdot I_{diff} \quad (I_1 - I_2)$$

$I_0 = A_i \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} - I_0 \right)$

↳ $I_0 (1+A_i) = A_i \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} \right)$

↳ $I_0 = \frac{A_i}{(1+A_i)} \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} \right)$

$V_0 = R_F I_0$

(Book p. 334)

Practically, it is simpler to think that the *current mode* OpAmp "reads" the current of the positive input (both AC and DC) and acts to impose it on the negative input with the same sign and the amplitude.

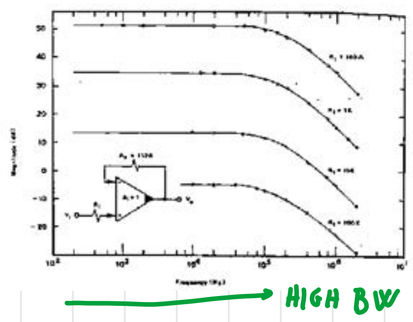
It is appropriate to highlight that the closed-loop bandwidth is wide and that it does not change by varying the closed loop. The circuit in Fig. 4.63 has the behavior depicted in the same figure. In this case, A_i=1, therefore, for R_1=1kΩ the voltage gain will be:

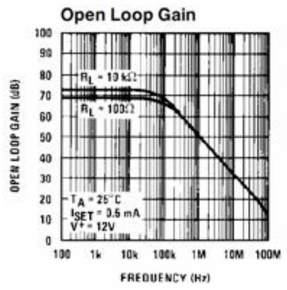
$$\frac{R_F}{R_1} \cdot \frac{Ai}{1+Ai} = \frac{1}{2} \cdot \frac{R_F}{R_1} = \frac{112}{2} = 35dB$$

(from the VOA config.)

and not R_F/R_1, as you would instinctively think!

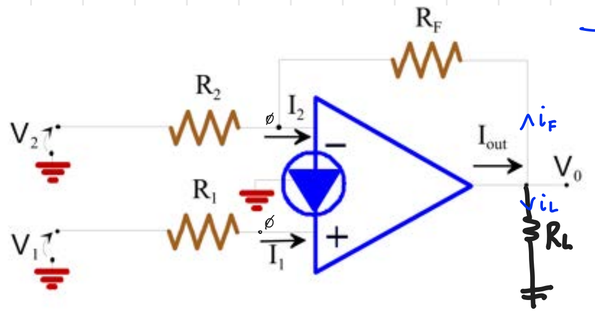
Often, these OpAmps have a control pin into which a current can be injected, controlling the biasing of the internal current mirrors. In this way, you can change the performance of the Op-Amp, in terms of: Gain-Bandwidth Product, Slew Rate, Input Bias Current, Supply Current, and Output Sink Current in order to better adapt the OpAmp to the specific application of interest. For details, please consult the attached data-sheets.





HIGH BW

↳ If we add an external load



$I_{out} = I_F + I_L = A_i (I_1 - I_2) = A_i \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} - I_F \right)$

$I_2 = \frac{V_2}{R_2} + I_F$

$I_1 = \frac{V_1}{R_1}$

$V_0 = R_L I_L$

$V_0 = R_F I_F$

$i_L = \frac{R_F}{R_L} I_F$

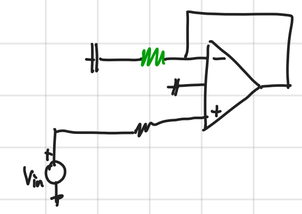
$I_F \left(1 + A_i + \frac{R_F}{R_L} \right) = A_i \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} \right)$

the gain depends also on the load R_L

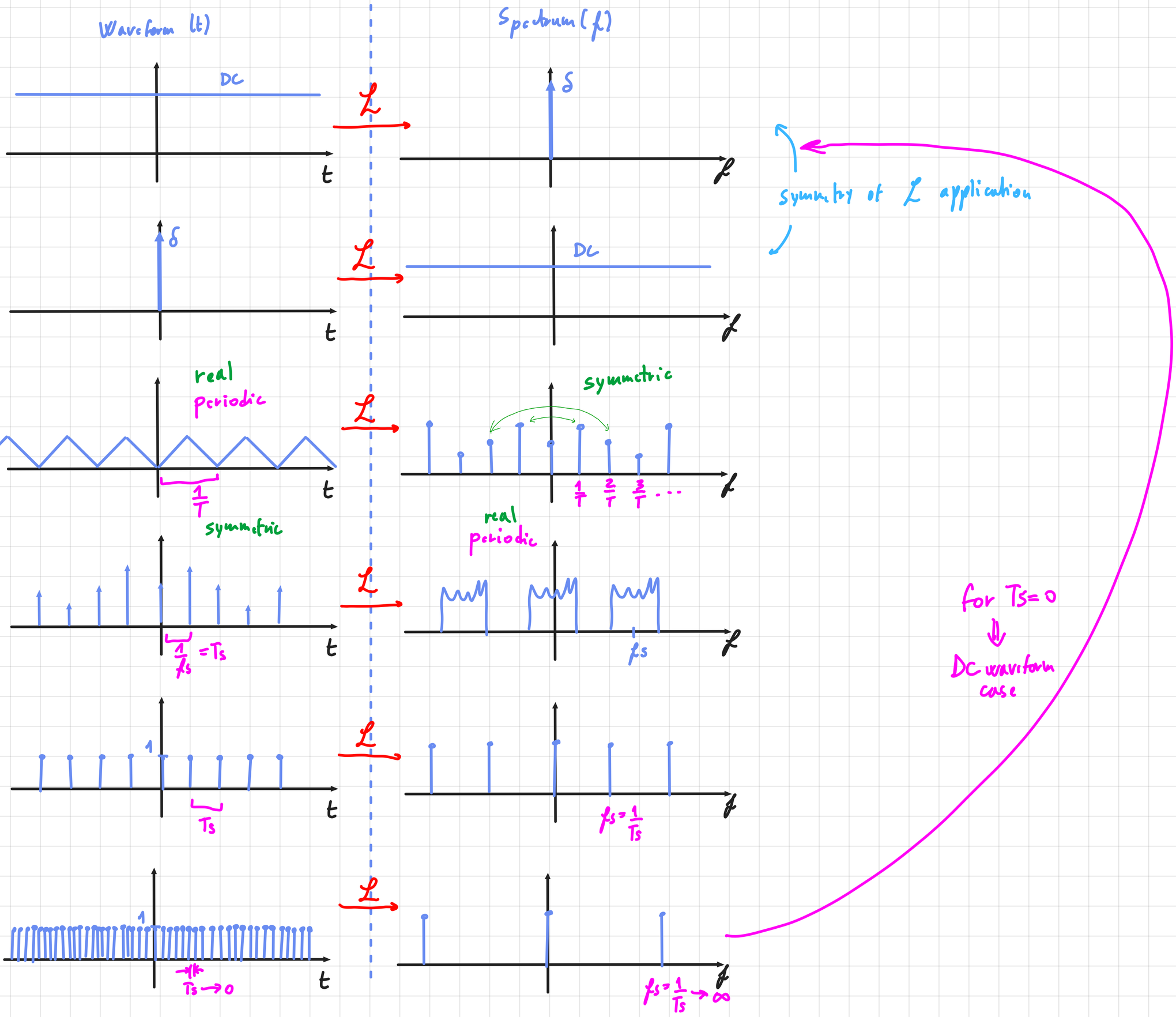
$I_F = \frac{A_i}{1 + A_i + \frac{R_F}{R_L}} \left(\frac{V_1}{R_1} - \frac{V_2}{R_2} \right)$

$V_0 = R_F I_F$

Note: If we don't want to connect any voltage to an input pin, let's put any way a resistance, because due to the mismatches of the internal circuit there will be any way a small volt. therefore a big curr.



Waveform vs Spectrum



Sampling theory

(Book p. 443)

Sampling consists in the observation of analog quantities (time and amplitude variable) at certain time moments. If these moments are periodically and uniformly distributed, we talk about *free-running sampling*; otherwise, if the sampling is non-periodic, it is named *single-shot*. In the case of sampling with a period of T_S, i.e. at a sampling frequency $f_S=1/T_S$, the Sampling Fundamental Theorem gives the condition to be met in order that the original signal and the sampled signal have the same information content. The importance of the theorem is attributable to the fact that a discrete-time signal can represent a continuous-time signal without altering the information. This theorem is attributed to C.E. Shannon, who first published a paper on this topic in 1949.

Fig. 6.3 depicts the spectrum of a sampled signal for various sampling frequencies. The graph shows that only a sampling that respects Shannon's Theorem allows preserving the starting base-band spectrum (i.e. below f_{max}) and then reconstructing the original signal by means of a simple low-pass filter, as those shown in Fig. 6.4.

Generally, the sampling operation relies on a circuit named *Sample&Hold* (described in detail in this chapter), which does not generate analog pulse samples, but periodically picks the signal setting the amplitude at the output for a period T_S, as shown in Fig. 6.5.

Because of this, unlike what was found by the ideal sampling, satisfying also the constraint imposed by the sampling theorem, after the reconstruction of a signal from an S&H, you get a distorted signal shown in Fig. 6.6. Note that, assuming to sample with a frequency $f_S=2\cdot f_{max}$, the error caused by the *sinc(f)* distortion is -3.9dB, which is non- negligible. In the case of an input signal with a frequency lower than the sampling one, the introduced distortion cannot be neglected. (For example, with $f_{in}=f_S/100$, the gain is reduced to 0.9998, corresponding to an error of about 1 LSB in a 12 bit conversion, as we are going to see in the next chapters about digital-to-analog and analog-to-digital converters).

Luckily, amplitude and phase distortions are linear, and it is possible to cancel them, equalizing the reconstruction filter output. Fig. 6.7 and Fig. 6.8 show the blocks needed for reconstruction and equalizing, respectively.

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 $T_S=1/2BW$"*

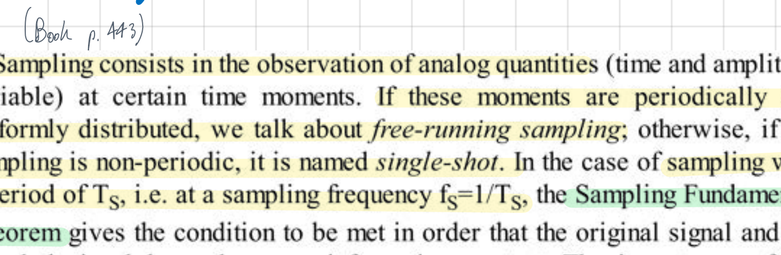


Fig. 6.3

Sampling:

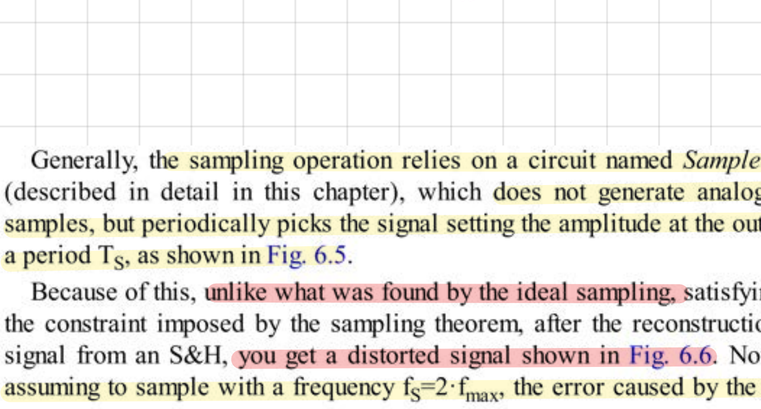


Fig. 6.4

Reconstruction:

Fig. 6.5

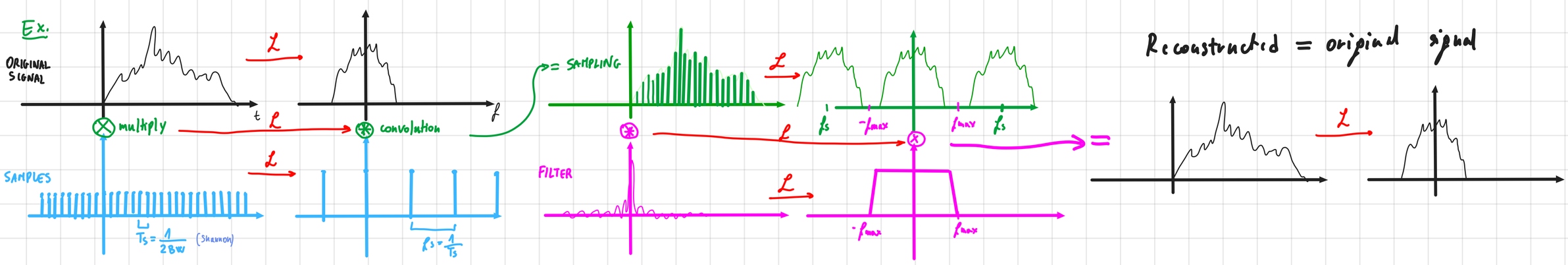
It is easier to employ "rectangular" samples instead of delta-like ones → Explained better later

… but this causes spectral distortion:

Fig. 6.6

Therefore **equalization** is compulsory:

Fig. 6.8



Aliasing

(Book p. 453)

If one samples with a sampling frequency not high enough, the discrete-time samples, once sent to the final user (the time-continuous signal reconstruction), will generate a signal different from the desired one. This problem must be avoided because signal equivocations can rise otherwise. Intuitively, we can understand the aliasing effect by taking into account Fig. 6.13 and Fig. 6.14. In Fig. 6.13, the sampling frequency f_S=10ksps (T_S=100µs) is not high enough to give the correct sinusoid at f_{in}=13kHz. With slack samples, the obtained input signal will thus be considered with a frequency lower than the original one, particularly $f_{aliased}=f_{in}-f_S$=3kHz. Only with $f_S>2\cdot f_{in}$ (i.e. $T_S<½\cdot T_{in}$, when at least two values in the minimum period of the input sinusoid are sampled), the sampled signal maintains the correct information on the original signal. This happens, for example, in Fig. 6.14 when f_S=32.5kHz (T_S=30.8µs, i.e. with about 2.5 samples for every input signal period).

It is the same effect described in the previous paragraph and exemplified in Fig. 6.10. It can be quantitatively explained by analyzing the sampling effect in the frequency domain. The sampling of a signal with a frequency f_S results in the original spectrum repetition (around the integer multiples of the sampling frequency). So, if f_{max} is the bandwidth of the signal, then the minimum sampling frequency must be at least equal to $f_S>2f_{max}$ in order to have no superimposition of the repeated spectra.

The Fig. 6.15 (at the top) shows a generic analog signal and its spectrum; in the centre of the figure, the sampling frequency is lower than the minimum allowed, causing aliasing; finally, at the bottom of the figure, the sampling frequency is increased, i.e. many more samples are taken, and the spectra are more distanced from each other without any superimposition.

To reconstruct the signal, we must use a reconstruction filter that makes a band-pass filtering (low-pass as seen until now) on the obtained samples. We can understand that it is simpler to reconstruct the signal if the sampling frequency is significantly higher than the limit imposed by the Sampling Theorem. Thus, the sampling shown in Fig. 6.16 is really good for the signal reconstruction because the filter that selects the bandwidth can have relaxed features, i.e. it can have a corner frequency not very sharp at the frequency f_{max} (can be less selective).

On the other side, the repeated sampled spectra are superimposed, and we will have *aliasing*, which causes an information distortion and then a wrong reconstructed signal. In Fig. 6.17, we see, in the bandwidth of the signal (more precisely between f_S-f_{max} and f_{max}), that the spectrum is different because of the presence of the replica tails.

Fig. 6.15

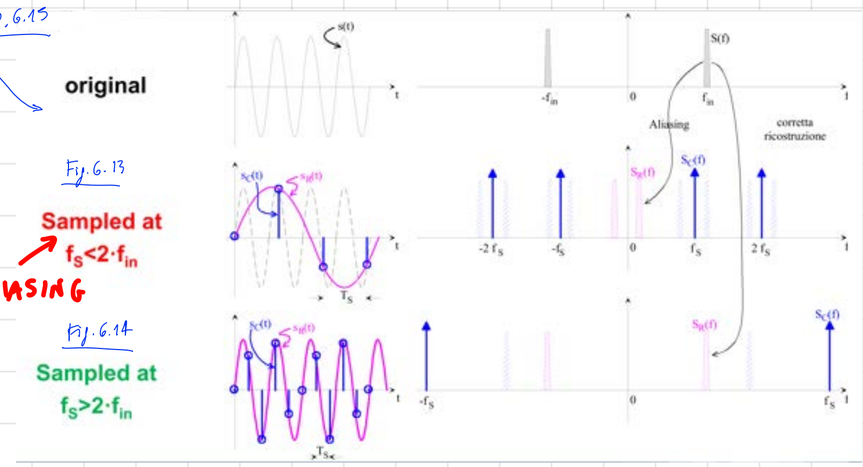


Fig. 6.16

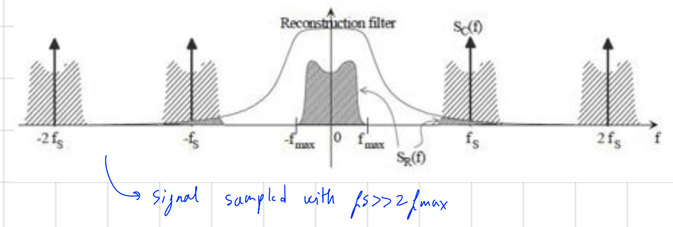
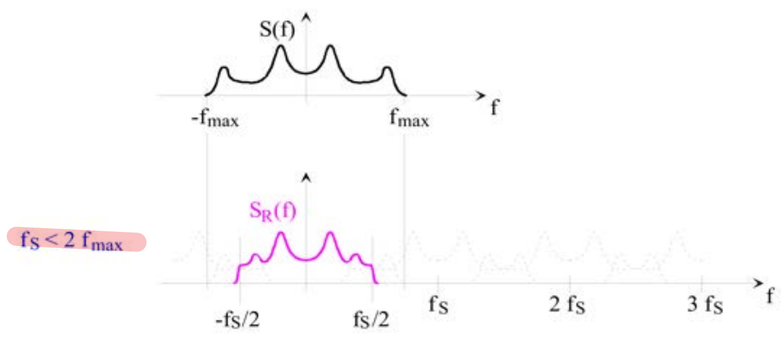


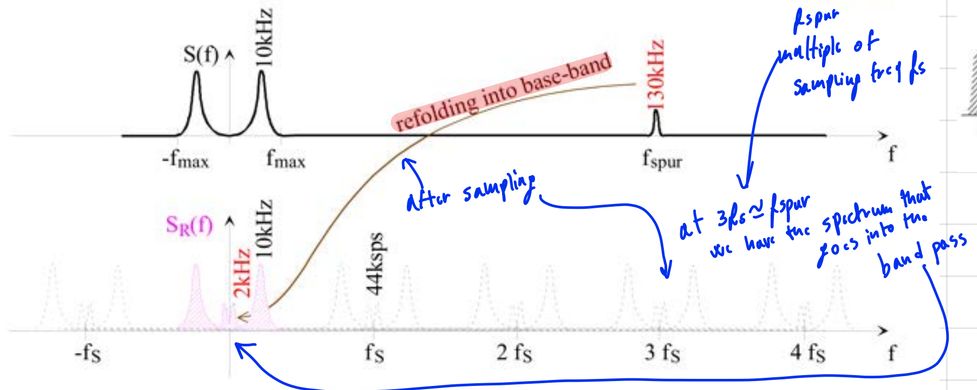
Fig. 6.17

Therefore aliasing drastically deforms the original spectrum



Avoid aliasing BEFORE it comes into play, otherwise no way to remove it !

Aliasing stems out also from unexpected spurious disturbances with $f_{spur}>f_S/2$



Never trust just on the bandwidth self-limitations of the input welcome signal …

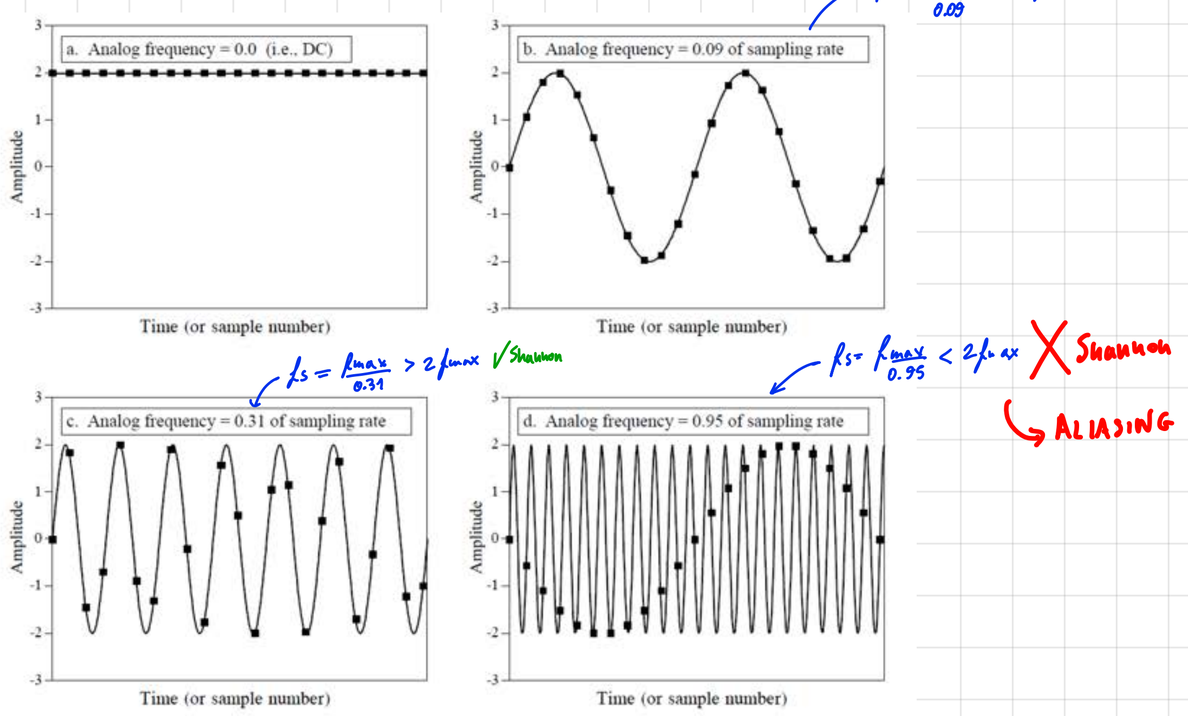
… anyway, bandpass filtering is a must !

Ex.

Misinterpretation (aliasing)

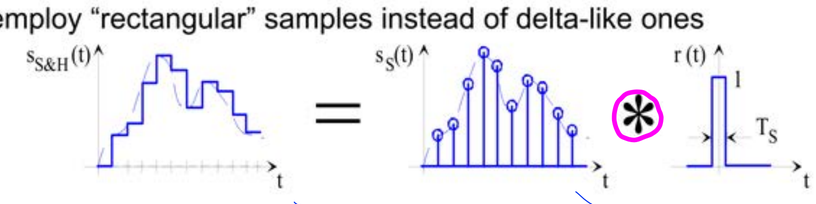
when $f > 0.5\ f_S$

Avoid !

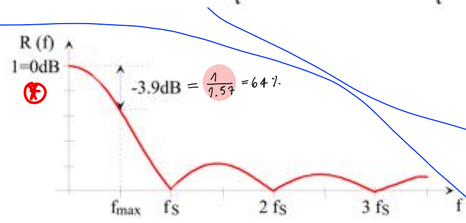


Equalization

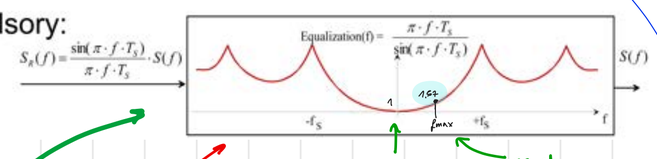
It is easier to employ "rectangular" samples instead of delta-like ones



… but this causes spectral distortion:



Therefore **equalization** is compulsory:



use this filter

can't use just this

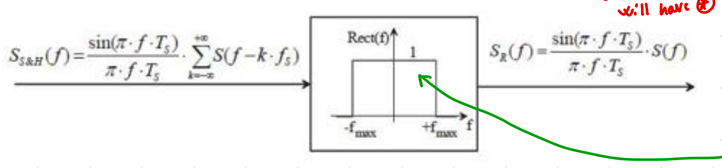
we have to equalize it due to the distortion it will have ⊛

1 at 0 freq.

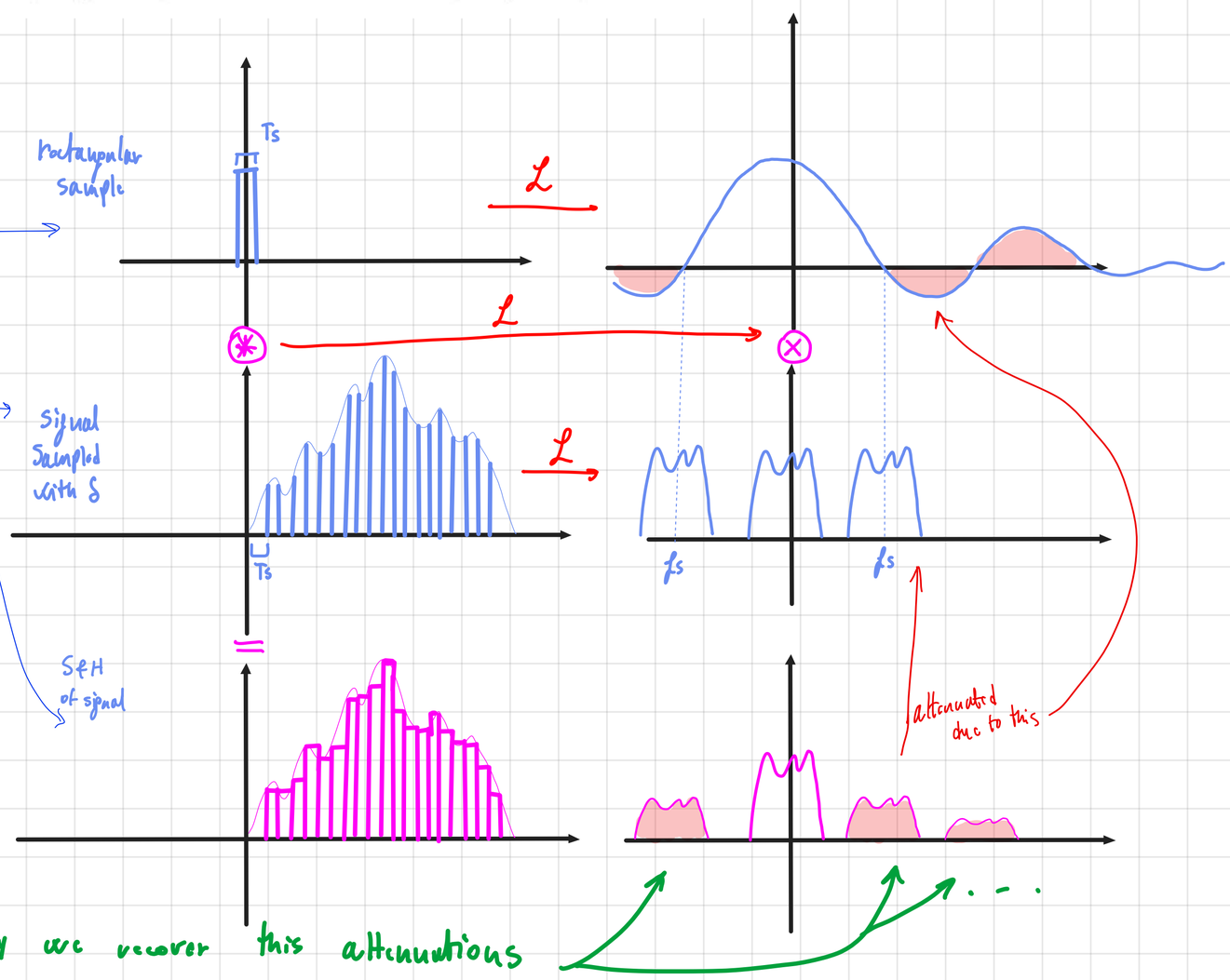
1.57 at fmax ⇓ in this way we equalize the distortion of $\frac{1}{1.57}$ at fmax

because in this case $\frac{1}{1.57}\cdot 1.57 = 1$

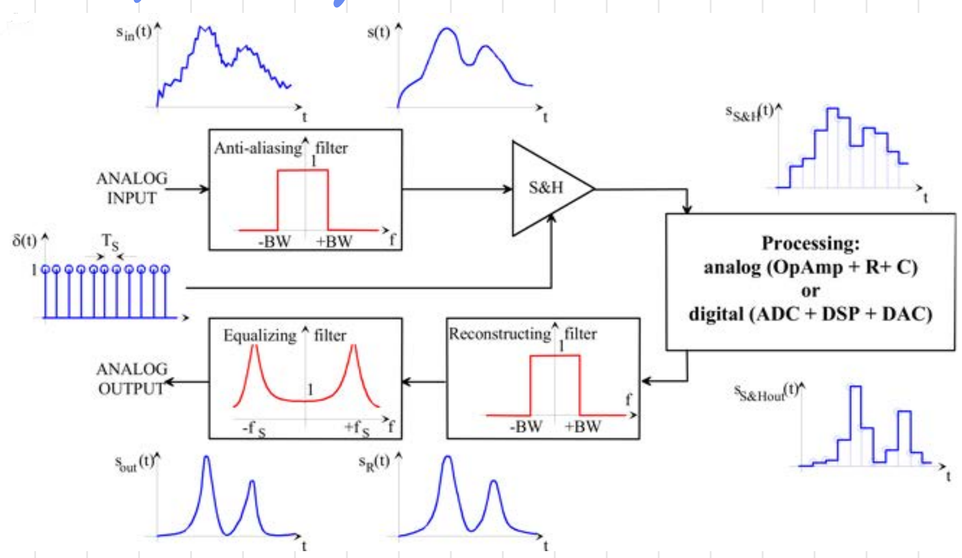
like it should be



In this way we recover this attenuations



RECAP: Signal Processing Chain



Analog Filtering: review

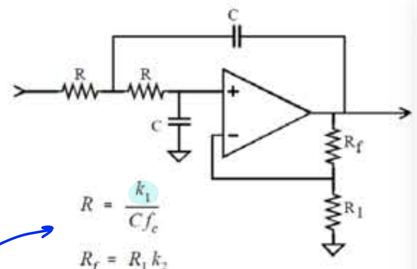


FIGURE 3-8
The modified Sallen-Key circuit, a building block for active filter design. The circuit shown implements a 2 pole low-pass filter. Higher order filters (more poles) can be formed by cascading stages. Find k_1 and k_2 from Table 3-1, arbitrarily select R_1 and C (try 10K and 0.01μF), and then calculate R and R_f from the equations in the figure. The parameter, f_c, is the cutoff frequency of the filter, in hertz.

depending on k_1, k_2 coeff. we can have different filters

TABLE 3-1
Parameters for designing Bessel, Butterworth, and Chebyshev (6% ripple) filters.

# poles		Bessel k_1	Bessel k_2	Butterworth k_1	Butterworth k_2	Chebyshev k_1	Chebyshev k_2
2	stage 1	0.1251	0.268	0.1592	0.586	0.1293	0.842
4	stage 1	0.1111	0.084	0.1592	0.152	0.2666	0.582
	stage 2	0.0991	0.759	0.1592	1.235	0.1544	1.660
6	stage 1	0.0990	0.040	0.1592	0.068	0.4019	0.537
	stage 2	0.0941	0.364	0.1592	0.586	0.2072	1.448
	stage 3	0.0834	1.023	0.1592	1.483	0.1574	1.846
8	stage 1	0.0894	0.024	0.1592	0.038	0.5359	0.522
	stage 2	0.0867	0.213	0.1592	0.337	0.2657	1.379
	stage 3	0.0814	0.593	0.1592	0.889	0.1848	1.711
	stage 4	0.0726	1.184	0.1592	1.610	0.1582	1.913

poles for stages of the filter

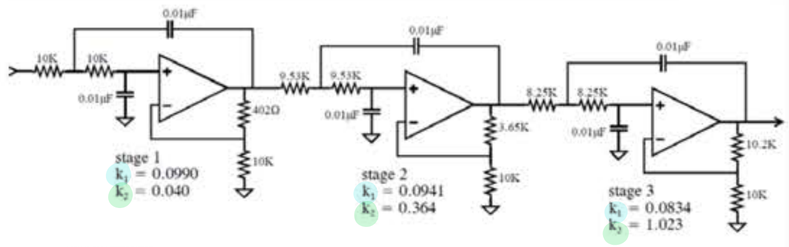


FIGURE 3-9
A six pole Bessel filter formed by cascading three Sallen-Key circuits. This is a low-pass filter with a cutoff frequency of 1 kHz.

↳ Based on requirements we can have different choices of k_1, k_2, so filter type and we then can also choose the number of poles and stages

Log scale — Softer freq. response — Linear scale

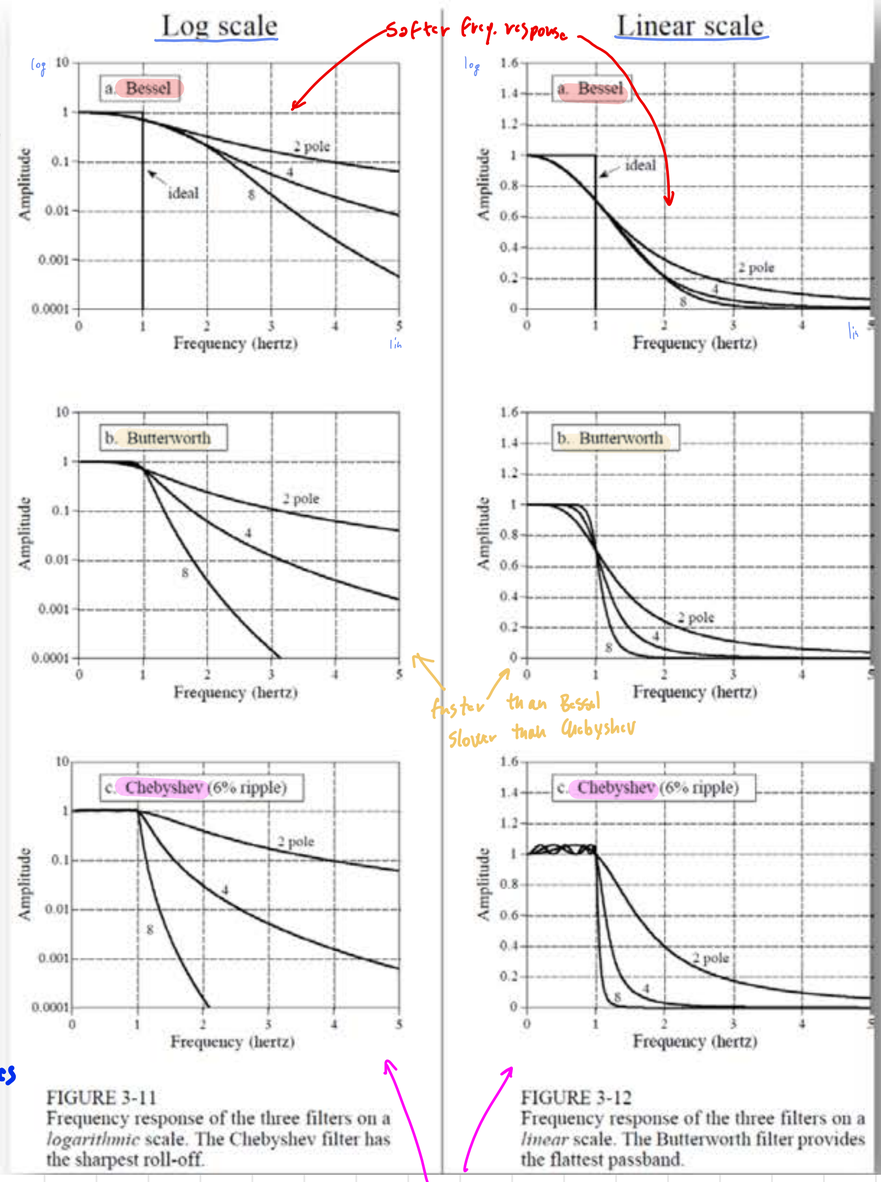


FIGURE 3-11
Frequency response of the three filters on a *logarithmic* scale. The Chebyshev filter has the sharpest roll-off.

FIGURE 3-12
Frequency response of the three filters on a *linear* scale. The Butterworth filter provides the flattest passband.

faster than Bessel, slower than Chebyshev

Sharpest fall → more similar to the ideal rectangular filter

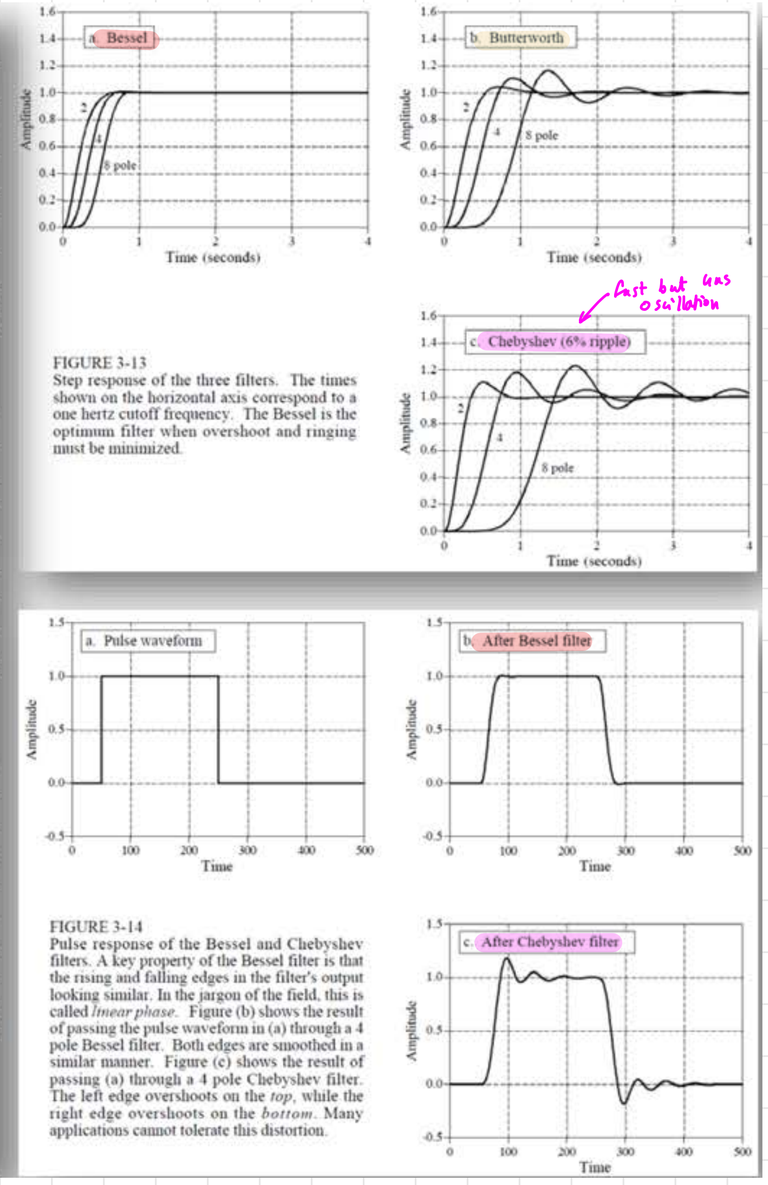


FIGURE 3-13
Step response of the three filters. The times shown on the horizontal axis correspond to a one hertz cutoff frequency. The Bessel is the optimum filter when overshoot and ringing must be minimized.

fast but has oscillation

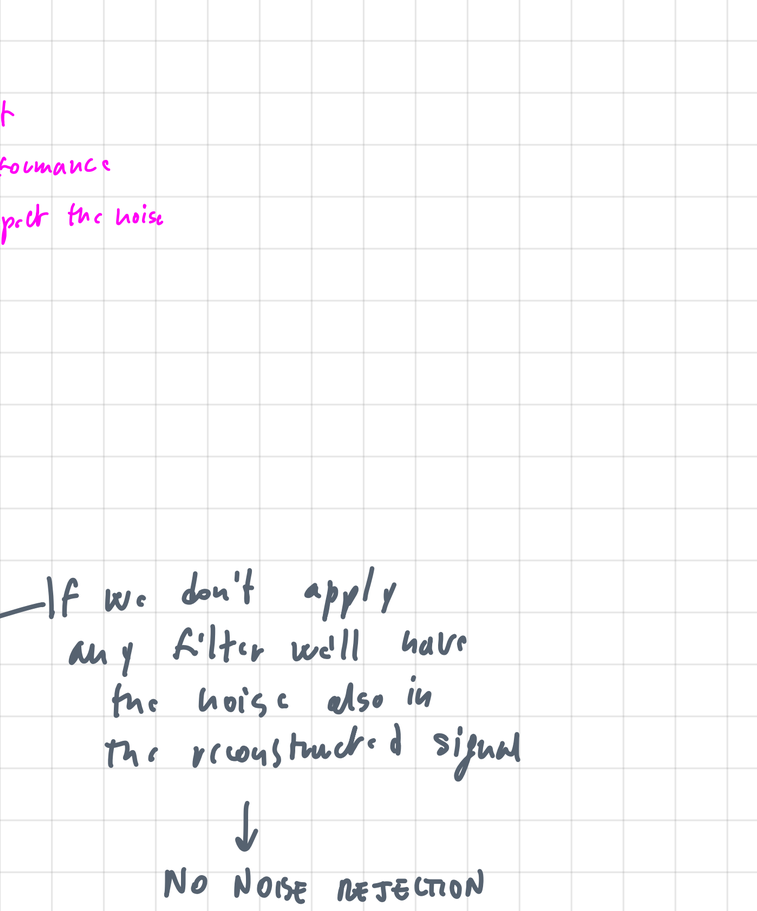
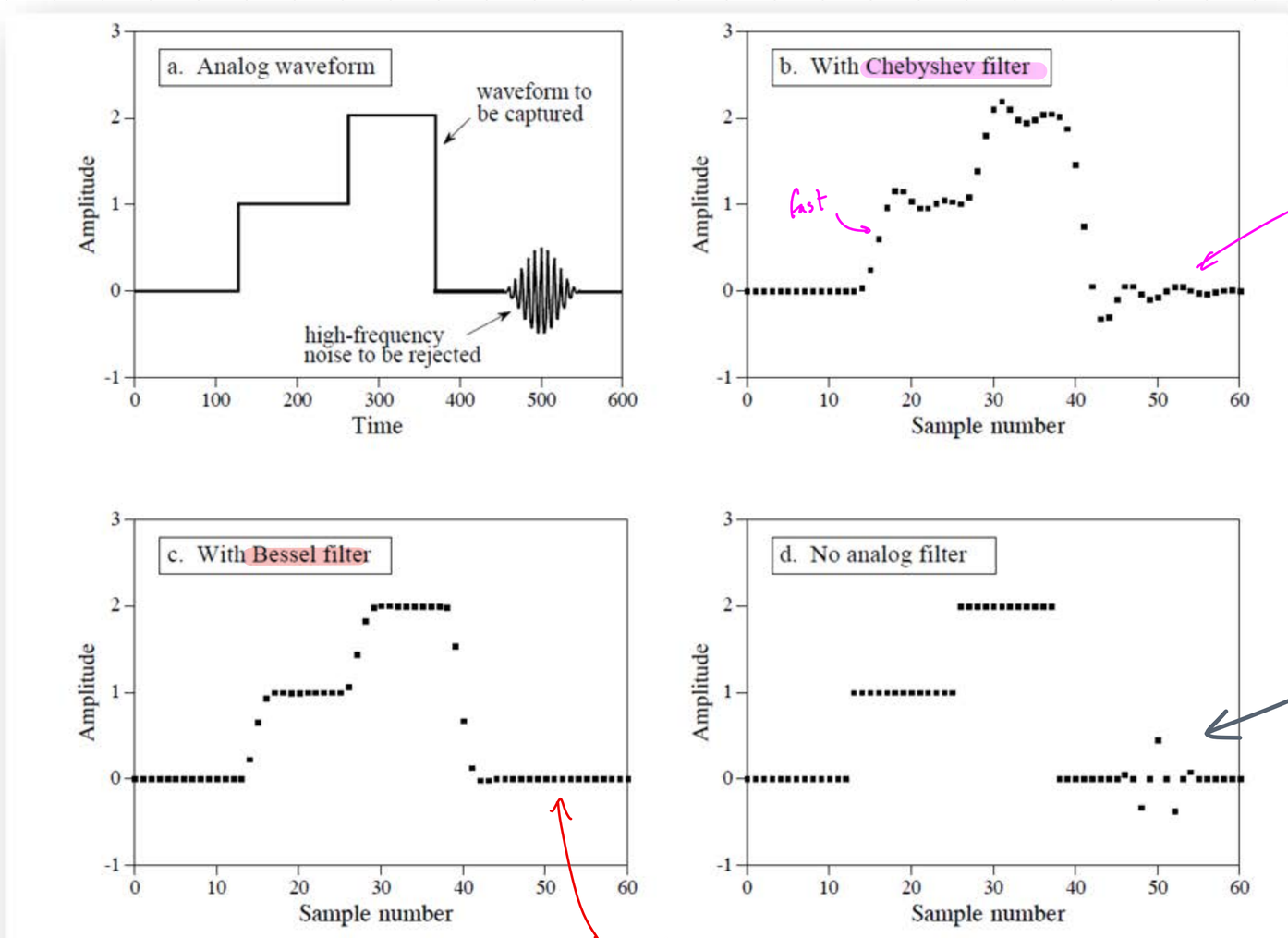


FIGURE 3-14
Pulse response of the Bessel and Chebyshev filters. A key property of the Bessel filter is that the rising and falling edges in the filter's output looking similar. In the jargon of the field, this is called *linear phase*. Figure (b) shows the result of passing the pulse waveform in (a) through a 4 pole Bessel filter. Both edges are smoothed in a similar manner. Figure (c) shows the result of passing (a) through a 4 pole Chebyshev filter. The left edge overshoots on the *top*, while the right edge overshoots on the *bottom*. Many applications cannot tolerate this distortion.

Anti-aliasing filtering



completely remove the high freq. noise

FIGURE 3-15
Three antialias filter options for time domain encoded signals. The goal is to eliminate high frequencies (that will alias during sampling), while simultaneously retaining edge sharpness (that carries information). Figure (a) shows an example analog signal containing both sharp edges and a high frequency noise burst. Figure (b) shows the digitized signal using a *Chebyshev filter*. While the high frequencies have been effectively removed, the edges have been grossly distorted. This is usually a terrible solution. The *Bessel filter*, shown in (c), provides a gentle edge smoothing while removing the high frequencies. Figure (d) shows the digitized signal using *no antialias filter*. In this case, the edges have retained perfect sharpness; however, the high frequency burst has aliased into several meaningless samples.

a little bit worse performance with respect the noise

If we don't apply any filter we'll have the noise also in the reconstructed signal

↓

NO NOISE REJECTION

FFT

(Book p. 459)

To calculate the spectrum of a sampled sequence, you must use the Discrete Fourier Transform (DFT) or its improved version named the Fast Fourier Transform (FFT). The calculation of the FFT is based on the assumption that the sequence is regular; therefore, the FFT spectrum will provide the spectrum of the sequence obtained from the periodic original sequence, i.e. the sequence is repeated infinitely many times and is always the same, as shown in Fig. 6.20.

Given N samples in the time domain, the calculation of the FFT will provide as many samples as in the frequency domain according to an algorithm that is described in the chapter about digital signal processing and digital signal processor (DSP). As in the time domain, it is assumed that the original sequence is periodically repeated every $N \cdot T_S$, also the spectrum obtained from the FFT will be repeated every $f_S = 1/T_S$ (Fig. 6.20).

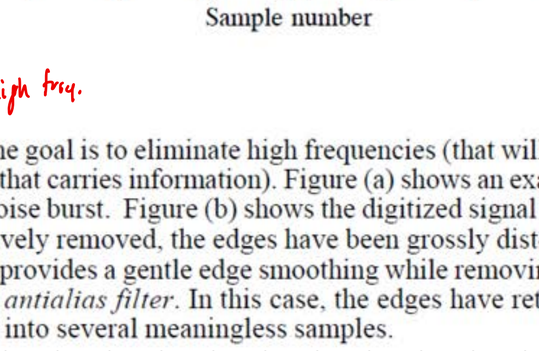
Because of this periodicity of the original sequence, implicit in the spectrum calculation with the FFT, the choice of the samples number N must not be underestimated. A different number of samples N can determine the imprecision in the periodic sequence and, ultimately, deformation in the computed spectrum. Fig. 6.21 shows a sinusoidal signal at 1kHz sampled at f_S=20kHz and its FFT for a number of N=20 samples that follow one another every $1/f_S$=50μs. Notice that the obtained spectrum is made of a perfect Dirac's delta function at 1kHz and another Dirac's delta function symmetrically placed with respect to $f_s/2$, i.e. at 20kHz-1kHz=19kHz. If we had used N=21 samples of the sinusoid, although everything else was unchanged, we would have had a very different FFT spectrum, as shown in Fig. 6.22. Although it may seem to have correctly sampled the input signal in both cases, the two sequences computed by the FFT algorithm are very different, as can be seen in Fig. 6.23. The great junction between the beginning and the end of the sequence causes the spectral deformation and a different representation in the frequency domain. Notice that the most intense histogram in Fig. 6.22 is at f_S/N=20kHz/21=953Hz, not at 1kHz (and 19kHz) any more.

Fig. 6.20

The FFT algorithm trusts on the periodicity of the input sequence
... hence FFT spectrum is for the periodic sequence and not the original one!

Fig. 6.23

Due to uneven junction of the input sequence (as assumed by the FFT algorithm)



Therefore, let become independent of the junction... → windowing →

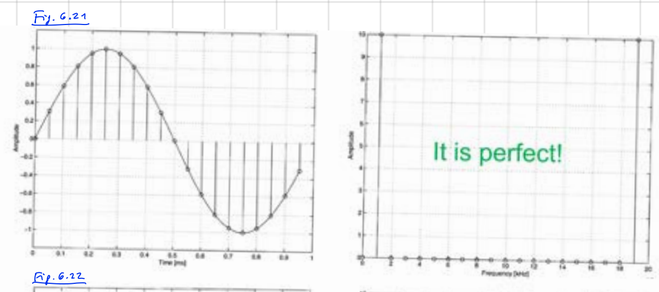
If T=Tmeasurement ↑ N ↑

To increase the frequency resolution (i.e. the number of histograms in the FFT), it is possible to increase the number N of collected data. Fig. 6.24 and Fig. 6.25 show the processing of the previous sinusoid, which was performed with a higher number of samples, with N=200 and N=256, respectively, corresponding to the total measurement duration of $N \cdot T_S$=10ms and 12.8ms, respectively, with the same sampling frequency.

We want to increase N ↑ to increase the resolution but we don't want this to happen

Correct choice:

Fig. 6.21



Wrong (unlucky) choice:

Fig. 6.22

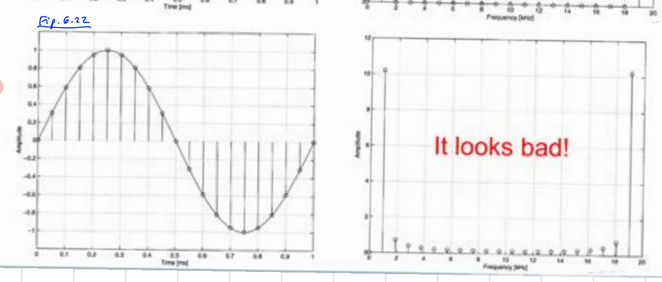


Fig. 6.24

N=200:

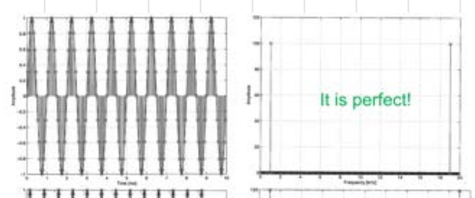
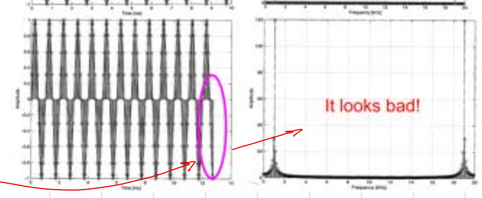


Fig. 6.25

N=256:



Windowing before FFT

The spectrum is described more finely with N frequency histograms and consequent bin-width of f_S/N=20kHz/200=100Hz (while in the previous figures, it was 1kHz). However, while the first spectrum is theoretically expected, the second one is different. The reason is that with N=200 samples, we have a perfect connection between the beginning and the ending of the sequence while with N=256 samples (or a number different from an integer multiple of f_S/f_{in}), the periodic sequence has an abrupt connection (see Fig. 6.25).

Notice that these truncations are non-predictable when we have an unknown input signal. To alleviate this problem, it is extremely useful to smooth the values of the sequence at the ending of the interval before applying the FFT algorithm. To do this, we can use the windowing technique. This operation consist of "smoothing out" the incoming samples, considering a weight becoming less and less as you approach the extremes of the interval, as shown in Fig. 6.26.

The windowing technique alleviates the problem of the connection and therefore improves the accuracy of the spectrum computed by the FFT. Unfortunately, it slightly distorts the spectrum of the original sequence, as evident in Fig. 6.27. In fact, multiplying the original sequence with a window in the time domain corresponds to convolving the original spectrum with the Fourier transform of the window in the frequency domain.

In literature, different windowing solutions have been proposed. The simplest, in addition to the rectangular one, is the Bartlett (triangular), which, for 2M+1 samples, is:

$$\text{window}(n) = 1 - \frac{|n|}{M} \quad \text{if } -M \le n \le M \quad \text{otherwise } 0$$

In the case of N samples, other windows (Fig. 6.27) are the Hanning (raised cosine):

$$\text{window}(n) = \frac{1}{2}\cdot\left(1+\cos\frac{2\pi\cdot n}{N}\right) \quad \text{if } -\frac{N-1}{2} \le n \le \frac{N-1}{2} \quad \text{otherwise } 0$$

the Hamming:

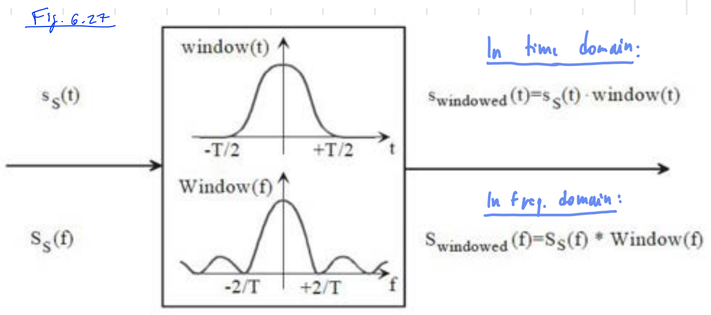
$$\text{window}(n) = 0.54 + 0.46\cdot\cos\frac{2\pi\cdot n}{N} \quad \text{if } -\frac{N-1}{2} \le n \le \frac{N-1}{2} \quad \text{otherwise } 0$$

and the Blackman:

$$\text{window}(n) = 0.42 + 0.5\cdot\cos\frac{2\pi\cdot n}{N} + 0.08\cdot\cos\frac{4\pi\cdot n}{N} \quad \text{if } -\frac{N-1}{2} \le n \le \frac{N-1}{2} \quad \text{otherwise } 0$$

The spectral trends can be found in textbooks.

Fig. 6.27



In time domain:

$s_{windowed}(t) = s_S(t) \cdot \text{window}(t)$

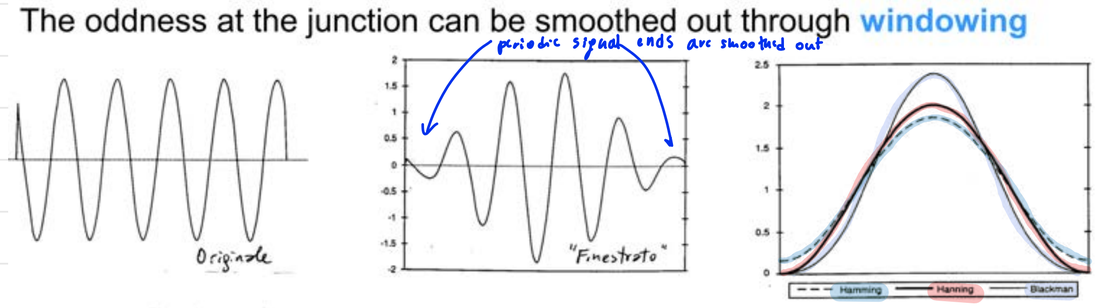
In freq. domain:

$S_{windowed}(f) = S_S(f) * \text{Window}(f)$

Fig. 6.27: Windowing effect on original sequence samples.

Fig. 6.26

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: $\text{window}(n) = 1 - \frac{|n|}{M}$ for $-M \le n \le M$ and 0 elsewhere

Hanning (rised cosin): $\text{window}(n) = \frac{1}{2}\cdot\left(1+\cos\frac{2\pi\cdot n}{N}\right)$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

Hamming: $\text{window}(n) = 0.54 + 0.46\cdot\cos\frac{2\pi\cdot n}{N}$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

Blackman: $\text{window}(n) = 0.42 + 0.5\cdot\cos\frac{2\pi\cdot n}{N} + 0.08\cdot\cos\frac{4\pi\cdot n}{N}$ for $-\frac{N-1}{2} \le n \le \frac{N-1}{2}$ and 0 elsewhere

Ex.

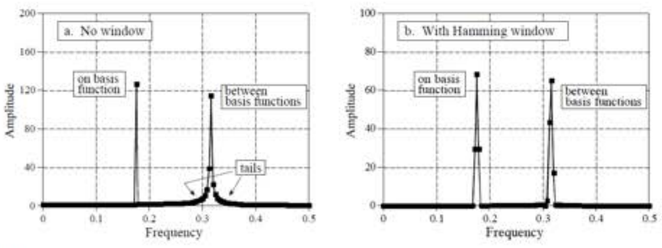


FIGURE 9-4
Example of using a window in spectral analysis. Figure (a) shows the frequency spectrum (magnitude only) of a signal consisting of two sine waves. One sine wave has a frequency exactly equal to a basis function, allowing it to be represented by a single sample. The other sine wave has a frequency *between* two of the basis functions, resulting in *tails* on the peak. Figure (b) shows the frequency spectrum of the same signal, but with a Blackman window applied before taking the DFT. The window makes the peaks look the same and reduces the tails, but broadens the peaks.

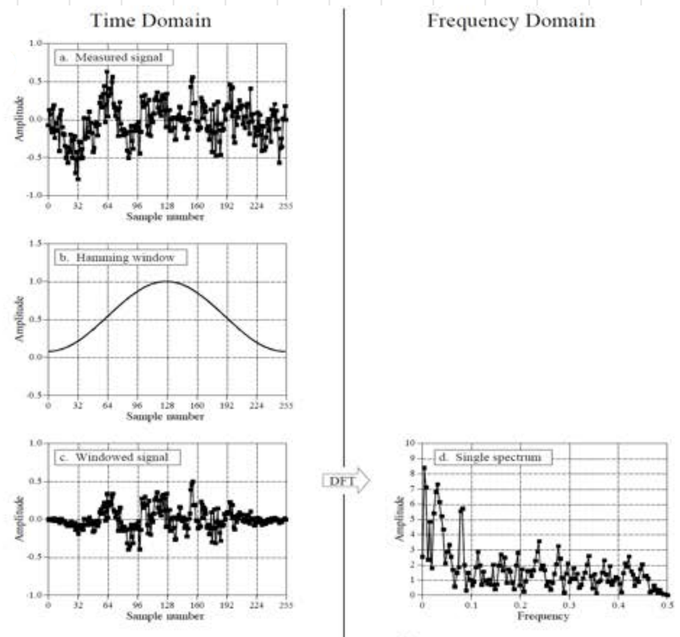


FIGURE 9-1
An example of spectral analysis. Figure (a) shows 256 samples taken from a (simulated) undersea microphone at a rate of 160 samples per second. This signal is multiplied by the Hamming window shown in (b), resulting in the windowed signal in (c). The frequency spectrum of the windowed signal is found using the DFT, and is displayed in (d) (magnitude only). Averaging 100 of these spectra reduces the random noise, resulting in the averaged frequency spectrum shown in (e).

Sample and Hold

(Book p. 463)

The *Sample & Hold* (S&H) is an analog circuit which, in correspondence of a control signal, *samples* the input voltage value and stores it until the next command. Such circuits are normally used to provide a sampled signal to an analog-to-digital converter that translates the amplitude into the relative binary code. Generally, the S&H is used wherever storing the instantaneous value of an analog signal is needed before its processing.

The operation of an S&H is divided into two steps: the first one is the actual *sampling* operation while the second one consists in the voltage maintenance (*holding*). As shown in Fig. 6.28, the first operation consists of charging the capacitance C_H to the input voltage to have $V_{out}=V_{in}$. In the second phase, the switch is open, and the capacitor C_H remains insulated to store the analog value. The switch is usually constructed by a MOSFET while the OpAmp is used as a buffer to provide the output current without changing the information stored in the capacitance.

The need of an S&H with an ADC, and sometimes a DAC, is correlated to the continuous variation in the input signal. In fact, to convert an analog signal into a digital code, we use a certain conversion time T_{conv}, during which it is very important that the signal applied to the input of the ADC does not vary significantly; otherwise, an error in the conversion occurs. Consider, for example, the conversion of a sinusoid $s(t)=V_p\cdot\sin(2\pi\cdot f\cdot t)$ with a peak amplitude V_p. The maximum variation speed is at the inflection point of the sinusoid with a slope equal to: $\left.\frac{dV}{dt}\right|_{max}=2\pi\cdot f_{max}\cdot V_{p\,max}$

As we can see in the following, in the case of an ADC with n=12bits and a maximum excursion V_{FSR}=5V (Full Scale Range), the value of the smallest part of the output signal (resolution) is:

$$LSB=\frac{V_{FSR}}{2^n}$$

equal to 1.22mV. To avoid the ADC error during the conversion, we can tolerate a maximum variation less than $\Delta V<\frac{1}{2}LSB=610\mu V$. Assuming to apply as an input a sinusoid with a peak-to-peak amplitude equal to the whole range of the ADC ($V_{pmax}=\frac{1}{2}FSR$) and to use a conversion time of $T_{conv}=10\mu s$, the maximum frequency allowed for the input signal is:

$$f_{max}=\frac{\Delta V}{T_{conv}}\cdot\frac{1}{2\pi\cdot V_{p\,max}}=3.9Hz$$

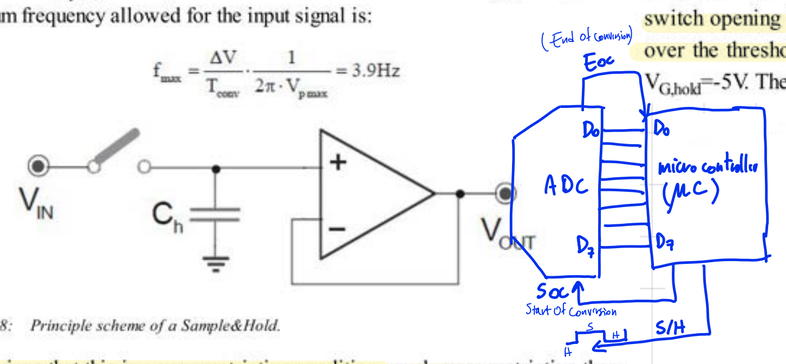


Fig. 6.28: Principle scheme of a Sample&Hold.

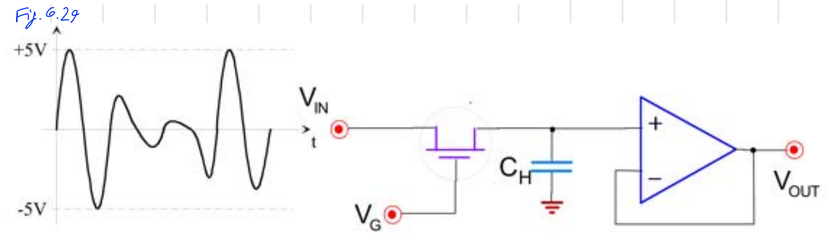
It is obvious that this is a very restrictive condition, much more restrictive than the theoretical limit imposed by the Sampling Theorem of $f_{max}<\frac{1}{2}f_S=(2\cdot T_{conv})^{-1}=50kHz$.

Connecting an S&H before an ADC, the restrictions on the input signal's frequencies become relaxed. In fact, the signal conversion takes place during the hold phase when the switch is open, and the ADC input does not vary. At least ideally, there is no error during the conversion, and the maximum frequency of the input signal can reach the theoretical limit $f_{max}=\frac{1}{2}f_S$.

Project the S&H shown in Fig. 6.29, using devices with the following characteristics: a MOSFET with V_t=2V, R_{on}=50Ω, C_{gs}=0.5pF, and C_{ds}=0.1pF; an OpAmp with A_O=110dB and I_{bias}=50pA. The input signal is variable between -5V and +5V, whose bandwidth is 20kHz. We require that the maximum be less than 0.01% of the maximum input dynamic (Full Scale Range, FSR), i.e. less than 1mV. The sampling frequency is 100kHz.

To ensure (during the sampling phase) the switch closing at all times, the control voltage must exceed $V_{in,max}+V_t=+7V$. To ensure that the MOSFET conduction resistance is small enough, we choose to work with at least 3V of overdrive, i.e. we choose $V_{G,sampling}=+10V$. In the same way, to ensure the switch opening during the hold phase, it is needed that V_G does not exceed V_{in} over the threshold voltage, i.e. $V_{G,hold}<V_{in,min}+V_t=-3$, we choose, for example, $V_{G,hold}=-5V$. The control voltage swing will be 15V.

Fig. 6.29

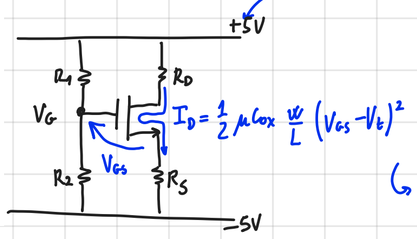


Specs:

Input signal:	-5V÷+5V 20kHz bandwidth
maximum admitted error:	< 0.01% FSR = 1mV
sampling frequency:	100kHz

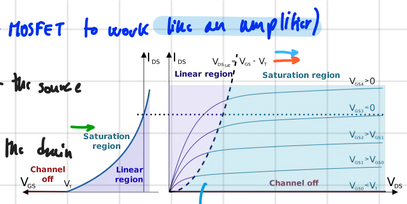
Components: OpAmp: A_O=110dB I_{bias}=50pA

MOSFET: V_t=2V $R_{on_{min}}$=50Ω C_{gs}=0.5pF C_{ds}=0.1pF



Operating conditions (for the MOSFET to work like an amplifier)

- $V_{GS}>V_t$ → channel at the source
- $V_{GD}<V_t$ → No channel at the drain
- ↳ $V_{DS}>V_{OD}=V_{GS}-V_t$

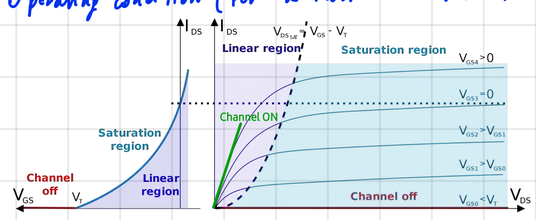


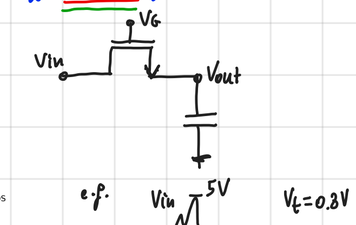
$V_D-V_S>V_G-V_S-V_t$

↳ $V_{OD}<V_t$

Note on MOSFET:

Operating condition (for a MOSFET to work like a SWITCH)





→ ON: $R_{ON}=\frac{dV_{DS}}{dI_D}$ low, $V_{DS}\simeq 0$, $V_G>V_{in}+V_t$

→ OFF: $V_{GS}<V_t$, $V_G<V_{in\,min}+V_t$

e.g. Vin 5V V_t=0.3V −5V

- $V_{GON}\geq 5+0.3$
- $V_{GOFF}\leq -5+0.3$

Charge injection

This first problem is attributable to the capacitive coupling between the control of the MOSFET and the capacitance C_H during the transition from sampling to hold, as shown in Fig. 6.30. Following the switch opening, there is an undesirable injection of charge into C_H and a consequent change in the potential, equal to:

$$V_{injection}=\Delta V_G\cdot\frac{C_{gd}}{C_{gd}+C_H}$$

The value ΔV_G represents the amplitude of the transition since the capacitor left floating (switch open). To limit the introduced error, we need to choose C_H high enough, unfortunately, penalizing the input bandwidth during the sampling phase, depending on the time constant $\tau=R_{ON}C_H$. For example, to ensure an error $V_{injection}\leq 0.01\%\cdot FSR=1mV$ and supposing $\Delta V_G=V_{G,sampling}-V_{G,hold}=15V$ (the worst case), we will choose $C_H\geq 9nF$, determining the bandwidth as 354kHz.

If this error is constant, it behaves as the OpAmp offset (since this error has a behavior similar to an offset, we decided to call it *induced offset* or *pedestal error*) and will be eliminated by the subsequent electronics or at the software level after the ADC.

Aperture-induced non-linearity

Actually, as shown in Fig. 6.30, not all of the V_G transition constitutes ΔV_G coupled on the capacitance because while the MOSFET conduces ($V_G>V_{in}+V_t$), the input generator V_{in} imposes the C_H voltage. Because this value depends on the input voltage, the error is not constant and then determines a non-linearity of the transfer function of the S&H (for this reason, we gave the name *aperture-induced non-linearity* to this error), which is more difficult to eliminate as in the case of a simple systematic offset.

In this case, the worst condition is verified with a high input voltage. In fact, with V_{in} equal to 5V, the MOSFET will be turned off at $V_G=5+V_t=7V$ and then $\Delta V_G=12V$ (from which $V_{injection}=0.83mV$). The best condition is when $V_{in}=-5V$ because the MOS is turned off at $V_G=-3V$ and thus $\Delta V_G=2V$ (corresponding at $V_{injection}=0.28mV$).

Signal-feedthrough

The MOS capacitance C_{ds} is the cause of another unwanted coupling (Fig. 6.31) between the input and the capacitance C_H, which in the hold phase, determines a perturbation of the stored voltage due to the input transition. The value of the error is:

With a capacitor of 9nF and in the worst condition, i.e. the input fluctuation equal to the whole FSR (ΔV_{in}=10V), the error is only 111μV peak-peak, equal to a tenth of the admitted error, thanks to the favorable ratio between C_{ds} and C_H. Choosing a smaller C_H subjects this situation to a deterioration.

Droop

We need to consider also the C_H discharge during the hold phase due to the OpAmp bias current and the MOS leakage current. With a leakage current of 100pA and capacitance of 9nF, the voltage varies 11mV/s. This means that working with a sampling frequency of 100kHz, the variation is 0.11μV during each hold phase. The error, negligible due to our specifications, can become meaningful when working with longer hold phases (low sampling frequency unless you lengthen the tracking phase) and is subjected to strong temperature dependence because the inversely biased junction's leakage doubles every 5°C.

Fig. 6.30

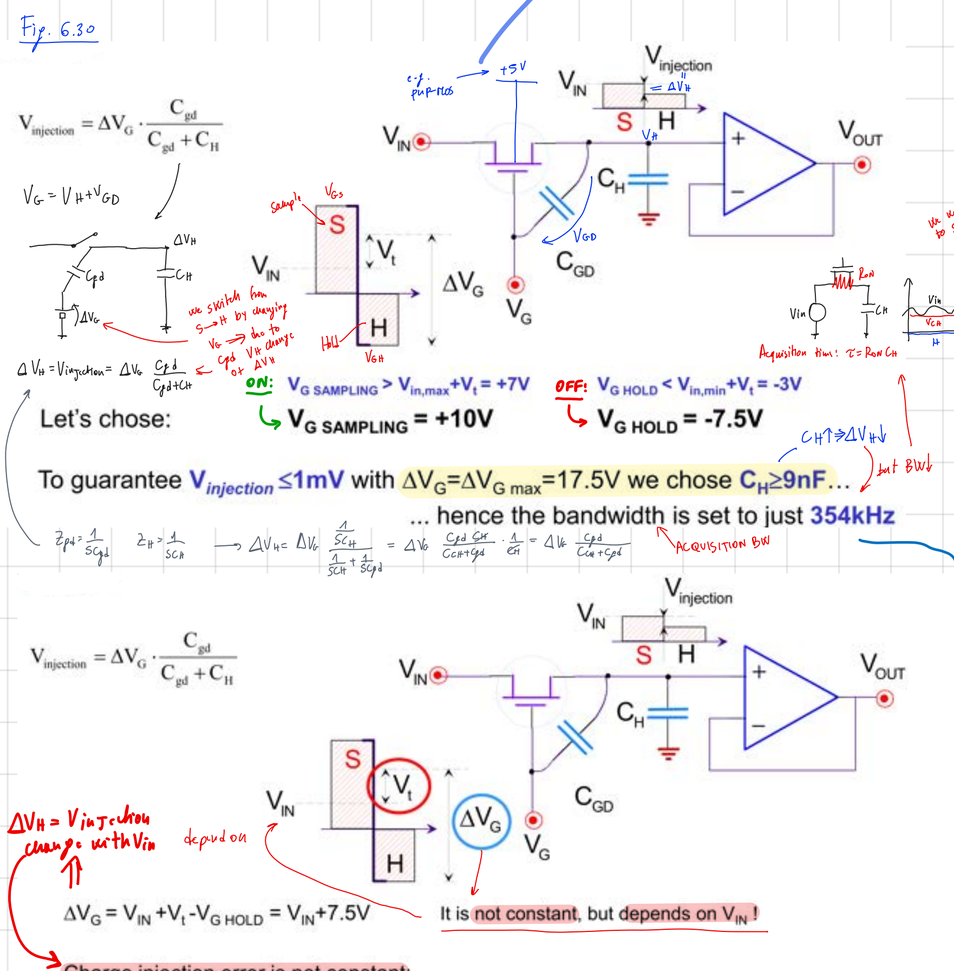
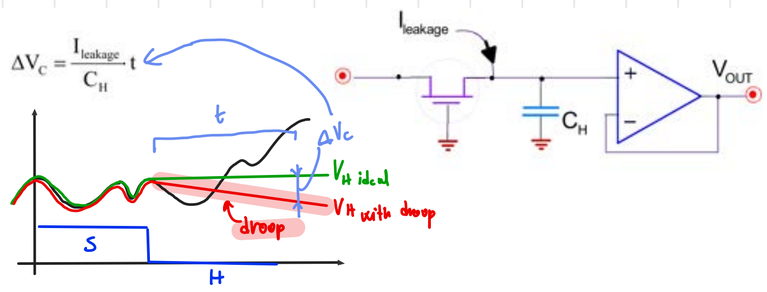


Fig. 6.31

With $\Delta V_{IN\,max}$ = 10V we get $V_{injection}$ = 111μV

↳ C_H sufficiently large

… negligible (in this case) compared to the requirements of <1mV

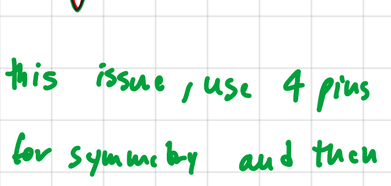


With $I_{leakage}$=100pA and C_H=9nF, the stored voltage will droop by 11mV/s

Hence with f_S=100kHz we get ΔV_C=0.11μV every 10μs Hold duration

If the MOS transistor is fully symmetric is useless to specify which is the source and which is the drain:

↳ But the thing is that with just the 3 pins S,G,D the transistor won't be symm.



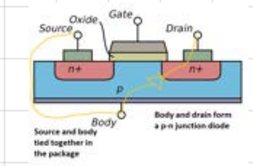
connecting bulk-source together create symmetry

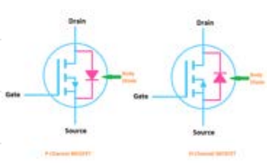
↓

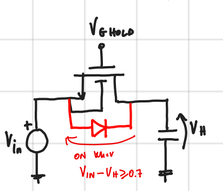
But introduces a parasitic diode/inverse bias

(no problem when using MOS as an amplifier because $V_D>V_S$ → DIODE OFF)

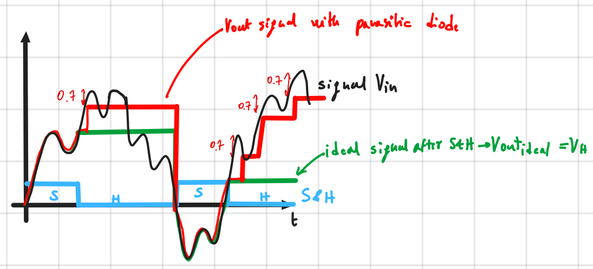
↳ Problematic when using MOS as SWITCH in a S&H



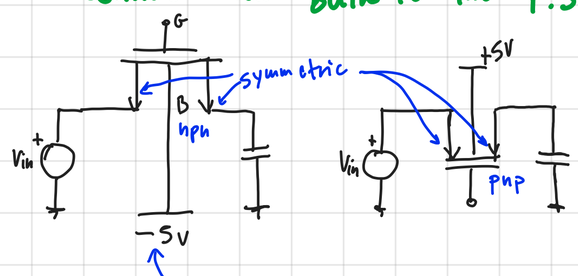




when $V_{in}>V_H$ such that $V_{in}-V_H>V_{t\,diode}$



↳ to solve this issue, use 4 pins transistors for symmetry and then connect the bulk to the P.S.

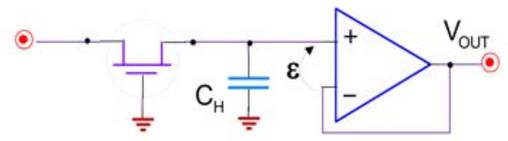


in such way that we reverse bias both the diode forming b/w B and S and the one b/w B and D

↳ so parasitic diodes OFF

Buffer-induced non linearity

Because of the finite OpAmp gain, the output voltage will always have an inaccuracy in respect of the stored voltage on the capacitance because of the residual differential voltage between the two terminals. To ensure an error lower than the 10% of the admitted one, i.e. $\varepsilon<100\mu V$, it is necessary to choose an OpAmp with at least $A_0>V_{out,max}/\varepsilon=50000=94dB$. Also this error, depending on the output voltage, introduced a non-linearity effect.



ε → due to the fact that A_0 is not ∞ but limited in reality

To guarantee $\varepsilon<1mV$ we must have $A_0 > V_{out,max}/\varepsilon = 5{,}000 = 74dB$

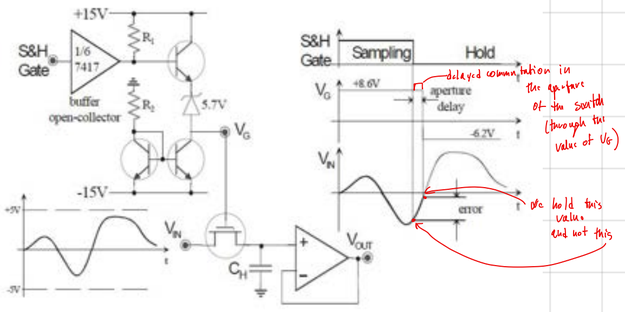
Aperture delay-time → dynamic error

Other inaccuracies in the sampled signal are caused by non-ideal timing control signal. A first effect is caused by the delay with which, in the transition from sampling to hold, the opening command can actually halt the switch conduction. In this regard, note that although the control sampling command has often CMOS logic levels (for example, it comes from a µC or DSP), the voltages applied to the MOS gate are very different and much wider. A stage which makes the appropriate level-shifting is therefore necessary.

Consider, for example, the driver shown in Fig. 6.32. One of the six buffers (a non-inverting buffer) contained in the IC 7417; the buffer has an open-collector output. The tail *npn* transistor impose the Zener bias current; the upper transistor determines the diode cathode voltage and, thus, on the MOS gate.

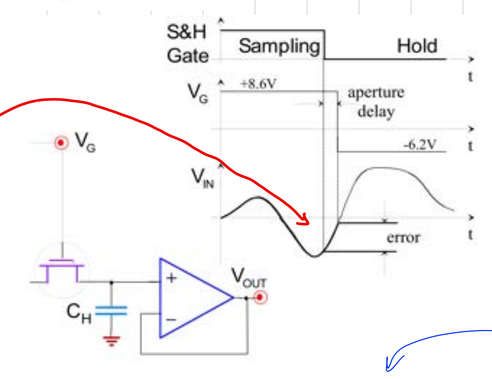
When the buffer input is low, its output has a voltage equal to +0.2V and then $V_G\cong0.2\text{-}0.7\text{-}5.7=-6.2V$ (hold phase). Indeed, with a high input, the buffer open-collector output is interdict, and the resistance R is free to increase the base voltage up to the supply voltage; the output goes around $V_G\cong15\text{-}0.7\text{-}5.7=+8.6V$ (sampling phase). Both levels are suitable to drive the S&H.

Fig. 6.32



$$\Delta V_{aperture}=\left.\frac{dV_{in}}{dt}\right|_{max}\cdot T_{aperture}=2\pi\cdot f_{max}\cdot V_{in,max}\cdot T_{aperture}$$

highest error when Vin switches fast → $\left.\frac{dVin}{dt}\right|_{max}$ ⇒ max error



With $T_{aperture}$=1ns we get $\Delta V_{aperture}=2\pi\cdot20kHz\cdot5V\cdot1ns=0.63\,mV$

The propagation of control command along the circuit will be in a finite time, in the order of several nanoseconds. The resulting effect is that the actual switch opening will happen with a certain average delay compared to the sampling command assertion. This corresponds to sampling the signal at a different time than expected. In the case of free-running sampling (continuous acquisition with constant cadence), the effect is merely to have a sampling comb delayed compared to the theoretical one, but is always uniform.

Indeed, in the case of single-shot sampling (irregular acquisition with unpredictable results), it is extremely important to have a sampling exactly at the desired moment. Therefore, the aperture delay $T_{aperture}$ determines an acquisition error $\Delta V_{aperture}$, depending on the slope of the input voltage V_{in} at the desired sampling moment (Fig. 6.32).

$V_{in}=V_{in,max}\sin(2\pi f_{max}t)$

Assuming to have a sinusoid with a peak-to-peak amplitude $V_{in,max}$ with the maximum frequency $f_{in,max}$, we can extract the maximum amplitude error:

$$\Delta V_{aperture}=\left.\frac{dV_{in}}{dt}\right|_{max}\cdot T_{aperture}=\overbrace{2\pi\cdot f_{max}\cdot V_{in,max}}^{V_{in,max}\cdot2\pi\cdot f_{max}\cdot\overbrace{\cos(2\pi f_{max}t)|_{max}}^{=1}}\cdot T_{aperture}$$

For example, with an average delay $T_{aperture}$=1ns, we have $\Delta V_{aperture}=2\pi\cdot20kHz\cdot5V\cdot1ns=628\mu V$, close to the limit of the accuracy imposed by the specifications.

Aperture time jitter → dynamic error

Actually, the aperture delay is not a deterministic value because of the electronic noise, of the supply voltage fluctuations, of the thermal drift, of the components' tolerance degrees and so on. Therefore, in addition to the average value described previously, we need to consider also its fluctuation, the *time jitter*.

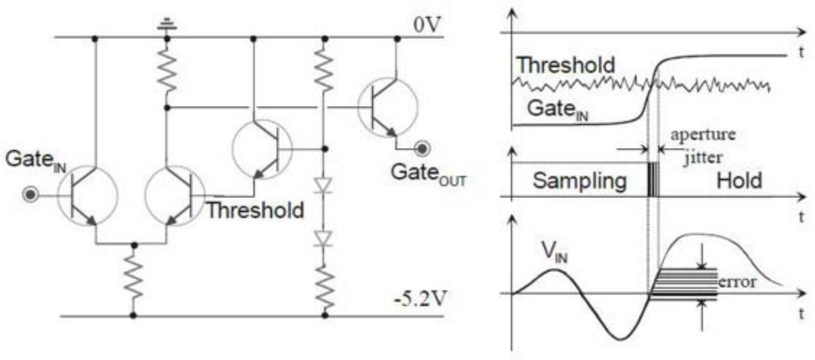
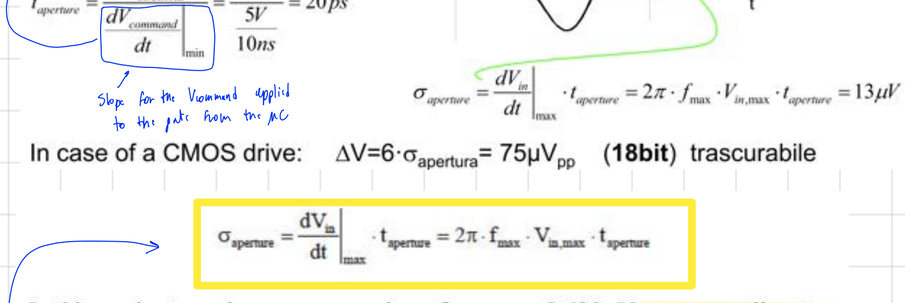


Fig. 6.33: Example of a differential stage for the sampling command.

For example, consider the comparator shown in Fig. 6.33, which is made of a differential stage, and consider the noise superimposed on the switching threshold equal to $\sigma_{threshold}$=1mV. If the applied input command has a swing of 0.4V in 400ps, the corresponding time jitter for the threshold crossing will be equal to:

$$t_{aperture}=\frac{\sigma_{threshold}}{\left.\frac{dV_{ECL}}{dt}\right|_{min}}=\frac{1mV}{\frac{0.4V}{0.4ns}}=1ps$$

As it has been obtained, this jitter is expressed as *rms*. Similar to what has been done previously; the resulting amplitude error is given by:



In case of a CMOS drive: $\Delta V=6\cdot\sigma_{apertura}=75\mu V_{pp}$ **(18bit)** trascurabile

$$\sigma_{aperture}=\left.\frac{dV_{in}}{dt}\right|_{max}\cdot t_{aperture}=2\pi\cdot f_{max}\cdot V_{in,max}\cdot t_{aperture}$$

In this project, we have an *rms* value of $\sigma_{aperture}$=0.628µV, corresponding to a statistical error with a ranging of about 3.8µV peak-to-peak, which is non-negligible in this case.

Acquisition time → dynamic error

Consider, however, the sampling phase when the switch conduces, and the S&H must track the input signal. It is important to pay attention to the time required for the S&H to make an accurate acquisition of the input signal so that V_{out} reaches the value of the input voltage within the range given by the specifications.

The main limit to the charging of C_H is given by the circuitry before the MOSFET (Fig. 6.34); this circuitry must provide the required (charge) current. For example, a possible OpAmp before the S&H or the same switch could have a finite $I_{O,max}$. In both cases, indicating the maximum current available for charging with I_{max}, the charging speed of the capacitor, named the *slew-rate* SR, could reach the value:

$$SR=\frac{I_{max}}{C_H}=\frac{25mA}{9nF}=2.8\,V/\mu s$$

Indeed, without a current limit, an increasing Δ of the voltage on the capacitance would be with the typical exponential law, with the time constant $\tau=R_{ON}\cdot C_H$, i.e. with an initial maximum speed equal to:

$$\left.\frac{dV_{C_H}}{dt}\right|_{max}=\left.\frac{d\left[\Delta\cdot\left(1-e^{-t/\tau}\right)\right]}{dt}\right|_{max}=\frac{\Delta}{\tau}$$

In the case where this slope exceeds the available SR, the S&H will remain limited by the slew-rate until the two slopes are equal, and the charge will continue exponentially, as shown in Fig. 6.35. The connection will come when $\Delta/\tau=SR$, i.e. when for the charge across the capacitor remains for only a transition of $\Delta=R_{on}\cdot I_{max}$.

In the example, the worst case is verified when the voltage across the capacitance must make the maximum excursion, going from the previously sampled value -5V to the new value +5V that must be sampled. Assuming to have an S&H with an OpAmp with $I_{O,max}$=25mA, we should have $\Delta=50\Omega\cdot25mA=1.25V$. The charging is slew-rate limited (linear ramp) for a time equal to: $t_{SR}=\frac{10-\Delta}{SR}=\frac{(10-1.25)V}{2.8V/\mu s}=3.1\mu s$

To estimate the time required to complete the charging transition, which should happen in a linear way (exponential law), we should establish the desired precision on the final voltage value quantifiable through the error:

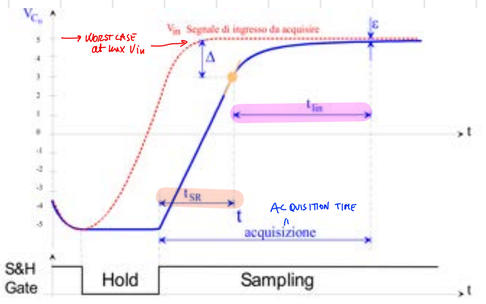
$$\varepsilon=\Delta-\Delta\cdot\left(1-e^{-\frac{t_{lin}}{\tau}}\right)$$

To obtain a maximum error less than ε, we should wait the time: $t_{lin}=\tau\cdot\ln\frac{\Delta}{\varepsilon}$

Imposing in our case ε<1mV, we obtain t_{lin}>3.2µs.

In conclusion, the total acquisition time will be $T_{acquisition}=t_{SR}+t_{lin}=$ 3.1µs+3.2µs=6.3µs. This time is limited by the maximum working frequency of the S&H and, consequently, the maximum input signal frequency.

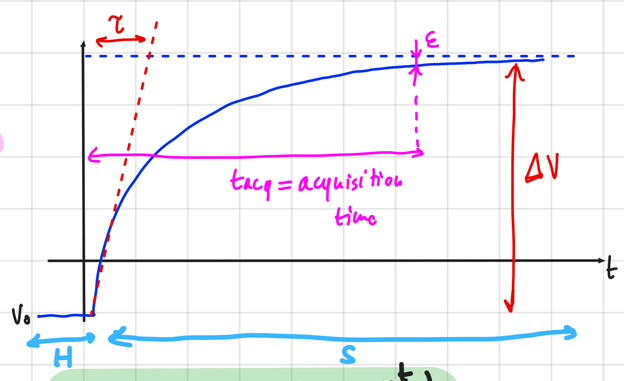
Fig. 6.35



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 $t_{lin}=\tau\cdot\ln\frac{\Delta}{\varepsilon}$ ← $t_{acquisition}$

For example: with $I_{Out.max}$ = 25mA we get Δ = 50Ω·25mA = 1.25V



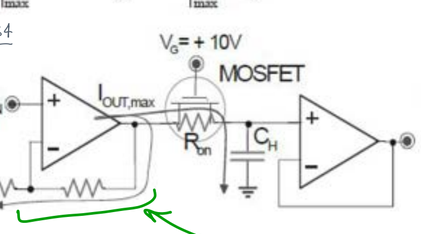
$V(t)=V_0+\Delta V\left(1-e^{-\frac{t}{\tau}}\right)$

error: $\varepsilon=\Delta V\cdot e^{-\frac{t_{acq}}{\tau}}$

acquisition time: $t_{acq}=\tau\ln\frac{\Delta V}{\varepsilon}$

$\tau=R_{ON}C_H$

Further explanations

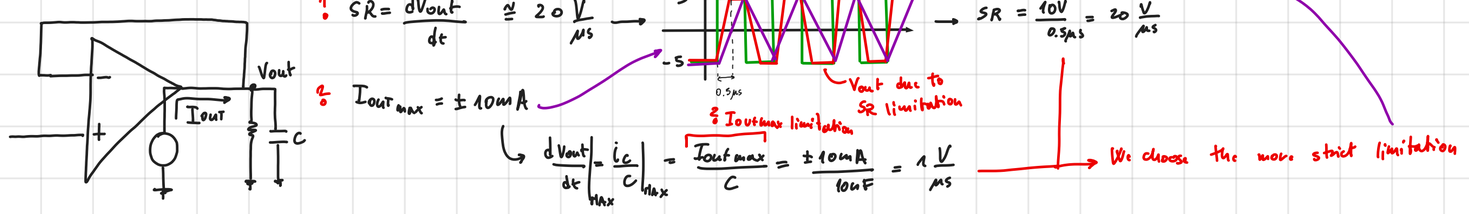


e.g.

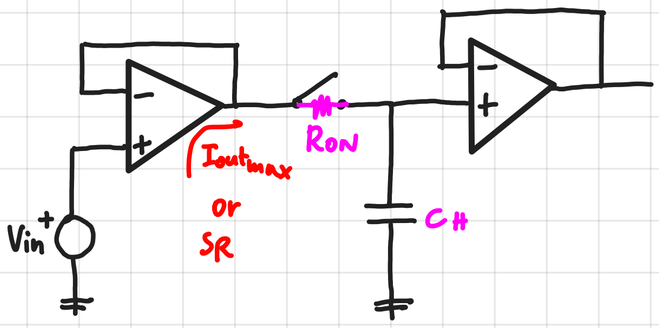
Slew rate

The S&H circuit could be driven by another stage → this can introduce another limitation in terms of Slew Rate (SR)

↳ consider an Op Amp bcs now that its 2 main limitations are:



So considering these limitations' effect on an S&H circuit:

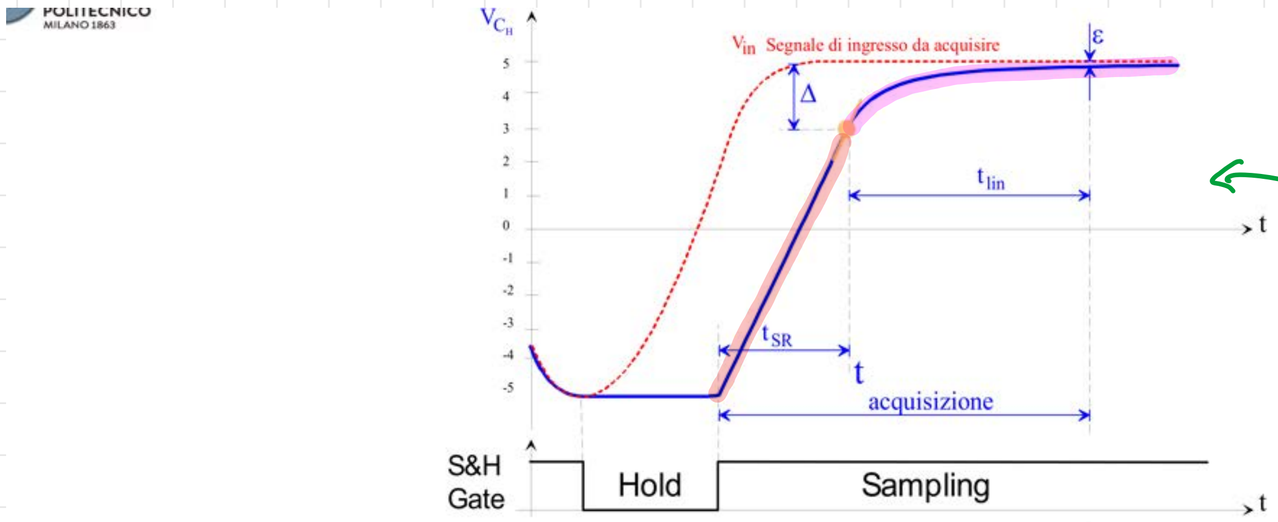


→ τ limitation

→ Ioutmax or SR limitation ($SR = 10 \frac{V}{\mu s} \overset{e.g.}{\geq} \frac{dV_{out}}{dt} \frac{I_{outmax}}{C_H} = 1 \frac{V}{\mu s} \overset{e.g.}{}$)

↑ choose the strictest/lowest limitation

↳ So the trend of V will no be just exponential but: ⟶ Indeed it cannot be just exponential:

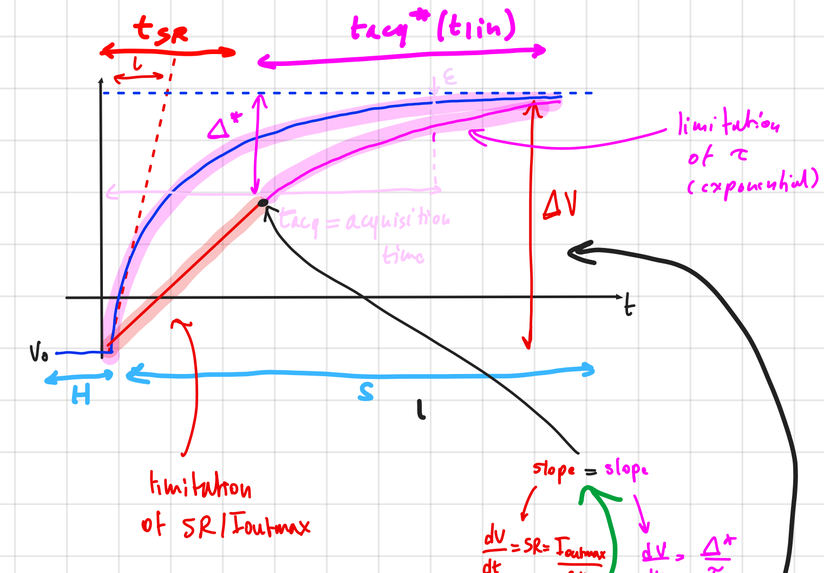


Acquisition (Sampling): linear ramp $t_{SR} = \frac{10-\Delta}{SR} = \frac{(10-1.25)V}{2.8V/\mu s} = 3.1\mu s$

and exponential charge $t_{lin} = \tau \cdot \ln \frac{\Delta}{\varepsilon}$

For example: with $I_{Out.max} = 25mA$ we get $\Delta = 50\Omega \cdot 25mA = 1.25V$

(Usually at the exam we aren't required to remember the SR limitations
↳ just exponential



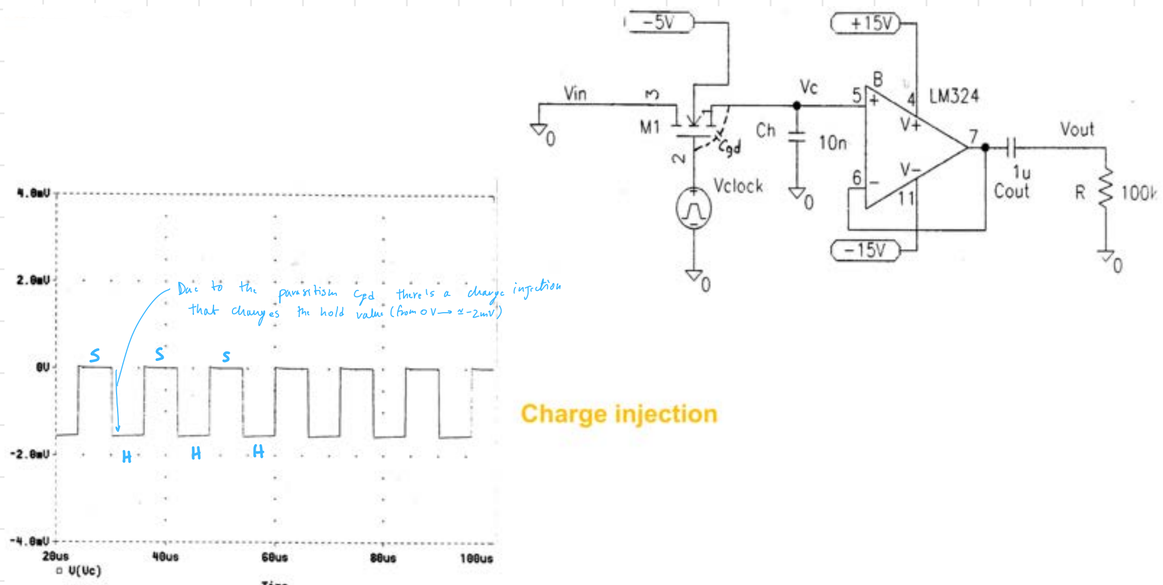
this happen when:

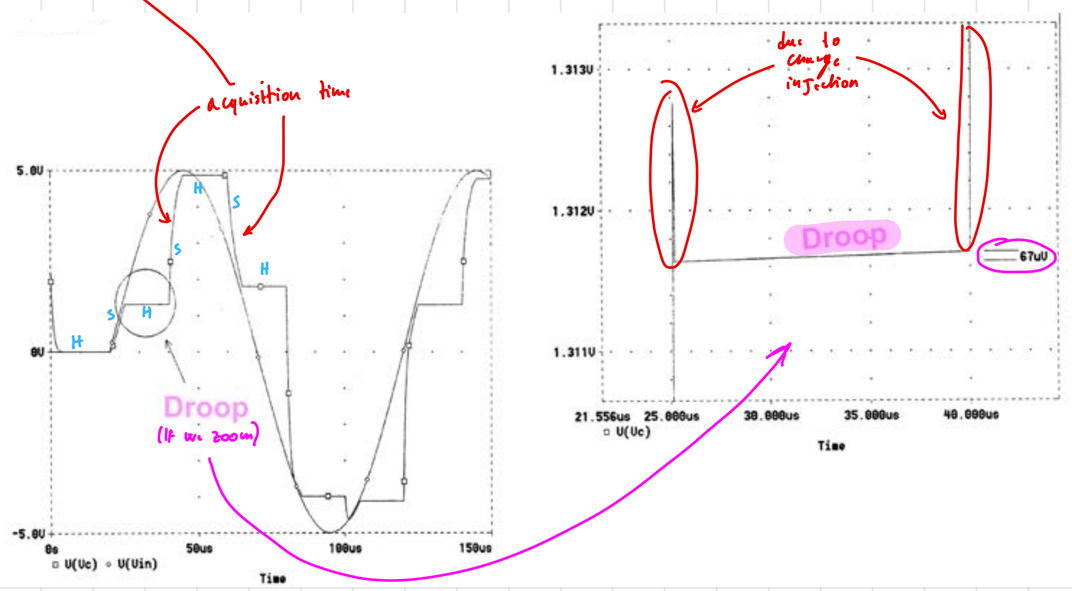
$\frac{\Delta^*}{\tau} = SR$ or $\frac{I_{outmax}}{C_H}$ (e.g. more strict)

↳ $\Delta^* = \frac{I_{outmax}}{C_H} \cdot \tau = I_{outmax} R_{on}$ ($\tau = R_{on} C_H$)

- $t_{lin} = t_{acq}^* = \tau \ln \frac{\Delta^*}{\varepsilon}$
- $t_{SR} = \frac{(\Delta - \Delta^*)}{\frac{I_{outmax}}{C_H}}$ ← or SR (depending on strictest limitation)

Examples





Improved S&H circuit

Charge injection compensation

(Book p. 476)

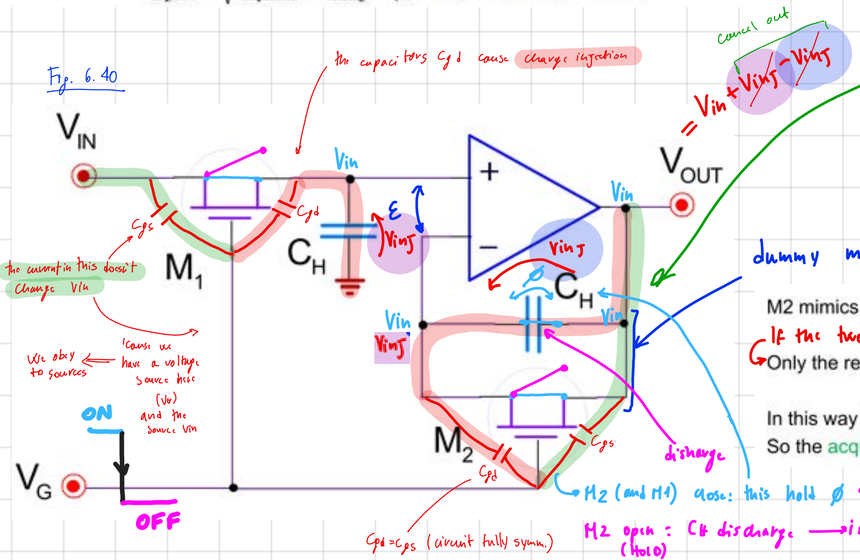
In the following will be described a more accurate and faster S&H in respect of the Fig. 6.40 prototype analyzed in other literature. The first goal is the acquisition time reduction. The more mundane method to obtain this result is that of the C_H reduction; unfortunately, in this way, also the other errors should increase consequently.

The solution shown in Fig. 6.40 employs a second capacitor on the OpAmp feedback, with the same capacitance as that of C_H. During the sampling phase, both MOSFETs are closed, and M2 discharges the feedback capacitance and re-establishes the buffer configuration. It is important to note that both transistors see the same voltages because $V_{source}=V_{drain}=V_{in}=V_+=V_{out}=V_-$. For this reason, the MOSFET opening occurs simultaneously (at the same moment and with the same transition applied between the channel and the gate) and causes an identical charge injection into both the capacitors. The total output effect is canceled because that the input voltage on the capacitor is lower than the one expected is counterbalanced by the slightly positive charging of the feedback capacitance. Only does the devices' mismatches cause a residual error of:

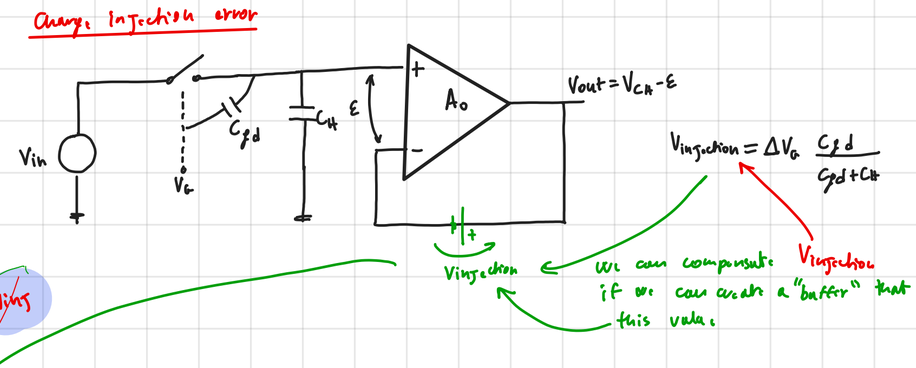
$$V_{injection} = \Delta V_G \cdot \frac{C_{gd}}{C_H} \cdot \left(\frac{\Delta C_{gd}}{C_{gd}} + \frac{\Delta C_H}{C_H} \right)$$

Also, the droop effect due to the OpAmp bias currents is canceled because, whatever their direction is, they will symmetrically charge the two capacitors, cancelling the output error. However, the proposed circuit does not reduce the signal feed-through which, on the contrary, increases because of the C_H reduction.

When the acquisition time due to the C_H charge becomes very small, the OpAmp settling-time becomes very important, which is linked both to its closed loop pole (through the GBWP) and to the intrinsic SR. Typical values are in the order of hundreds of ns. Thus, in this case, a good estimation of the total acquisition time is the sum of the RMS values of different delay contributions: $t_{acq,TOT} = \sqrt{t_{acquisition}^2 + t_{settling}^2} = \sqrt{(300ns)^2 + (300ns)^2} = 424ns$



Charge injection error



Current out = Vin + Vinj − Vinj = Vin

dummy matched pair (C_H and MOS transistor)

$V_{inj}^{I} - V_{inj}^{II} = \Delta V_G \frac{C_{gd}^{I}}{C_{gd}^{I}+C_H^{I}} - \Delta V_G \frac{C_{gd}^{II}}{C_{gd}^{II}+C_H^{II}}$

↳ $C_{gd}^{II} = C_{gd}^{I}(1+\Delta C_{gd})$

$C_H^{II} = C_H^{I}(1+\Delta C_H)$

M2 mimics M1 and eventually inject the same charge, but in oposite direction

If the two CH are not perfectly identical

Only the residual contriution due to the mismatches remains $V_{injection} = \Delta V_G \cdot \frac{C_{gd}}{C_H} \cdot \left(\frac{\Delta C_{gd}}{C_{gd}} + \frac{\Delta C_H}{C_H} \right)$

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

Simplify driving requirements

(Book p. 478)

Such an S&H is not fast enough if, downstream, it is followed by an ADC with a lower conversion time, for example T_C>5µs, so that the latter is the limiting factor of the speed and not the S&H (which weighs only for one tenth of the conversion time). Note that the proposed calculus is not true for feedback stages which are discussed in the next section.

Another very interesting architecture, which offers good performance levels not only in terms of speed, but also in terms of MOSFET command easiness, is depicted in Fig. 6.41. To understand how it works, note the first branch on the non-inverting terminal, which brings the voltage to the ground and serves as a dummy structure to balance the charge injection, as for the circuit shown in Fig. 6.40. Thus, also the inverting terminal should have a zero potential because it is a virtual ground. In the sampling phase, both MOSFETs are closed, and we obtain a simple inverting configuration with a gain $V_{out}/V_{in}=-R_F/R_F=-1$. In this situation, across the feedback C_H, there is the same output voltage, i.e. V_{in}.

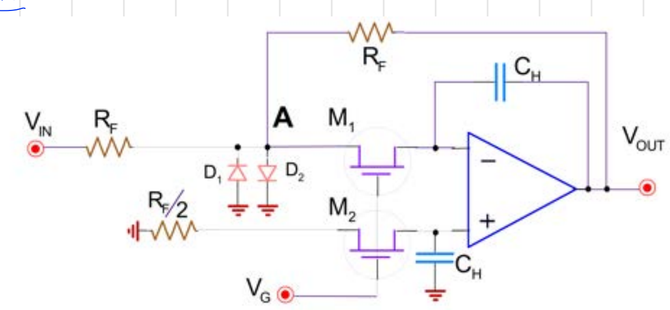
On the following hold phase, the MOSFETs are open without causing variations in V_{out} because any fluctuation in V_{in} impacts only the node A voltage, which has a potential equal to the average of the output voltage (set during the sampling phase) and the input one. However, to facilitate the conduction of M1 in the sampling phase, the node A cannot change the potential more than ±0.6V, introducing two anti-parallel diodes. ← called them X (*)

Note how the MOSFET command can assume TTL logic levels because the channel has a zero potential. The earning is dual: reduced amplitude V_G of the command signal is enough; the ΔV_G excursion during the MOSFET turning off is always constant, independently of the V_{in} value. Definitively, the error due to the charge injection in the virtual node (and thus in the hold capacitance) is significantly reduced with respect to the preceding cases. With equal admitted pedestal error, we can reduce C_H and, definitively, accelerate the S&H acquisition time, extending the bandwidth. In fact, the circuit time constant (which is the closed loop pole during the sampling phase and which is equal to the zero of the feedback block) results $\tau = C_H \cdot (2\ R_{ON} + R_F) = 24ns$. For example, with C_H=60pF, R_{ON}= 50Ω, and R_F=300Ω, the bandwidth is around 6.6MHz.

A drawback of the proposed circuit is the low input impedance: this is equal to R_F during the sampling phase and during the hold phase, is R_F+R_F if V_{in} is very similar to V_{out}; otherwise, with the diodes conducting, it decreases to R_F. ← not ideal voltage reader ($R_{in} \to \infty$)

This imposes to put a fast buffer upstream the S&H, which is capable of providing the necessary current to be injected into the virtual node; if R_F=300Ω and $V_{in,max}$=5V, we must select an OpAmp with $I_{O,max}$>17mA.

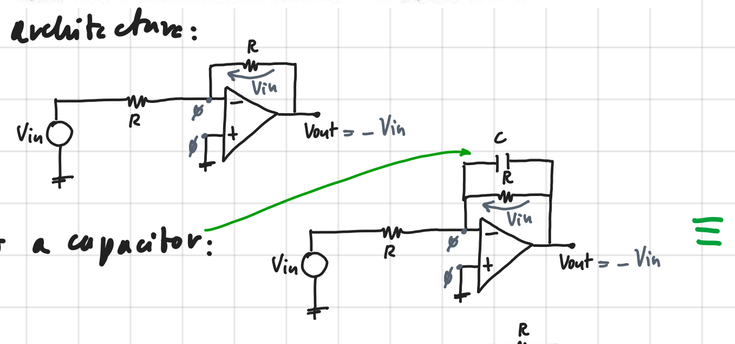
Fig. 6.41



- 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

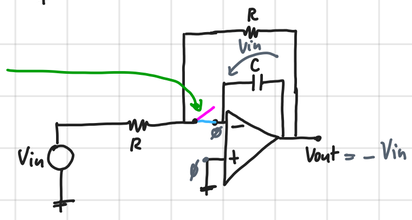
Obs. Proceeding by steps to understand this architecture:

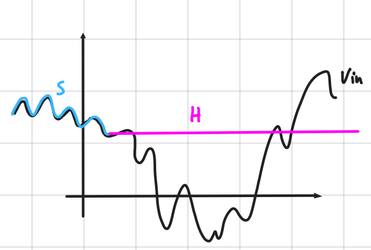
- Consider at first the simple stage:



- To hold that $V_{out} = -V_{in}$ voltage we put a capacitor:

- To control the S & H phases we should put a switch:

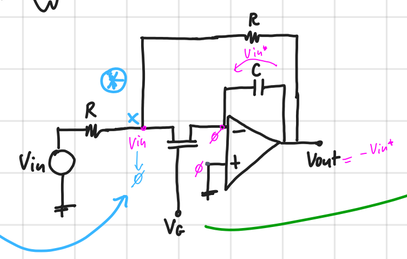


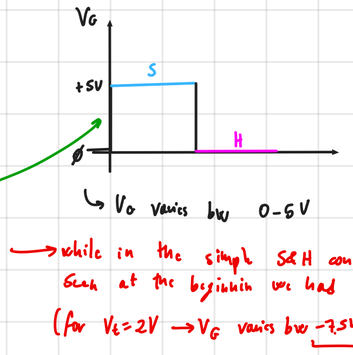


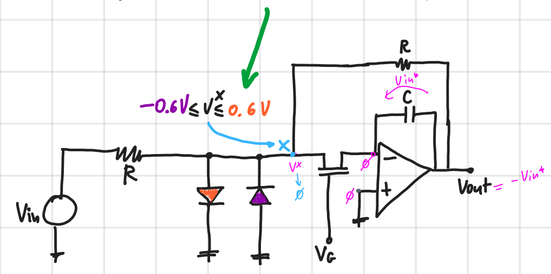
- To control the switch (MOS) we can apply a V_G (GATE) connected to a std CMOS level:

- To go to virtual ground at node X when we the MOS turns ON we can put diodes to force on that node a voltage which is not Vin and not far away from ø (virtual ground):

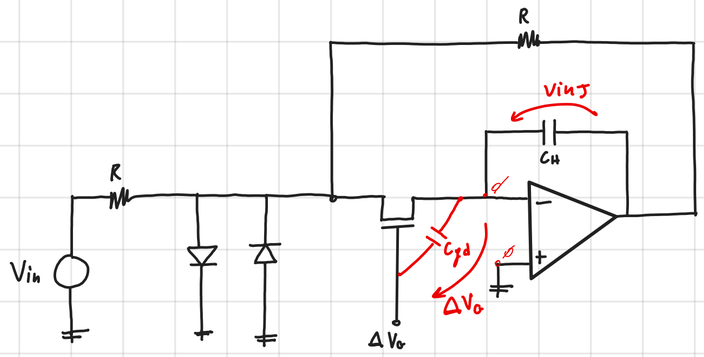
when we close the switch (MOS) we want this node to go to virtual ground from Vin → ø







- But we this configuration will have charge injection:

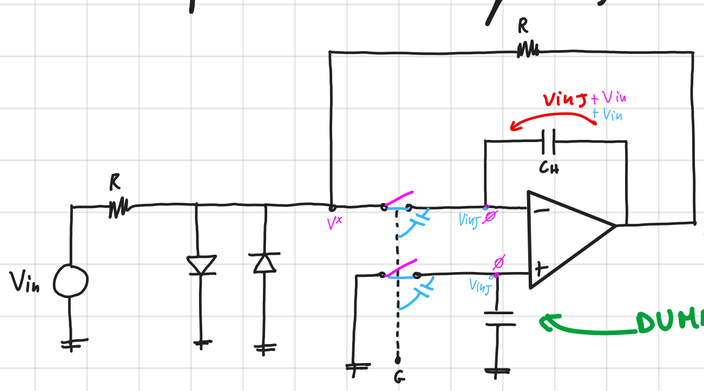


→ $V_{in_J} = \Delta V_G \frac{C_{gd}}{C_H}$

with: • $C_{gd} = \frac{Q}{\Delta V_G}$ → charge Q flows into C_H

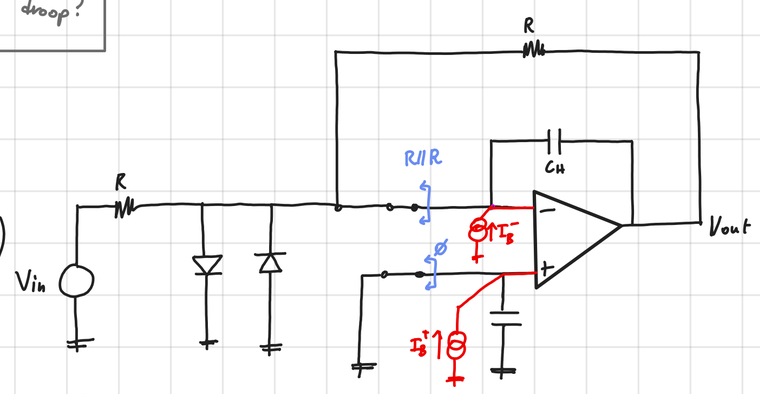
↳ $Q = V_{in_J} C_H$

- To compensate this charge injection we can move the virtual ground to a value that compensate the V_{in_J} when needed:



Bonus question for the oral exam (to get 30 L):
Does these improve circuits also compensate droop?

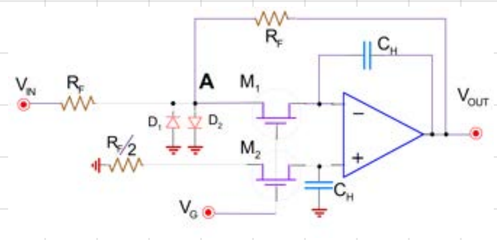
- Now consider the problem that might arise from bias currents: (consider the switches closed)

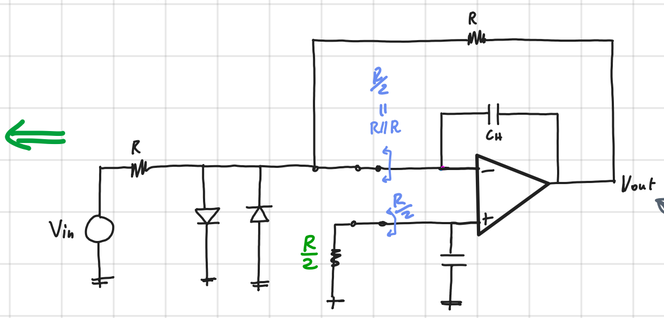


I_{B+} sees ø impedance → no voltage drop

I_{B-} introduces instead a voltage drop on the output due to impedance R//R ≠ 0

Final architecture





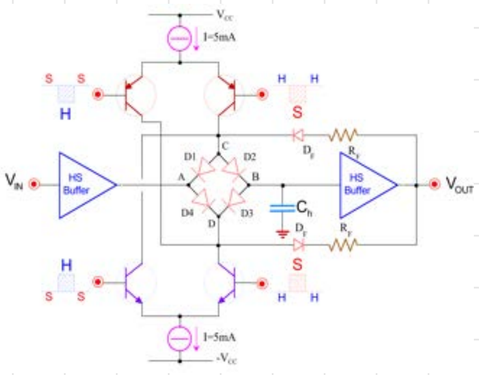
add $\frac{R}{2}$ resistance on non inverting pin

to compensate this we add an opposite effect from I_B^+

considering that $R//R = \frac{1}{\frac{1}{R}+\frac{1}{R}} = \frac{R}{2}$

Speed up the switch

Fig. 6.42



- fast diodes
- criss-cross

- poor accuracy
- timing issues to solve

Analyze the S&H topology shown in Fig. 6.42, which is made with bipolar technology and characterized by a high signal acquisition speed. Because the BJTs are not good voltage switches as the FETs, they can be employed as current switches within differential stages as current-steering. In the sampling phase, BJTs on the right branch of the differential stage are turned on, and the current I flows in the diode bridge. The potential of the two collectors and the nodes A and B are forced by the input, lower than the transistor saturation. If the diodes are identical with the same current I/2, then V_{CH} perfectly coincides with V_{in}. This condition is reached from any initial condition. Suppose that, for example, V_{in} = +5 V and V_{CH} = 0 initially holds. D1 cannot be turned on while D2 takes all the current I, which will charge C_H. Conversely, V_D can only be at V_{in}-0.6V=4.4V, causing the interdiction of D3, because the current I imposed by the *npn* will flow into D4.

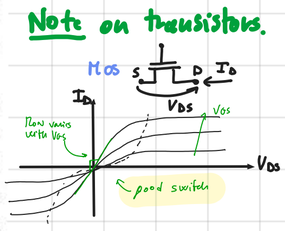
The process will end only when V_B reaches 5V. It is understandable therefore that the choice is dictated by the speed with which you want to charge the hold capacitance according to the ramp: $dV_{CH}/dt=I/C_H$. In the case of I=5mA and having Vin=±5V (i.e. FSR=10V), to complete the acquisition in less than 50ns, it is required to have C_H<25pf.

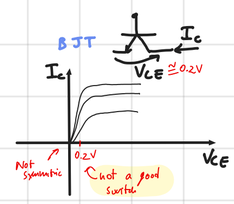
During the hold phase, the left of the differential stages are turned on instead of the transistors. Now, the voltage across C_H is no longer affected by V_{in}, as no current is supplied to the capacitor because of the polarity of the diode bridge. Indeed, instead of leaving the diode bias null (or slightly negative), we adopt the solution shown in Fig. 6.42, forcing the current in the opposite nodes of the bridge since the output cannot vary, i.e. $V_{out}=V_{CH}$. In this way, during the switching moment, the point C is brought to the voltage $V_{out}-R_F \cdot I-0.6V$ while the point D rises to $V_{out}+R_F \cdot I+0.6V$, i.e. the bridge is inversely biased with a value of $2 \cdot R_F \cdot I+0.6V+0.6V$, equal to 11.6V if we use R_F=1kΩ. This shrewdness is to reduce the diode depletion capacitance and to contain the effect of the signal feed-through. Moreover, the inverse excursion applied to across every diode is constant, independently of V_{in}, to avoid the introduction of non-linearity due to residual charge-injection.

Instead of normal *pn* junction diodes, employing Schottky diodes (metal-semiconductor junction) or hot-carrier diodes or diodes made with semiconductors with high electronic mobility (III-V compounds as GaAs or InP), we can significantly reduce the direct junction parasitic capacitance, allowing the improvement of the speed during the sampling phase, i.e. the inverse driving of the bridge. Thanks to the lower diodes threshold (around 0.5V) with respect to that of an FET switch (V_t of few volts), with the approach just described, we can obtain $t_{aperture}$<10ps.

Unfortunately, this circuit is very fast, but suffers from high inaccuracy due to the mismatches between the components. In fact, if the upper stage current is different from that of the lower branch during the sampling phase, the difference should vary the potential of the hold capacitance. Another equally problematic issue is the synchronization of the bipolar couples command signals, which can cause a very high error. For example, a delay Δt of only 10ps in the switching of one of two currents I=5mA causes a discharge or a charge of C_H=25pF and, definitively, an offset of 2mV.

Note on transistors.



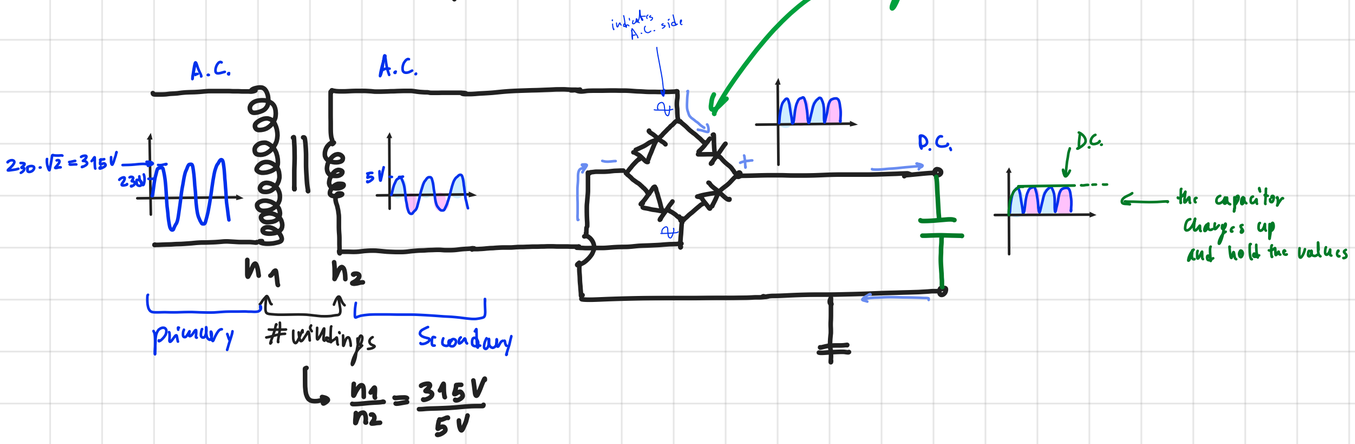


DIODES' BRIDGE

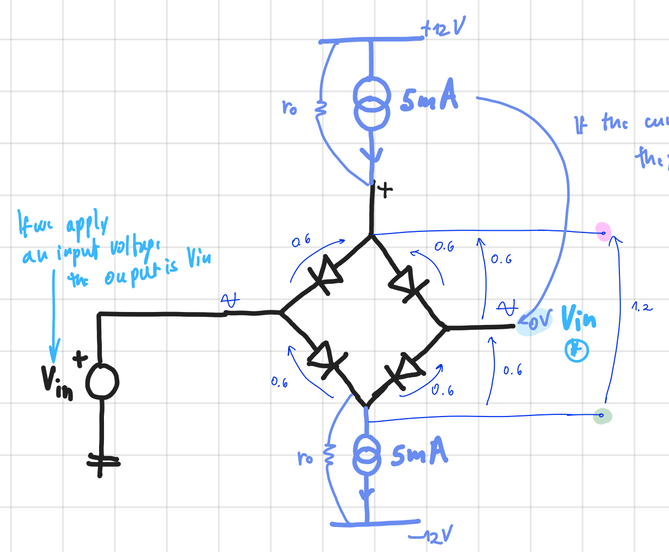
(next p.)

Note on diodes bridges

Consider the P.S. example:



Now consider:



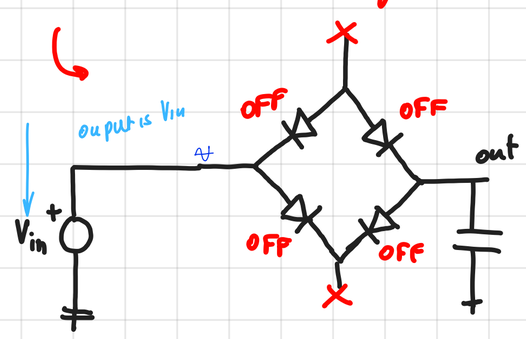
(without an input Vin)

If the current generators have their to (in this case equal) they can set the voltage here → so also here and here through voltage differences

|||| with current

switch ON

→ If we remove the current generators → no current through diodes



DIODES OFF

↓

no control over the output

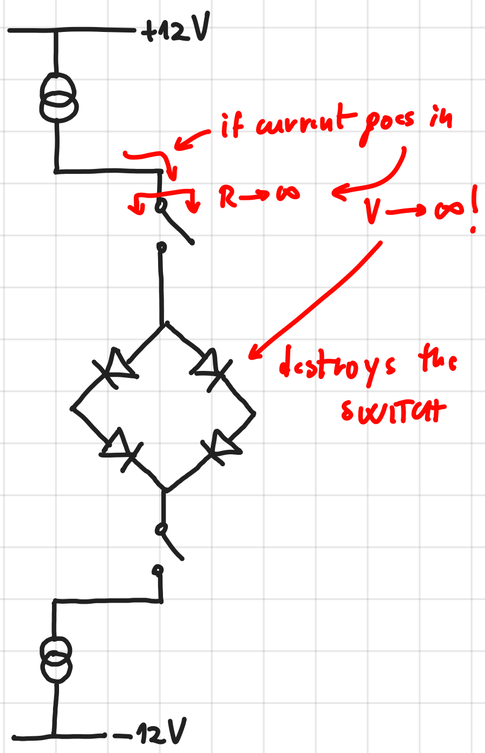
|||no current

switch OFF

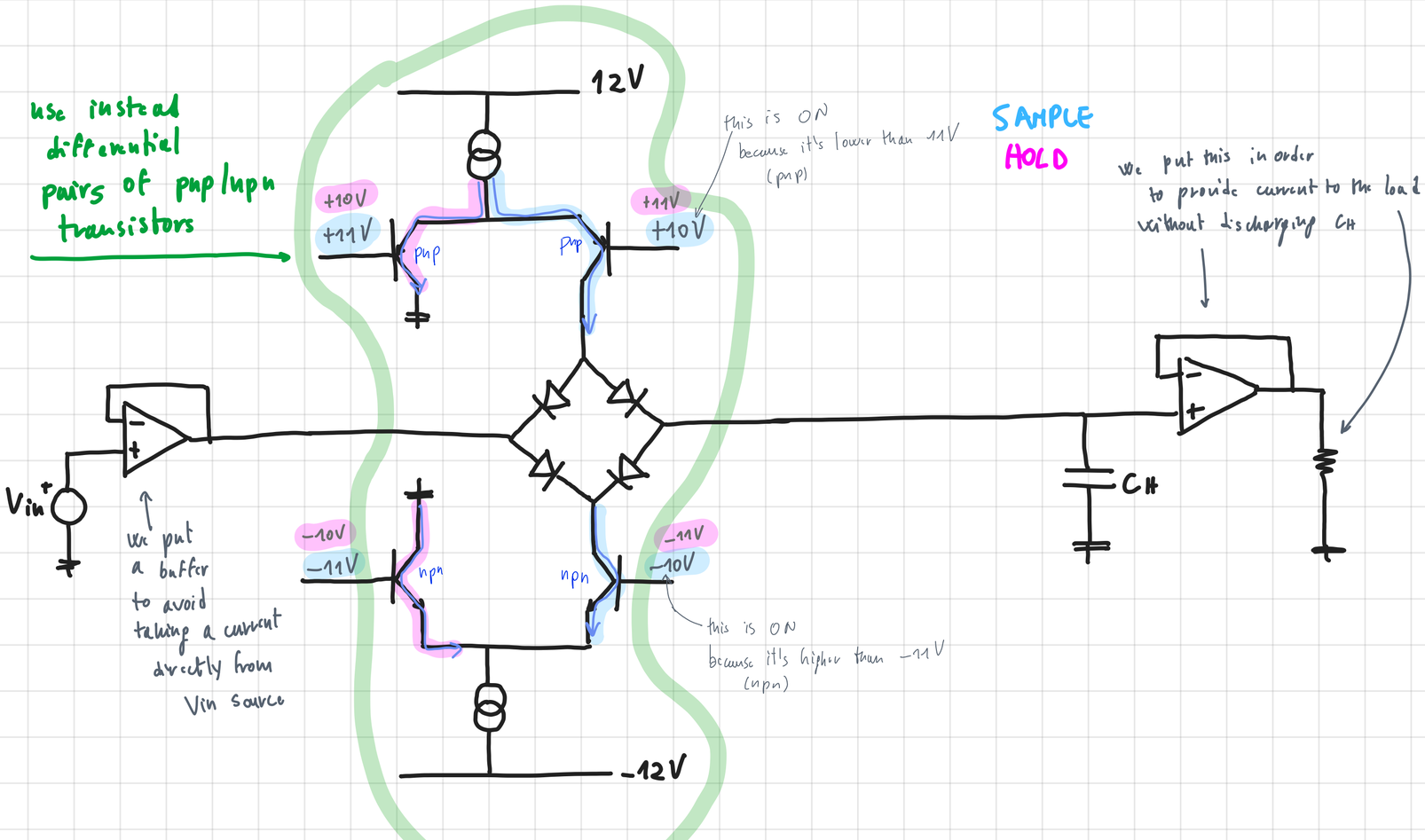
↳ To control the current, so the bridge/switch ON and OFF modes, we can put another (different) switch bw the current sources and the bridge:

SWITCH in the SH circuit

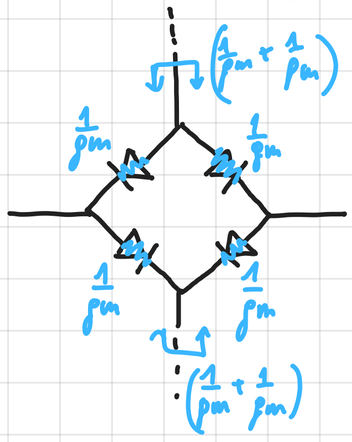
Wrong connection



use instead differential pairs of pnp/npn transistors



↳ Resistance of the diodes



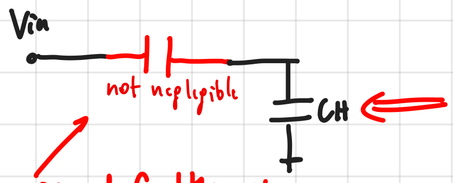
$\equiv$

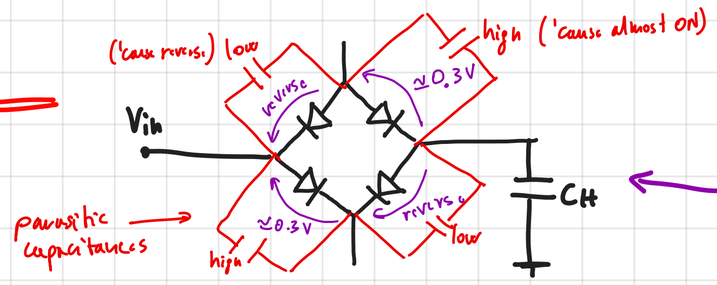
$\left(\frac{1}{g_m} + \frac{1}{g_m}\right) \parallel \left(\frac{1}{g_m} + \frac{1}{g_m}\right) \simeq$ just few Ω (low resistance) → good!

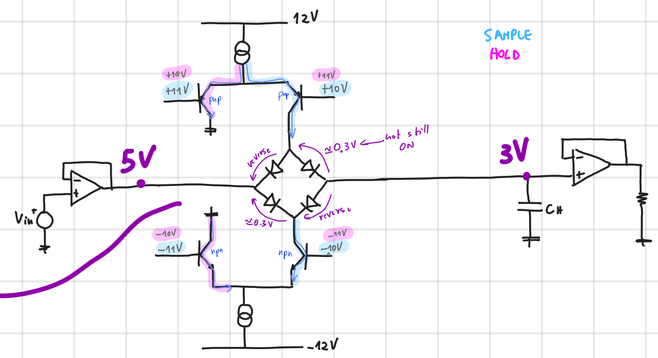
↳ This switch presents a disadvantage in terms of feedthrough → indeed consider the case where the CH has a value smaller than the one of Vin → 2 diodes will be forward bias and 2 not.

almost → not still ON

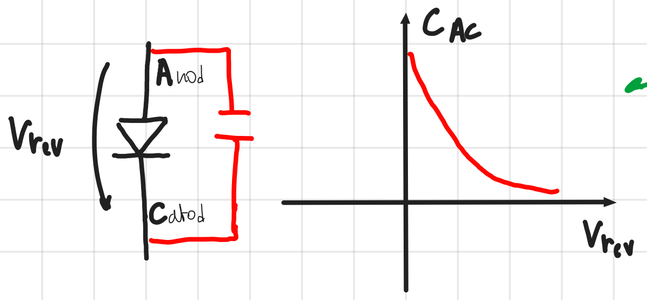
reverse-bias



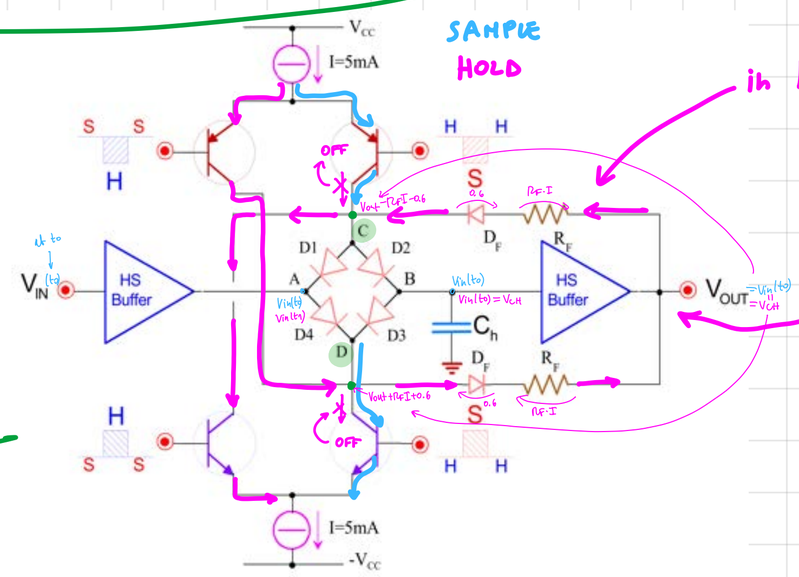




↳ To reduce this effect we need to decrease the parasitic capacitance → we need to increase the reverse bias



So we connect the bridge in this way

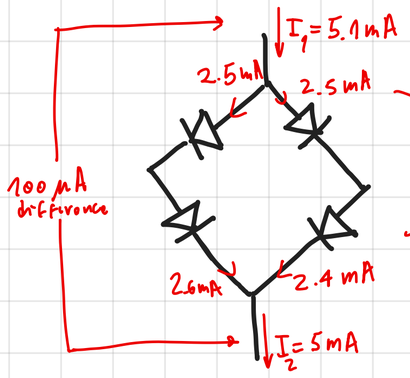


in HOLD / switch OFF mode we let the currents flow into a diode D_F and resistor R_F and the currents compensate each other

now at points C and D we have a higher voltage
↳ D2 and D4 go into reverse too

we reduce the parasitic capacitances

↳ Another issue could be if there are mismatches b/w the currents generators, the current difference will go into C_H and charge it till it saturates



we don't know how it will be divided b/w the diodes
↳ it changes the voltages across the diodes

the voltage on C_H will have an offset due to these mismatches and will not be precisely Vin (S&H)

↳ Another issue could be a timing issue b/w the switching of the transistors → so the two currents are delayed
↳ can cause a charge/discharge of C_H
↳ error

Feedback structures

To improve an S&H, it is very useful to increase the availability of the current to rapidly charge the capacitor and treat wideband signals. For this reason, we use structures as those shown in Fig. 6.43. In this structure, if the output resistors of the first OpAmp and R_{ON} of the MOSFET are neglected, the acquisition time should be limited only by the slew-rate and the maximum output current of the OpAmp.

more strict limitation b/w SR/Iout max

The presence of the two buffers introduces a double error in the sampling accuracy, caused by the errors ε between the inputs of every OpAmp. The first buffer OpAmp loses its meaning because it is possible to employ a structure with full feedback, shown in Fig. 6.44. To further increase the acquisition speed, a follower realized with a BJT downstream the switch is inserted to reduce the series resistance to the value $1/g_m+R_{ON}/\beta$ of the conductive BJT (*npn* for $V_{in}>0$, *pnp* for $V_{in}<0$).

We must note that, in this circuitry, the hold capacitance affects the S&H performance levels and is very important to determine the closed loop stability. In fact, the network $C_H \cdot 1/g_m$ on the forward branch adds a pole in the transfer function A(s) of the first OpAmp while the second, a buffer, presents a closed loop pole in its GBWP. The consequent Bode diagram for the forward gain and for the unity feedback is depicted in Fig. 6.45.

To compensate the S&H, we could use the pole due to C_H after the OpAmp's GBWP, reducing C_H or $1/g_m$ or R_{ON}. We can understand that the buffer can improve the S&H speed and, indirectly, the stability, too (neglecting singularities introduced by the parasitisms).

Another possibility for the S&H compensation is obtained by means of a reduction in the feedback effect, i.e. increasing the β block attenuation from β=1 in Fig. 6.44 to $\beta=R_1/(R_1+R_2)<1$ in Fig. 6.46. In this way, we ensure the stability of the overall S&H, as shown in the Bode diagram in Fig. 6.47. Now, the S&H has a gain equal to $1/\beta=1+R_2/R_1$.

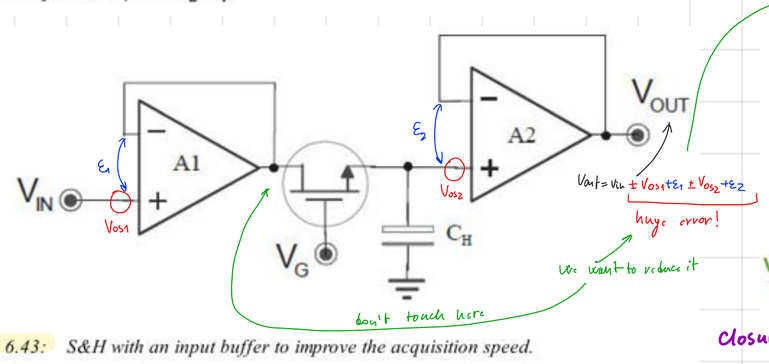


Fig. 6.43: S&H with an input buffer to improve the acquisition speed.

We can understand that the main limitation of the feedback architectures is due to the trade-off between the speed and the stability. In fact, we must ensure that the pole originated from the capacitance C_H (together with those introduced by the OpAmps) does not cause an excessive phase shift in the input signal. Therefore, feedback architectures are usually compensated conservatively in order to avoid the worst cases for components' tolerance levels and cannot reach the maximum speed for a given technology. We observe how the pole in Fig. 6.45 and Fig. 6.47 depends on V_{in} for g_m and the MOSFET R_{dsON}.

Other circuits

Other circuitry solutions for feedback S&Hs use an Operational Transconductance Amplifier (OTA) in which fully-differential architectures allow an efficient error rejection. Moreover, using the current command of the OTA, it is possible to save also the external switch, as shown in the scheme shown in Fig. 6.54. In the case depicted in figure, during the sampling phase, the voltage stored in the capacitor C_H is $V_{IN}+2\cdot V_{BE}$.

C_H, connected between the ground and the floating node, is sensitive to the non-linearity problems due to the pedestal error. To avoid this drawback, we can use the configuration shown in Fig. 6.55, in which the switch is connected to the virtual ground node, and the change between sampling and hold causes a constant output offset. To reduce it, one can use the redundant architecture depicted in Fig. 6.56.

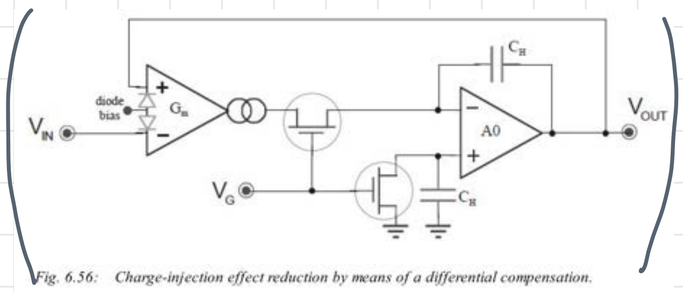
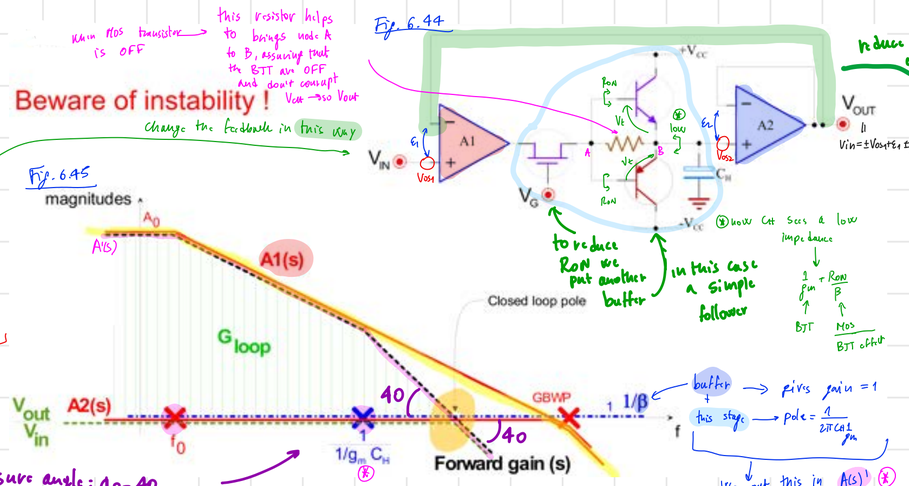
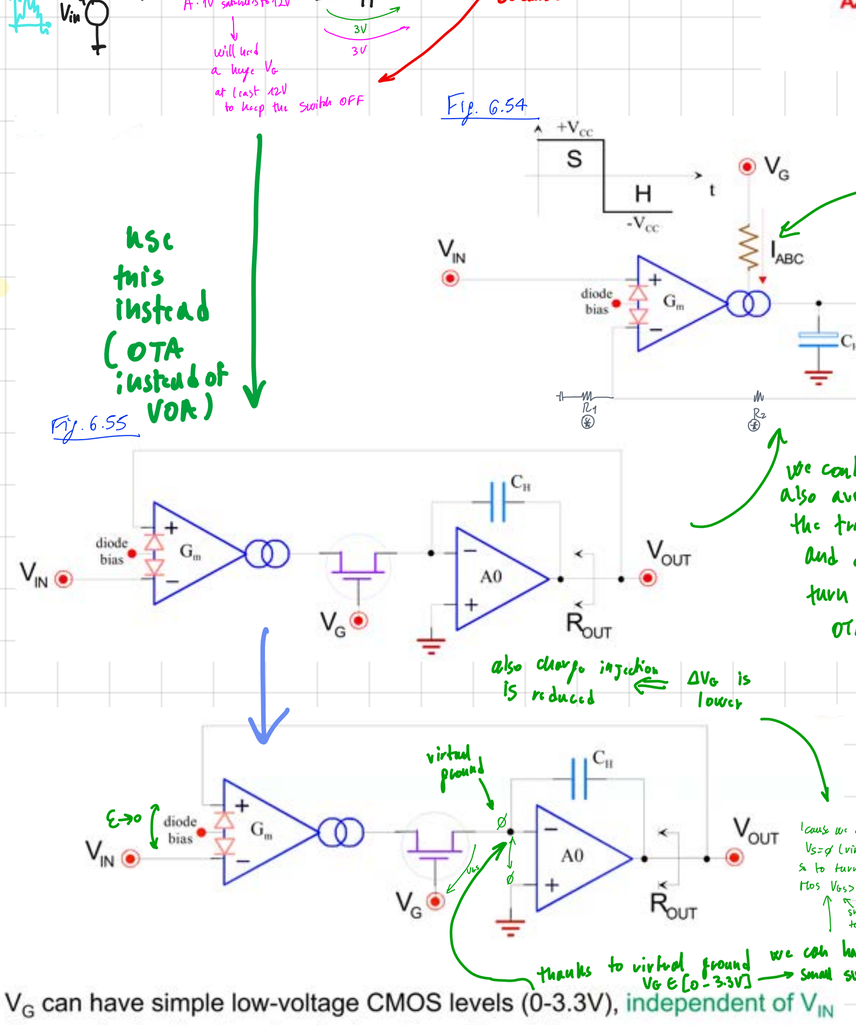


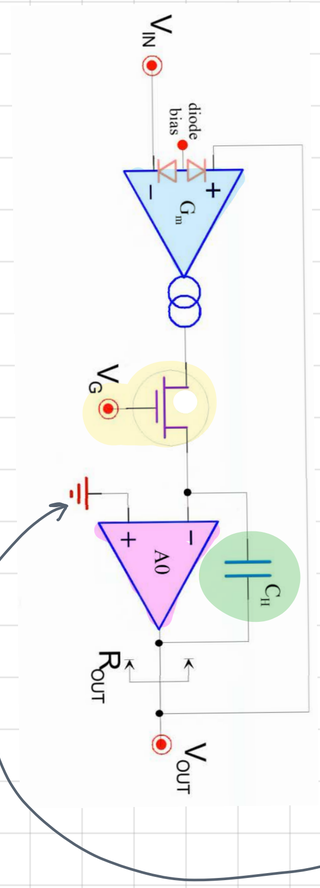
Fig. 6.56: Charge-injection effect reduction by means of a differential compensation.



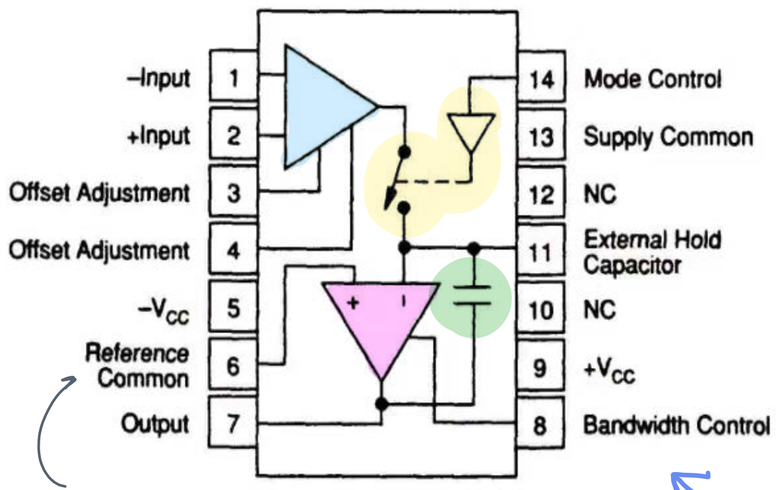


- V_G can have simple low-voltage CMOS levels (0-3.3V), independent of V_{IN}
- (thus reducing the *Aperture-Induced non-linearity*)
- What about stability? ← We should always check it
- What about the need of linearization through I_{DIODE}? We actually don't need linearization, because the feedback will make ε→0 ⇒ No problems from OTA Non-linearity

Datasheets



general-purpose S&H



If we apply ground $V_g \in [0, 3.3V]$

High-Speed Bipolar Monolithic SAMPLE/HOLD AMPLIFIER

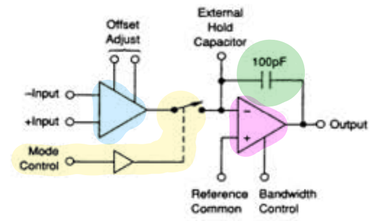
FEATURES

- ACQUISITION TIME TO 0.01%: 1.5µs max
- HOLD MODE SETTLING TIME: 350ns max
- DROOP RATE AT +25°C: 0.5µV/µs max
- TTL COMPATIBLE
- FULL DIFFERENTIAL INPUTS
- INTERNAL HOLDING CAPACITOR
- TWO TEMPERATURE RANGES:
  –40°C to +85°C (KH, KP, KU)
  –55°C to +125°C (SH)
- PACKAGE OPTIONS: 14-pin Ceramic, Plastic DIP, and 16-pin SOIC

↖ Surface mounting IC

APPLICATIONS

- PRECISION DATA ACQUISITION SYSTEMS
- DIGITAL-TO-ANALOG CONVERTER DEGLITCHER
- AUTO ZERO CIRCUITS
- PEAK DETECTORS



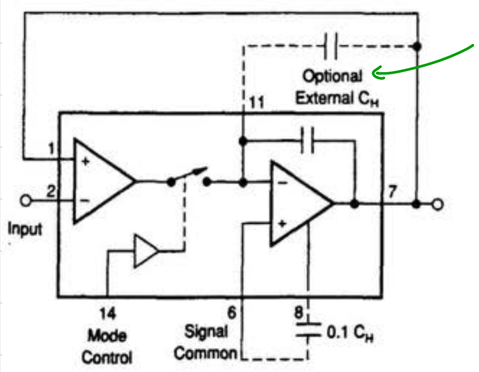
DESCRIPTION

The SHC5320 is a bipolar monolithic sample/hold circuit designed for use in precision high-speed data acquisition applications.

The circuit employs an input transconductance amplifier capable of providing large amounts of charging current to the holding capacitor, thus enabling fast acquisition times. It also incorporates a low leakage analog switch and an output integrating amplifier with input bias current optimized to assure low droop rates. Since the analog switch always drives into a load at virtual ground, charge injection into the holding capacitor is constant over the entire input voltage range. As a result, the charge offset (pedestal voltage) resulting from this charge injection can be adjusted to zero by use of the offset adjustment capability. The device includes an internal holding capacitor to simplify ease of application; however, provision is also made to add additional external capacitance to improve the output voltage droop rate.

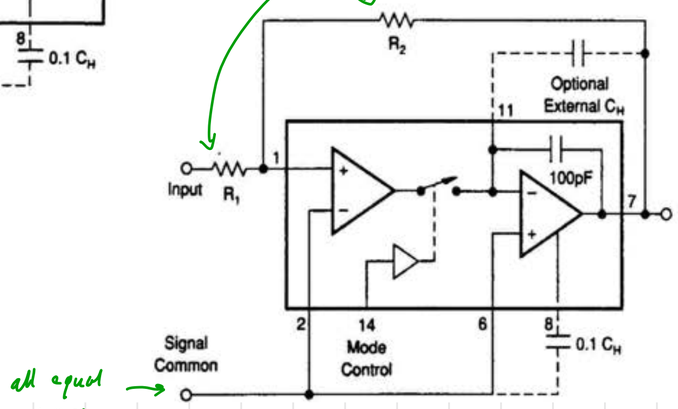
The SHC5320 is manufactured using a dielectric isolation process which minimizes stray capacitance (enabling higher-speed operation), and eliminates latch-up associated with substrate SCRs. The SHC5320KH, KP, and KU feature fully specified operation over the extended industrial temperature range of –40°C to +85°C, while the SHC5320SH operates over the temperature range of –55°C to +125°C. The device requires ±15V supplies for operation, and is packaged in a reliable 14-pin ceramic or plastic dual-in-line package, as well as a 16-pin surface-mount plastic package.

Examples of different configurations:

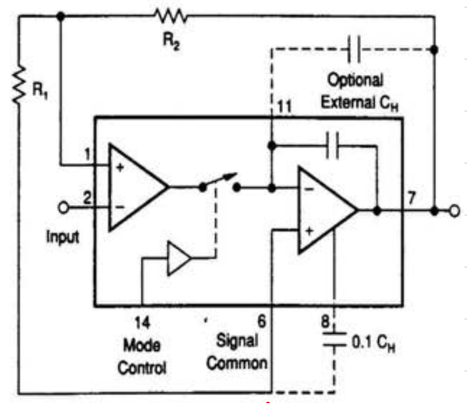


optional external capacitor not this

for gain amplification



all equal common



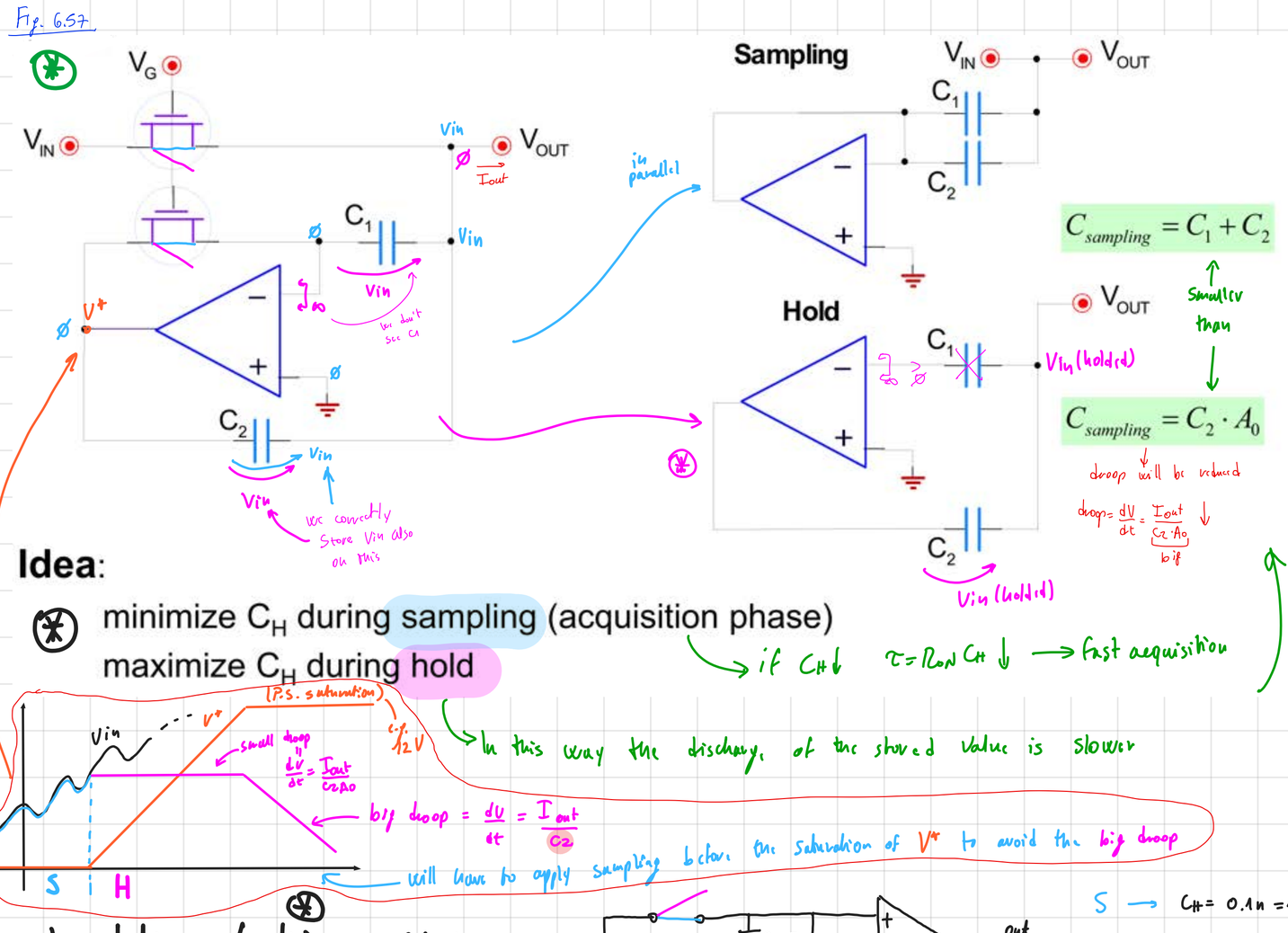
wrong design (or drawing) this should be connect to ground

Other improvement circuits

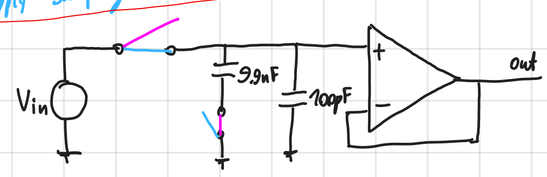
The trade-off between the speed and the accuracy can be avoided by distinguishing between the capacitance used in the acquisition phase, $C_{Sampling}$, and the hold one, C_H. This is implemented effectively in Fig. 6.57. It is a Miller amplifier, which is AC-coupled (through C_1 and C_2). During the sampling phase, the amplifier is a buffer with the inverting terminal as a virtual ground; for this reason, across the parallel capacitors $C_{Sampling}=C_1+C_2$, there is the input voltage. Moving then to the hold phase, the two capacitors and the OpAmp forms feedback. The equivalent C_{Hold} capacitance is that seen from the output node. To compute it, we can note that the output is a loop node; therefore, the impedance seen is that seen with an open loop ($1/sC_2$), which is reduced by the factor $1-G_{loop}$. Because the loop gain is A_0, the equivalent capacitance connected to the output node is $C_H \cong A_0 \cdot C_2$. For example, with typical values of $C_1=C_2=5pF$ and $A_0=10{,}000$, we have $C_{Sampling}=10pF$ while $C_{Hold}=50{,}000pF=50nF$.

In this way, we obtained hold capacitance several orders of magnitude greater than the charged sample capacitance. There is not any trick; the feedback is such that any injected or extracted current into or from the output node is not taken by C_1, but only by C_2; the other C_2 pin has not a fixed voltage, but the OpAmp output changes it with an opposite direction during its discharge, maintaining the virtual ground on the non-inverting terminal. During the hold phase, the absorbed output current $I_{out}=1\mu A$ determines a V_{out} discharge with a speed $dV_{out}/dt=I/C_{Hold}=20V/s=20\mu V/\mu s$. However, to discharge a stored voltage of 5V, it is not possible to sink a current for 0.25s because the OpAmp has an output limited by the supply voltage, for example 5V. Therefore, the stage works until the OpAmp output voltage reaches the maximum value of +5V starting from 0V, i.e. for a period less than $\Delta t=\Delta V \cdot C_2/I=25\mu s$.

Fig. 6.57



Idea (Step by step): We could think a good implementation of this could be:



S → $C_H = 0.1n = 100\,pF$ (small)

H → $C_H = 9.9nF + 0.1nF = 10nF$ (big)

↳ but this config. has a problem since in the S phase the 9.9nF is disconnected it means it doesn't store basically no voltage, so when in H phase we close it on 9.9nF it will be a very small voltage on it.

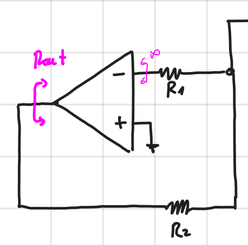
↳ charge will on 100pF → $Q = 0.1n \cdot 3V$ (e.g. voltage) $\left(C = \frac{Q}{V}\right)$

but now C during H is 10nF

$\Rightarrow Q = 10n\,V^* \longrightarrow V^* = 3V \frac{0.1n}{10n} = 30mV$ ← big error!

it's better to use this ✱

✱ Consider the simpler circuit:

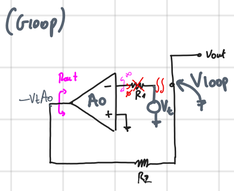


$\rightarrow R_{out} = \frac{(R_1 + \infty) \,//\, R_2}{1 - G_{loop}} = \frac{R_2}{1 - G_{loop}} = \frac{R_2}{1 + A_0}$

now consider Z_2 instead of R_2 with $Z_2 = \frac{1}{sC_2}$

$Z_{out} = \frac{\frac{1}{sC_2}}{1 + A_0} \cong \frac{1}{sC_2 A_0}$ Cout (big 'cause A_0 big)

(Gloop)



$\rightarrow V_{loop} = -V_t A_0$

$G_{loop} = -A_0$

ES11_ mux and digpot

Analog multiplexer

(Book p. 497)

Often, it is very important to acquire many analog signals from various sources and to process them by means of a single digital process, DSP or µC (or a memory, in the case of simple storage of acquired signal). A possible implementation of circuitry is shown in Fig. 6.60, in which there are N front-end circuits.

In this way, each channel has its *ad-hoc* designed analog stages for signal conditioning (filtering, amplifying, and possibly dynamic limitation and protection) and an S&H, with the appropriate synchronization circuitry. For economic reasons, it is better to employ a unique ADC converter, whose digital output can be connected to the processor.

To merge *n* analog inputs to a unique output, we need to introduce a new device named the Analog Multiplexer. Unlike the digital mux, which must only send the logic level of the selected input to the output, the analog mux must precisely pass the same analog information, introducing the minimum error on the voltage level (or current). The internal scheme of an analog mux is depicted in Fig. 6.61. Note the presence of the digital network for input selection, the appropriate drivers to drive the switch, and the *n* switches. The switches are realized with MOSFET transistors.

Fig. 6.60

Example of application:

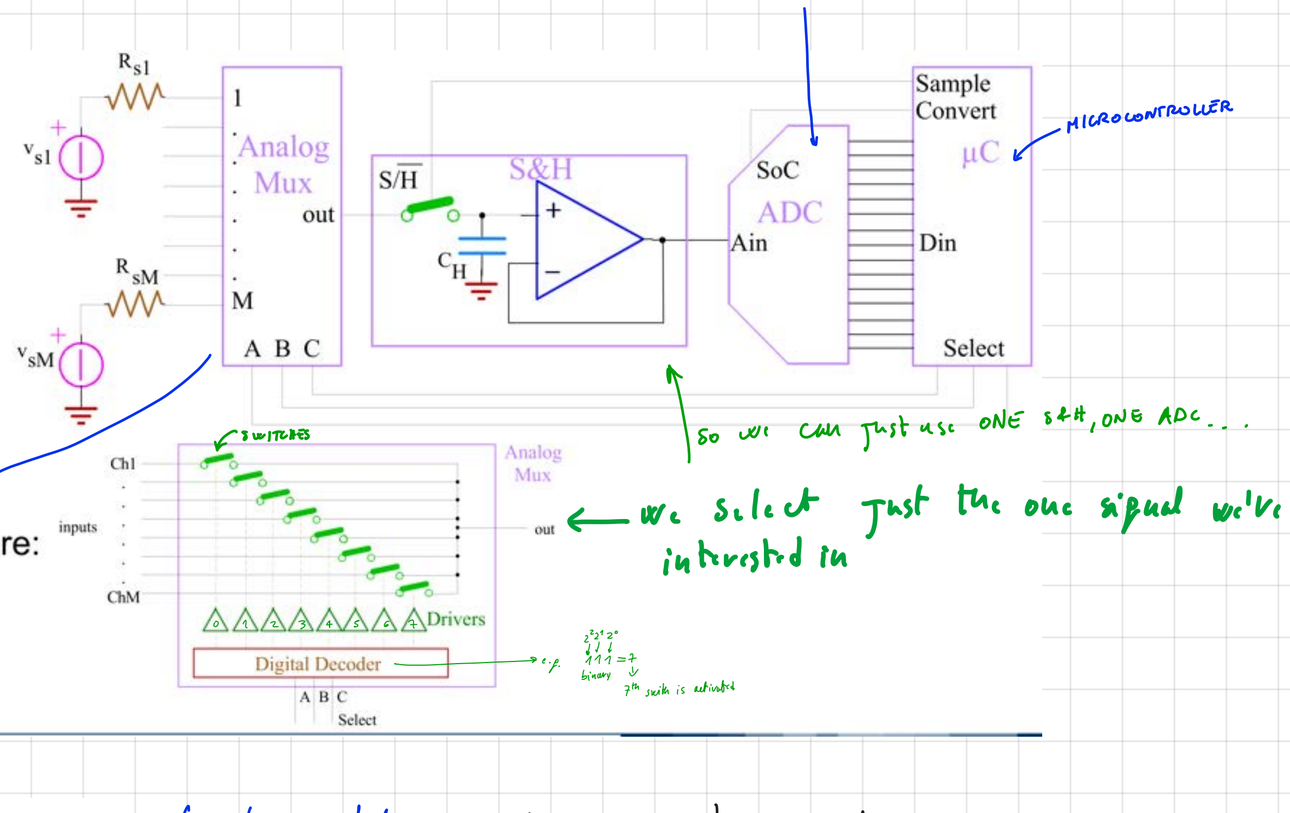
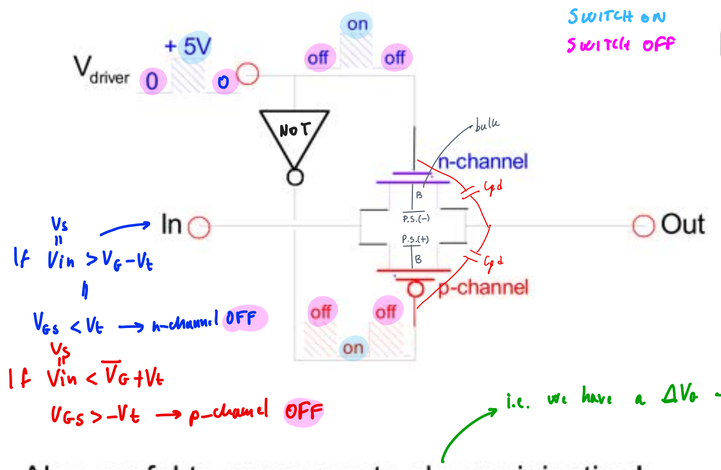


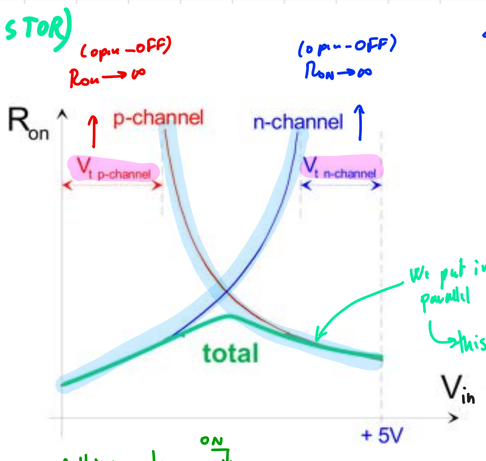
Fig. 6.61

Internal architecture:

Switch implementation:

Analog switch implementation: → (PASS-TRANSISTOR)





Also useful to compensate charge injection!

i.e. we have a ΔV_G → falling edge: ON → OFF on n-channel; rising edge: OFF → ON on p-channel → the charge injections are opposite and compensate each other

← for the switches: just one transistor is not enough, because analog inputs can vary a lot bw pos. and neg. values

↳ so to properly close the switch we shall use a transistor pair of p-channel and n-channel driven with opposite phases

Datasheet

Fig 6.62

19-0095; Rev 0; 9/95

MAXIM

Low-Voltage, 8-Channel/Dual 4-Channel Multiplexers with Latchable Inputs

MAX382/MAX384

1

General Description

The MAX382/MAX384 are low-voltage, CMOS, 1-of-8 and dual 4-channel muxes with latchable digital inputs. They feature low-voltage operation from a +2.7V to +16.5V single supply and from ±3V to ±8V dual supplies. Pin compatible with the DG428/DG429, these muxes offer low on-resistance (100Ω max) matched to within 4Ω max between channels. Additional features include off leakage less than 2.5nA at +85°C and guaranteed low charge injection (10pC max). ESD protection is greater than 2000V per Method 3015.7.

Features

- Pin-Compatible with Industry-Standard DG428/DG429, DG528/DG529, MAX368/MAX369
- Single-Supply Operation (+2.7V to +16.5V) Bipolar Supply Operation (±3V to ±8V)
- Low Power Consumption (<300µW)
- Low On-Resistance, 100Ω max
- Guaranteed On-Resistance Match Between Channels, 4Ω max
- Low Leakage, 2.5nA at +85°C
- TTL/CMOS-Logic Compatible

Applications

Battery-Operated Systems
Audio Signal Routing
Low-Voltage Data-Acquisition Systems
Sample-and-Hold Circuits
Automatic Test Equipment

Ordering Information

PART	TEMP. RANGE	PIN-PACKAGE
MAX382CPN	0°C to +70°C	18 Plastic DIP
MAX382CWN	0°C to +70°C	18 Wide SO
MAX382C/D	0°C to +70°C	Dice*
MAX382EPN	-40°C to +85°C	18 Plastic DIP
MAX382EWN	-40°C to +85°C	18 Wide SO
MAX382EJN	-40°C to +85°C	18 CERDIP**
MAX382MJN	-55°C to +125°C	18 CERDIP**

Ordering Information continued on last page.
* Contact factory for dice specifications.
** Contact factory for package availability.

Pin Configurations

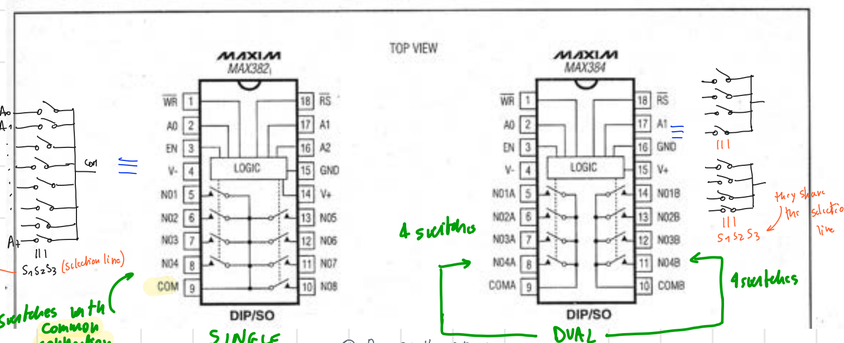
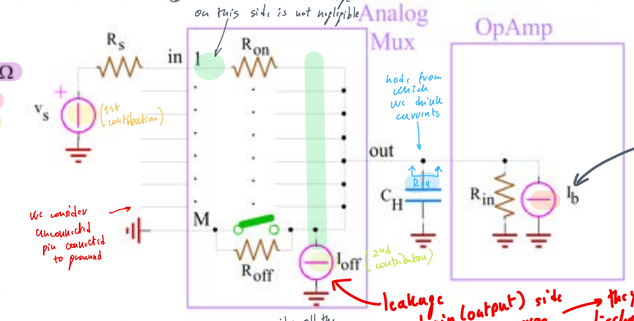


Fig. 6.63

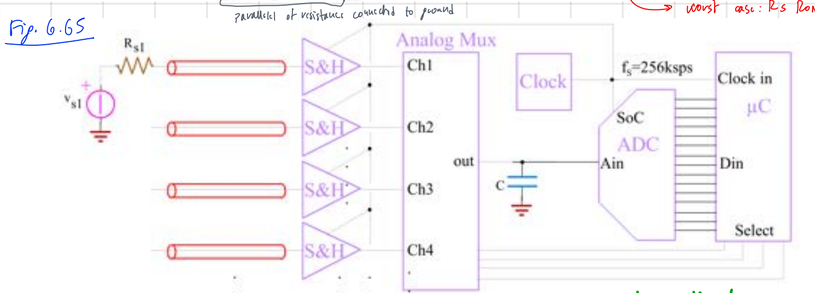
Sources of errors:

R_{on} 10Ω÷10kΩ
R_{off} >10MΩ
I_{off} <100nA
T_{on} T_{off} <100ns



$$\Delta V = \pm v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}} \pm (I_{off} \cdot (M+1) + I_B) \cdot \left\{(R_s + R_{on}) \| R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}$$

Fig. 6.65



Example: ADC with f_S=256ksps
4 "fast" and 12 "slow" input channels

S&H does not have to switch at $\frac{f_S}{M}$ (# channels)

we can relax the time constraint of the S&H (closed for most of the time when other inputs are ON)

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 main commercial analog multiplexers' static characteristics (for example, Fig. 6.62) are linked with the switches performance levels and are: R_{on}, maximum closed switch series resistance, usually between 10Ω and few kΩ; R_{off}, open switch minimum resistance, usually greater than some tens of MΩ; and, I_{off}, the maximum leakage current for every switch terminal (from drain/s to the bulk of the switches), in the order of few tens of nA. Among the dynamic performance levels, opening and closing times, T_{on} and T_{off}, respectively, are quoted; these values are usually in the range of few tens or hundreds of ns.

The non-ideal switches' parameters introduce an error in the output multiplexer value. To compute them, it is appropriate to refer to the equivalent circuit depicted in Fig. 6.63. Practically, the mux induced error is due to the difference between the source signal value v_s and those effectively downstream the mux, for example as an OpAmp input. The error occurs as a voltage drop across the switch R_{on} and the selected source R_s due to the resistive partition with the OpAmp input R_{in}, to the present leakage current leakage I_{off}, and to the OpAmp bias current I_b. Applying the superimposition principle, we have:

$$\Delta V = \pm v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}} \pm (I_{off} \cdot M \pm I_b) \cdot \left\{(R_s + R_{on}) \| R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}$$

in which the signs + or – are coherently taken with the effective directions. In general, increasing the number of inputs increases the error because of the progressive add up of the leakage currents and of the open switches' parallel parasitic resistance values. Moreover, increasing R_{in}, we reduce the contribution due to the partition of v_S, increasing the others until R_{in} is comparable with $R_{off}/(M-1)$.

For example, choosing the mux ADG508A, with M=8 channels, R_{on}=400Ω, R_{off}=10MΩ, I_{off}=100nA (at 125°C), v_s=±15V, Rs=1kΩ, and a BJT OpAmp with R_{in}=500kΩ and I_b=0.5µA, we obtain a maximum error of ΔV=±56mV±0.98mV±0.7mV. The first contribution corresponds to ½LSB if we use downstream a 10bit ADC while the sum of the other two is ½LSB in the case of a 15bit ADC. With a CMOS OpAmp with better performance levels (for example, R_{in}=10MΩ and I_b=50nA), we would have ΔV=±16.8mV±0.99mV±0.07mV, equal to ½LSB accuracy with 12 and 15 bit ADCs, respectively.

If the directions of the contributions were opposed, it would be possible to look for the compensation, imposing the equality between different components to have ΔV≈0:

$$v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}} = \left[I_{off} \cdot M + I_b\right] \cdot \frac{(R_s + R_{on}) \cdot \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}}{R_s + R_{on} + \left\{R_{in} \| \left[\frac{R_{off}}{(M-1)}\right]\right\}}$$

i.e. when:

$$R_{in} \| \left[\frac{R_{off}}{(M-1)}\right] = \frac{v_s}{I_{off} \cdot M + I_b}$$

In the preceding example, the parallel resistance should be 20MΩ. Unfortunately, this condition varies with the input signal and therefore cannot be ensured.

Conservatively, one can estimate the maximum error and sum all the contributions in their respective absolute values (in fact, the direction of I_{off} cannot be defined while the sign of v_s due to the partition is very clear). Assuming that the technologies used for the mux and the OpAmp are similar, i.e. $I_b \approx I_{off}$ and $R_{in} \approx R_{off}$, and that the series resistance R_s+R_{on} is effectively less than other resistance values, the preceding expression is simplified as follows:

$$\Delta V_{max} \approx \left|v_{s,max}\right| \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \frac{R_{off}}{M}} + \left|I_{off}\right| \cdot M \cdot (R_s + R_{on})$$

With the preceding values, we obtain $\Delta V_{max} \approx 17mV + 1.1mV$, in accordance with the results acquired by the most precise computations.

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 (error limit)

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 (error limit)

The use of an analog mux is convenient because it allows processing signals from various sources by a single ADC, as shown in Fig. 6.64. In this case, with a 16-to-1 mux and an ADC which can sample at f_S=256ksps, every channel can be scanned with a frequency of 16ksps and a total time per channel of only $1/f_S$=3.9µs (and not 1/16ksps!). This sensibly reduces the bandwidth that the system can acquire (less than 8kHz to avoid aliasing).

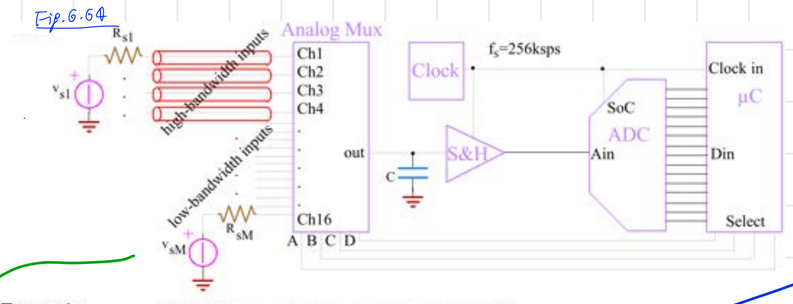
To solve this problem, we can intelligently manage the mux reading sequence, privileging the wideband signals in respect of the slowest signals. For example, if we have M=16 inputs and if only the first four (Ch1, Ch2, Ch3, and Ch4) have bandwidth of 20kHz, it should be possible to adopt a scanning as the following: Ch1, Ch2, Ch3, Ch4, **Ch5**, Ch1, Ch2, Ch3, Ch4, **Ch6**, Ch1, Ch2, Ch3, Ch4, **Ch7**… and so on, ensuring a sampling at $f_S/5$=51.2ksps for the first four while the other 12 inputs should be sampled at every $f_S/60$=4.2ksps, enough for bandwidth less than few kHz. → for low-BW inputs

Notice that, in case of high impedance sources, the settling time required to ensure good accuracy of the S&H acquired value can reach several tens or hundreds of µs, in addition to the time-slot available for each channel, which, also with the described sequencing, should be only 3.9µs per channel. Assume that the S&H samples the value at the end of the available time-slot for the *i-th* channel, and the ADC converts during the following acquisition *(i+1)th*. If the ADC has 12bits, requiring accuracy better than ½LSB, we should have a time constant less than:

$$\tau \le \frac{T}{\ln\frac{FSR}{0.5 \cdot LSB}} = \frac{T}{\ln\left(2 \cdot 2^n\right)}$$

(*) next p.

Fig. 6.64



Example:
M=16 (# channels)
n=12bit ADC, requested accuracy better than ½LSB

Acquisition time lower than: $\tau \le \frac{T}{\ln\frac{FSR}{0.5 \cdot LSB}} = \frac{T}{\ln(2 \cdot 2^n)}$ → τ < 3.9µs/9 = 434ns

with C_H=40pF we get R_s+R_{on}<11kΩ

i.e. τ<3.9µs/9=434ns. In case of parasitic capacitance of 40pF (maybe the same S&H C_H), this should impose a maximum limit on the series R_s+R_{on} of 11kΩ.

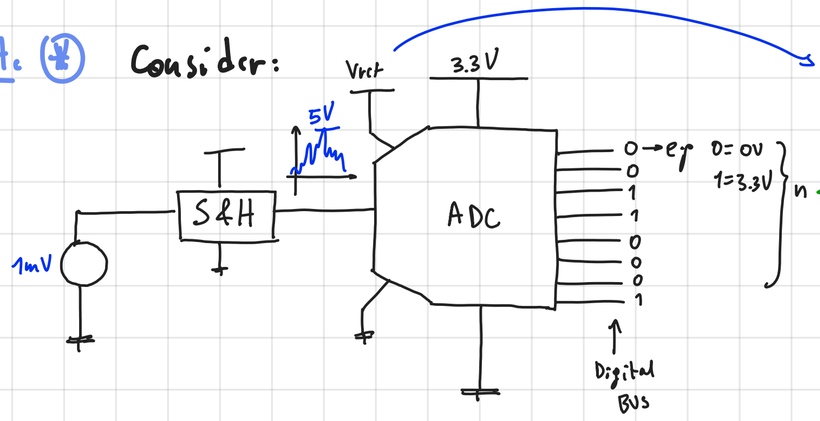
This is the reason for which, sometimes, it is necessary to put between the source and the analog mux a decoupling preamplifier stage (naturally with low noise and large bandwidth), as a general-purpose OpAmp buffer or, even, an INA.

To solve the limitation just presented, and for those applications in which you need to acquire the inputs at the same time and it is not possible to sample inputs at different moments, we must use the scheme shown in Fig. 6.65.

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

Note ① Consider:

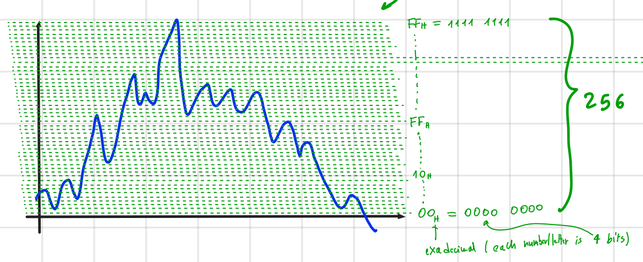


↳ I need to specify V_{ref} because for ex. the input signal is amplified from ±1mV → 5V

↳ Full Scale Range = 5V

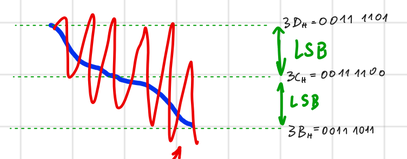
If we have n bits for the ADC → e.g. n=8 bit → $2^n = 2^8 = 256$

↳ This means that the input signal gets quantized in 256 levels



smallest level an ADC can convert = Least Significant Bit = $\frac{FSR}{2^n} = \frac{5V}{2^8} = 19mV = LSB$

Zoom:



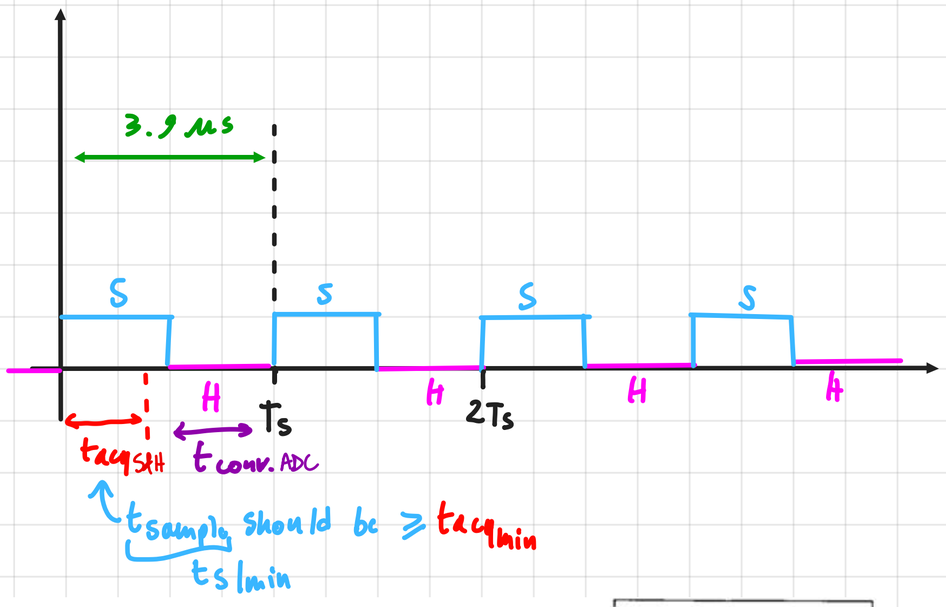
If we have an error > LSB the ADC will convert the actual value with a wrong level

We need that the error $\mathcal{E} \leq \frac{LSB}{2}$ to have a GOOD CONVERSION

Note ② Sample frequency of ADC and S&H vs. acquisition time

In the previous example: $f_s = 256\,ksps \rightarrow T_s = \frac{1}{f_s} = 3.9\,\mu s$

Time constraints:



- $t_{s_{min}} \geq t_{acq_{min}} = \tau \ln \frac{FSR}{\mathcal{E}} = \tau_{max}(n+1) \underbrace{\ln 2}_{\sim 0.7} = 0.7\,\tau_{max}(n+1)$

  $\mathcal{E} = \frac{1}{2} LSB = \frac{FSR}{2 \cdot 2^n} = \frac{FSR}{2^{n+1}}$

  total $R_{ON} = R_S + R_{ON}$

  $(R_{ON} + R_S)|_{MAX} C_H|_{MAX}$

- $t_H \geq t_{conv_{max}}$ (hold) = e.g. 10ns (flash ADC), 10 µs (SAR ADC), 10 ms (dual slope ADC)

- $t_s + t_H = T_s$

→ based on these constraint and the f_s available we can choose the most appropriate components for the S&H and ADC

So it's important that the next switches closes before changing the channel ← we just don't want that all the switches are open ↳ otherwise the OpAmp will go in open loop! ← we use the MUX to choose the resistors → to select the gain

Fig. 6.66

MC145040
MC145041

8-Bit A/D Converters With Serial Interface

CMOS

The MC145040 and MC145041 are low-cost 8-bit A/D Converters with serial interface ports to provide communication with microprocessors and microcomputers. The converters operate from a single power supply with a maximum nonlinearity of ± ½ LSB with a 5 V reference and ±1 LSB with a 2.5 V reference. No external trimming is required.

The MC145040 allows an external clock input (A/D CLK) to operate the dynamic A/D conversion sequence. The MC145041 has an internal clock and an end-of-conversion signal (EOC) is provided.

- Operating Supply Voltage Range: V_{DD} = 4.5 to 5.5 Volts
- Successive Approximation Conversion Time:
  - MC145040 — 10 µs (with 2 MHz A/D CLK)
  - MC145041 — 20 µs Maximum (Internal Clock)
- 11 Analog Input Channels with Internal Sample and Hold
- 0- to 5-Volt Analog Input Range with Single 5-Volt Supply
- Ratiometric Conversion
- Separate V_{ref} and V_{AG} Pins for Noise Immunity
- Monotonic Over Voltage and Temperature
- No External Trimming Required
- Direct Interface to Motorola SPI and National MICROWIRE Serial Data Ports
- TTL/NMOS-Compatible Inputs May Be Driven with CMOS
- Outputs are CMOS, NMOS, or TTL Compatible
- Very Low Reference Current Requirement
- Low Power Consumption: 11 mW
- Internal Test Mode for Self Test

P SUFFIX
PLASTIC DIP
CASE 738

DW SUFFIX
SOG
CASE 751D

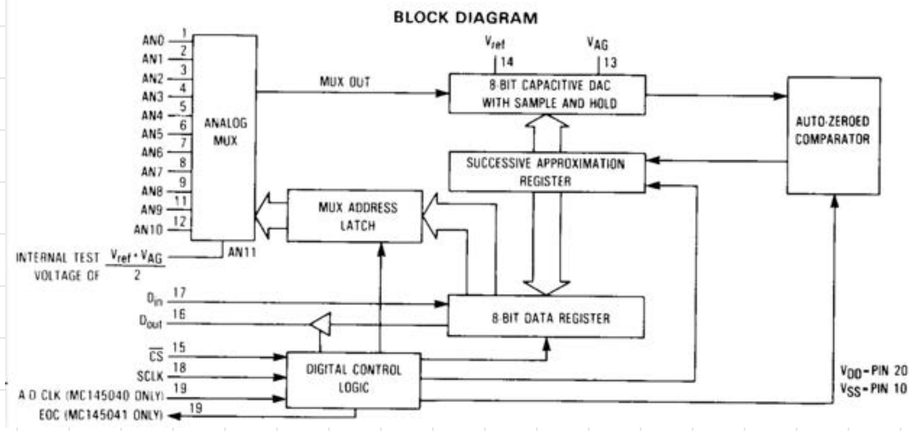
FN SUFFIX
PLCC
CASE 775

ORDERING INFORMATION

MC14XXXX

Suffix	
1	– 40 to 125°C
2	– 40 to 85°C
P	Plastic DIP
DW	SOG
FN	PLCC

BLOCK DIAGRAM



As was presented, now it is necessary to employ a number of S&Hs, which is equal to the number of channels for which we want to ensure the simultaneous acquisition. The sampling and hold signals are given to all the S&Hs, preserving the phase information for all the sources. It does not change anything about the number of samplings per channel while the time for each S&H to acquire the corresponding input is much expanded. For example, wanting the usual 12bit and 256ksps ADC with 4 "fast" channels (with large bandwidth signals) and 12 "slow" channels, every channel should have $5/f_S$=19.5µs to acquire the source. If all the 16 mux inputs should be connected to one S&H input, and the channel should be read in sequence, each one of the slowest should have the settling time of $16/f_S$=62.5µs.

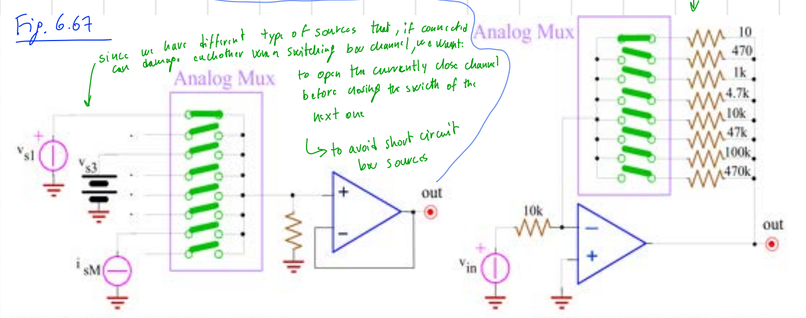
Since it is very useful to be able to acquire many inputs with the configuration depicted in Fig. 6.65, commercial ADCs have many analog inputs, one analog multiplexer, and only one S&H. Fig. 6.66 shows a data-sheet.

The analog multiplexers are important not only in multiple sampling circuits as those seen thus far, but also in all the applications where it is necessary to include or exclude analog devices or paths within a circuit. Fig. 6.67 depicts two typical uses. Practically, the first is that seen until now; indeed, the second employs a mux to change the OpAmp feedback resistance and, thus, to vary the gain.

It is important to make a clarification on the commercial multiplexers. Depending on the application, it can be necessary to ensure that the currently used switch opening and the following switch closing happen with a well-known timing. For example, in the application depicted in Fig. 6.67 (on the left), it is important that, before closing the following path, the currently used switch is open to avoid short-circuits between input sources. If you do not proceed in this way, the same sources (or even the analog mux) could be damaged permanently, or you could see intense spikes (even tens of mA) for each new input selection.

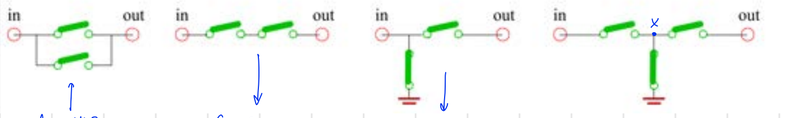
Vice-versa, in the application depicted in Fig. 6.67 (on the right), it is important that, before opening the currently used switch, the following is closed. Otherwise, for a certain time, the OpAmp should work in an open-loop configuration, and the output should saturate towards the supplies or, at least, should present intense voltage spikes (this time due to the charge injection). For these reasons, the manufacturers provide two different typologies of multiplexers, based on the command succession: Break-before-Make (BbM) and the opposite Make-before-Break (MbB). For example, Analog Devices offers two identical multiplexers: the ALD4201 as BbM and the ALD4202 as MbB.

Fig. 6.67



ALD4201 (**Break before Make**) B b M

ALD4202 (**Make before Break**) M b B



For MbB: we want to be sure that at least one switch closes (we don't want no open switches) ↓ the parallel switches should close/open together, but if one fail at least one it's close

For BbM: we want to keep the switches securely open ↓ so if one fails to open at least the other one it's open and we don't have a short circuit

If we want to be sure that we're open NO signal can reach the output (we connect it to ground through a switch that closes when the other opens)

Very secure connection ↓ we for sure kill any residual signal on X putting it in short-circuit

DigPot - Digital Potentiometer

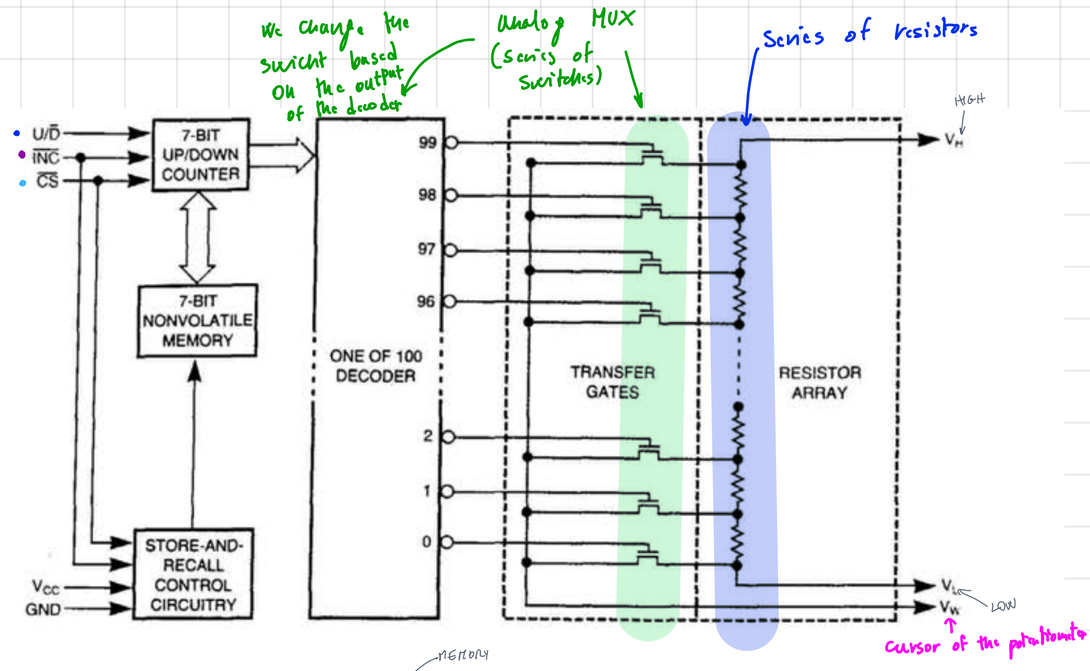
(Book p. 493)

In addition to the analog multiplexers and Universal Active Filters (UAF), covered by other textbooks, there are many useful devices for acquisition systems. An example of analog switches, in particular, is the digital potentiometer (DigPot). As for others, it is an extremely nice product and a very useful device.

The basic idea of the DigPot is to provide a three terminals analog device with the same functions as those of a sliding resistor with the central terminal continuously varying between the two extremes of the range. It is made with a series of N identical resistors, and the sliding terminal is taken with one of the N available switches (Fig. 6.58). The selection can happen by means of a binary address sent by an internal digital decoder or, smarter, a pulse that drives an internal up/down counter.

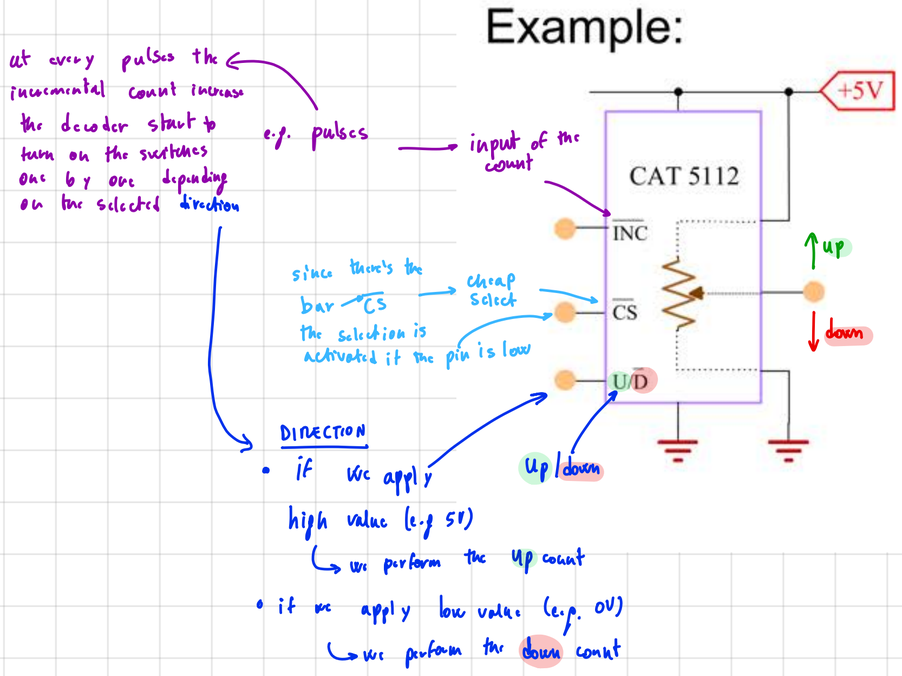
Fig. 6.58

Block diagram:



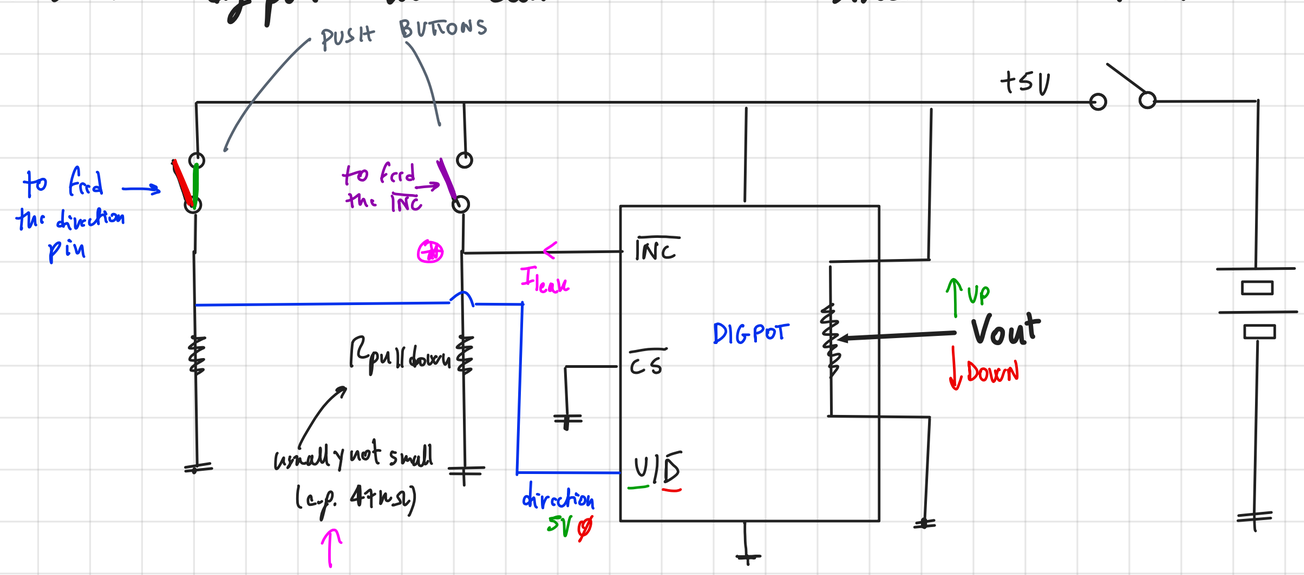
Model	Steps	Resistance (kΩ)	Configuration	Non-volatile (MEMORY)	Packages	Interface	Operating power	Features	Price (1000)
AD8402	256	10, 50, 100	Dual	No	DIP-14, SO-14	Three-wire	2.7 to 5.5V, 5 µA	Full ac specs, nA shutdown current	$1.66
AD8403	256	10, 50, 100	Quad	No	SOL-24	Three-wire	2.7 to 5.5V, 5 µA	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

Example:



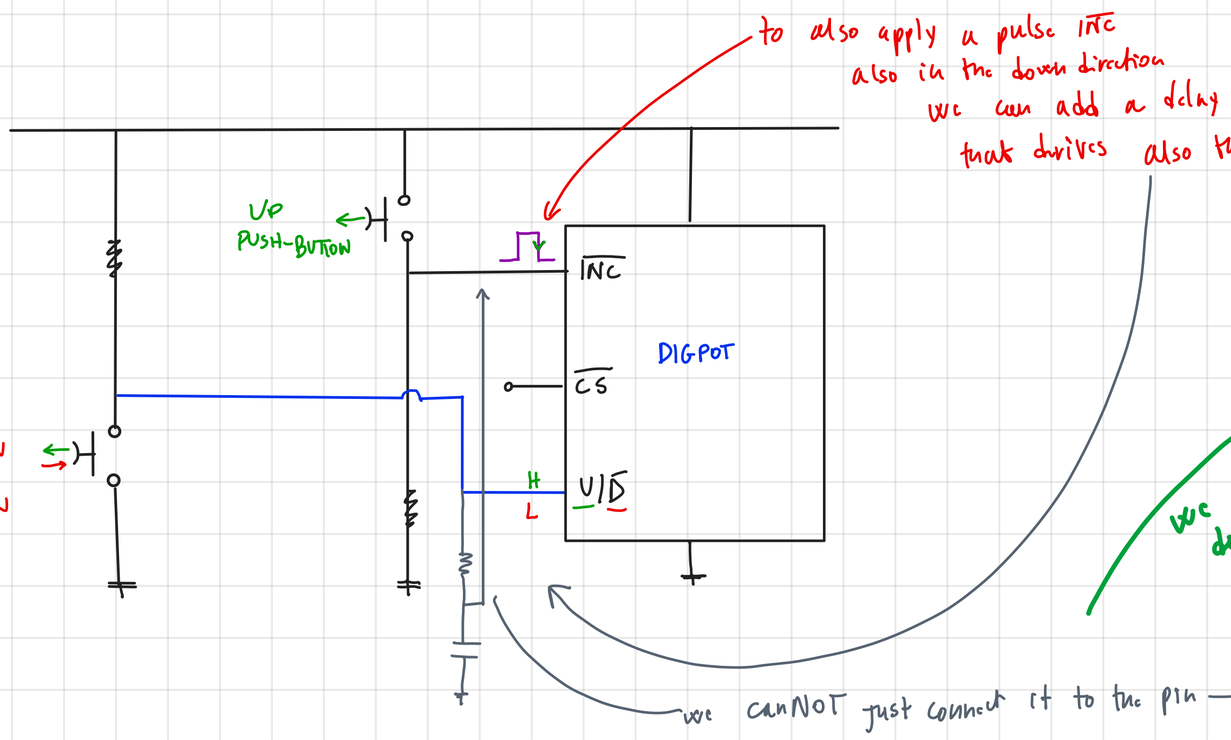
Examples

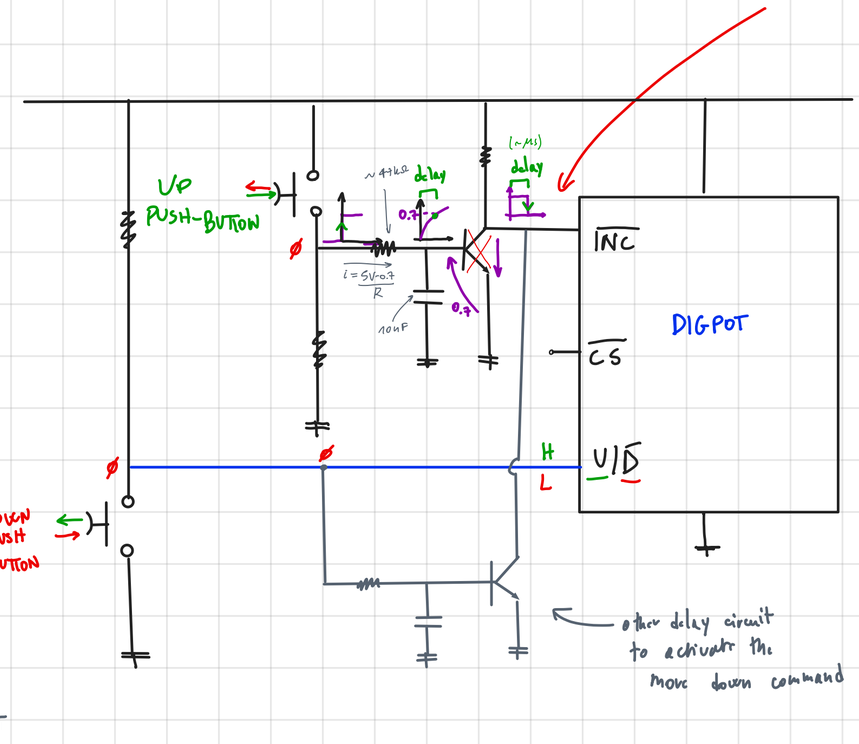
For the digpot we can consider a structure like this:



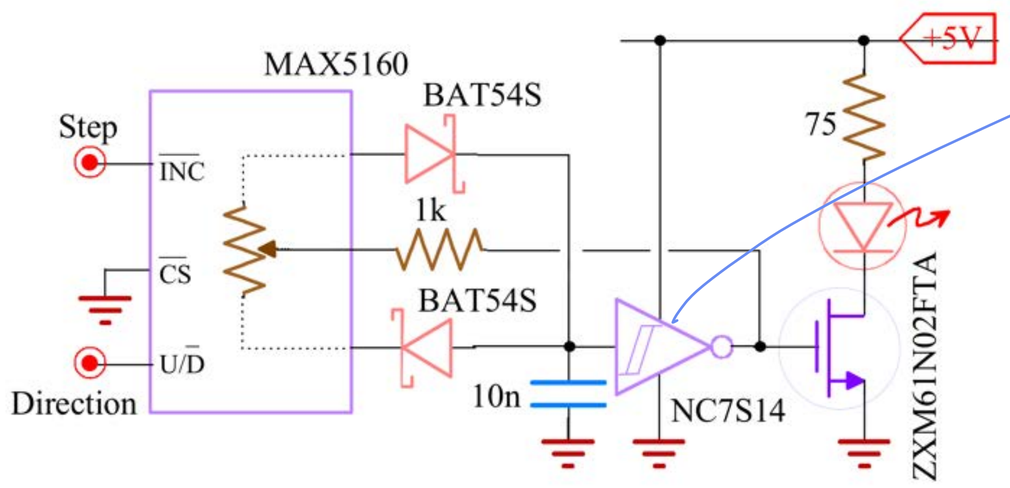
→ I want just one push-button to move ($\overline{INC}=1$) up ($U=1$) and one to move ($\overline{INC}=1$) down ($\overline{D}=1$)

(better way to represent push-buttons)

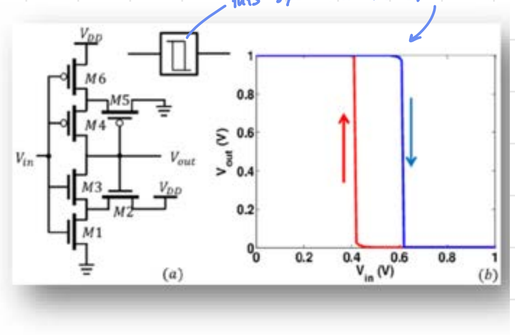




Possible applications



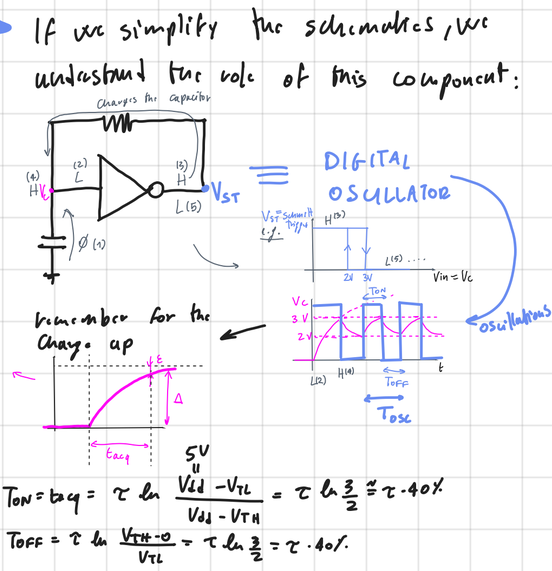
Schmitt trigger NOT gate:



Vin < VL → H
Vin = L → NOT

Vin > VH → L
Vin = H → NOT

If we simplify the schematics, we understand the role of this component:



In our case: $t_{acq} = \tau \ln \frac{\Delta}{\varepsilon}$ (lower threshold)

$\Delta = 5V - 2V$

$\varepsilon = 5V - 3V$

higher threshold →

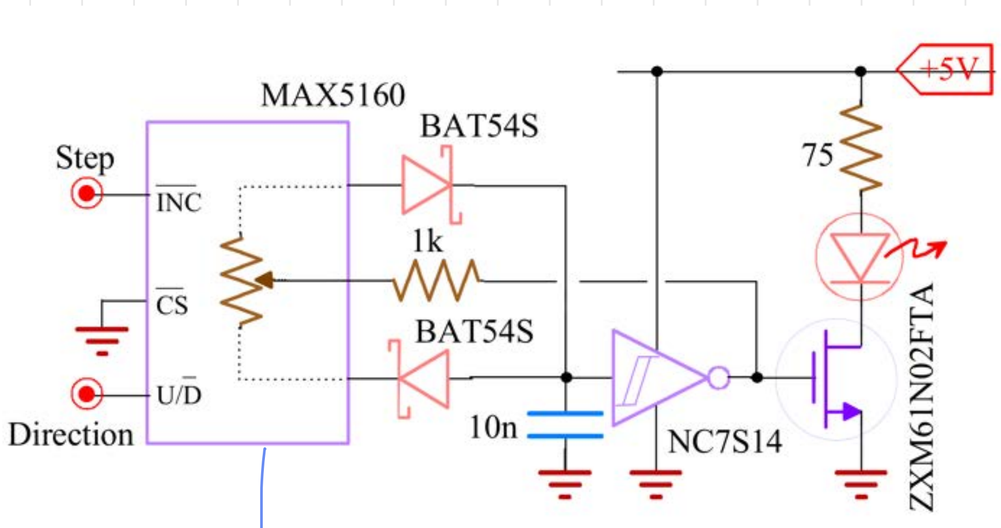
$T_{ON} = t_{acq} = \tau \ln \frac{V_{dd} - V_{TL}}{V_{dd} - V_{TH}} = \tau \ln \frac{3}{2} \cong \tau \cdot 40\%$

$T_{OFF} = \tau \ln \frac{V_{TH} - 0}{V_{TL}} = \tau \ln \frac{3}{2} = \tau \cdot 40\%$

$T_{osc} = 2.40\% \cdot \tau = 0.8\,\tau$

↳ $f_{osc} = \frac{1}{\tau} \cdot 1.25$

So focusing more on the whole circuit:

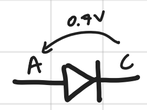


Dig Pot DATASHEET

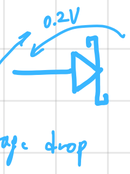
PART	TOP MARK	R (kΩ)
MAX5160NEUA	—	200
MAX5160MEUA	—	100
MAX5160LEUA	—	50
MAX5161NEZT	AAAC	200
MAX5161MEZT	AAAB	100
MAX5161LEZT	AAAA	50

c.f. 50kΩ impedance

Normal diodes: 0.9V

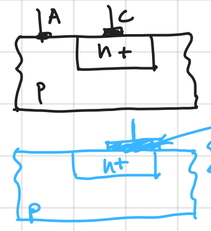


Schottky diodes: 0.2V

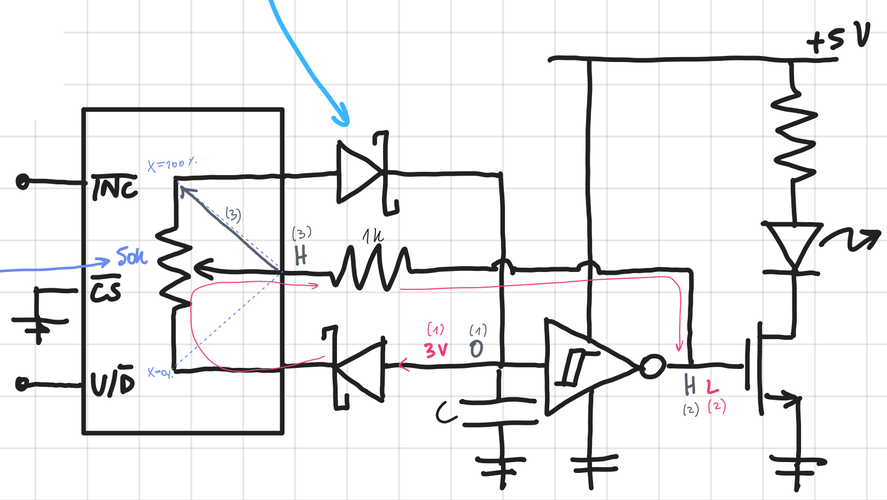


low voltage drop

used to reduce as much as possible the voltage drop on the diodes



In the connection bw a metal and a semiconductor it starts a reaction
↳ Shottky reaction



Seen by C

(2) → $R_{charge} = 1k + 50k(1-X) \cong 1k$ (if X = 100%)

(1) → $R_{discharge} = 1k + 50kX$

if C = 10nF

Computation of Tosc

$$T_{ON} = \tau_{charge} \cdot \ln\left(\frac{V_{dd} - 0.2}{V_{dd} - 3}\right) = C\,R_{charge}\,\frac{\ln\frac{2.8}{2}}{34\%}$$ (Schottky diodes)

$$T_{OFF} = \tau_{discharge} \ln\left(\frac{V_{TH} + 0.2}{V_{TL}}\right) = C\,R_{discharge}\,\frac{\ln\frac{2.8}{2}}{34\%}$$

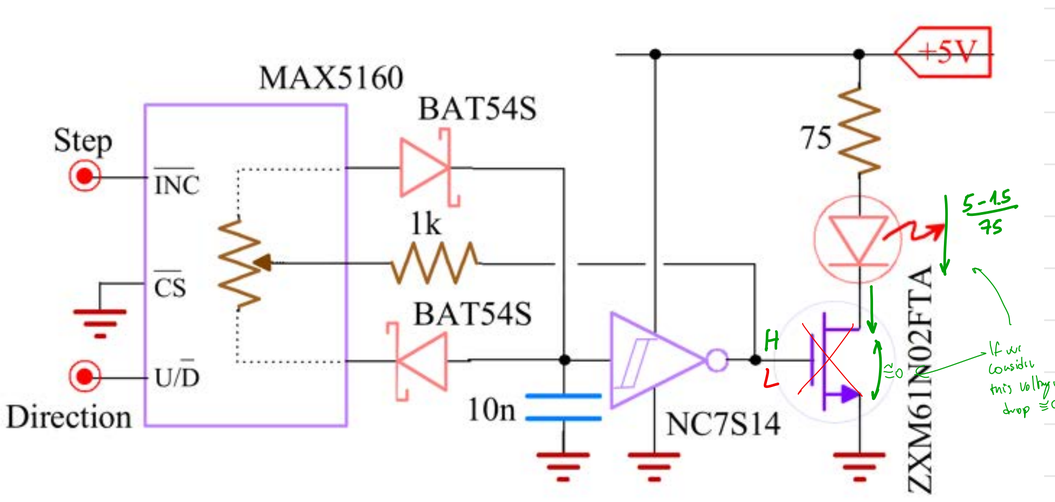
↕ now they're different since R change (dig pot)

$$T_{osc} = T_{ON} + T_{OFF} = 34\%\; C\,[R_{discharge} + R_{charge}] = 34\%\; C\,[1k + 50k]$$

↑ the oscillation doesn't depend on the variable (with the digpot cursor) resistance

↳ $f_{osc} = \frac{1}{T_{osc}}$

Then, in the circuit we see that the output of the Schmitt trigger NOT gate drives a MOS:



→ the LED will blink depending on the duty cycle of the oscillation

ES12_DAC

Electronic Systems

Introduction

(Book p. 550)

The D/A circuit has a set of n digital inputs (D_0–D_{n-1}) to which the binary value to be converted is transmitted, as shown in Fig. 7.1. From the pin 'Analog out', the result of the conversion is taken, which becomes available after a time fixed by the commutation of the signal 'Conv' (which determines the starting of the conversion). The voltage value corresponding to a binary input with all bits equal to '0' is set by the value V_{low} (in the following, we will assume that it is 0V) while that for a binary input of all '1' is determined by V_{ref} (often directly connected to V_{high}, usually equal to 5V). "Two quadrant" DACs allow the use of $V_{low}<0$, thus extending the output dynamics also to negative values symmetrically with respect to zero. If the reference can change over time, we speak of "Multiplying" DACs. For them, the analog output depends on the digital encoding, but its envelope is related to the instantaneous V_{ref} voltage, thus allowing easy modulation.

To prevent the spread of interference across the digital and analog parts, manufacturers provide two separate inputs for each power supply, denoted by V_{cc}, and for the grounds, marked with the initials A and D. Finally, ChipEnable/CE is the signal used to start the operation of the IC, thereby allowing us to keep it in stand-by condition in order to reduce the power consumption when not used.

(SoC) Conv

when the µC applies a rising edge on Conv pin the Din is converted to an Aout
↳ this is good because if µC continues to modify Din (for operations) also Aout will keep changing → we convert only when we want

Among the parameters that describe the performance of a DAC, there is the Full Scale Range (FSR) that indicates the maximum dynamics for V_{out} (basically, it is the full scale value). The output of the converter can assume 2^n different values between 0 and $(2^{-n}-1)\cdot V_{ref}$. The Most Significant Bit (MSB) brings at the output a contribution equal to $V_{ref}/2$ while the Least Significant Bit (LSB) affects it only by $V_{ref}/2^n$; this value also represents the highest resolution inherently achievable by the device. The Dynamic Range (DR) is the ratio between the maximum and minimum analog quantities that the converter can provide at its output; basically, it gives the number of discrete levels available at the output and is defined as $20\cdot\log_{10}2^n=6.02\cdot n$. For example, an 8bit DAC has 256 levels and a DR of 48dB. Note that, since the first available level is equal to $V_{low}=0V$, the maximum level will not be exactly V_{ref}, but it will be 1LSB lower in order to have 2^n levels. Note that the maximum value obtainable at the output is always "FSR–LSB"; thus, the number FSR is never represented since, of the 2^n combinations, one is used for the zero.

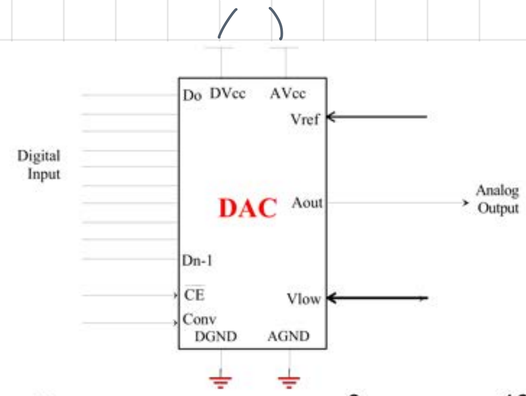
Tab. 7.1 shows the DAC resolution as the number of bits changes while Tab. 7.2 summarizes the different types of digital numbering.

explained later on in this lesson

integrated circuit

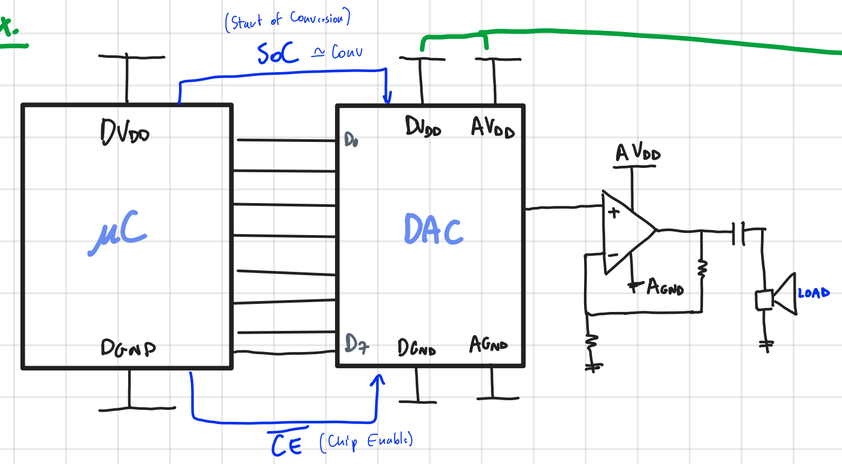
Separated in order to avoid that they disturb each other

Fig. 7.1

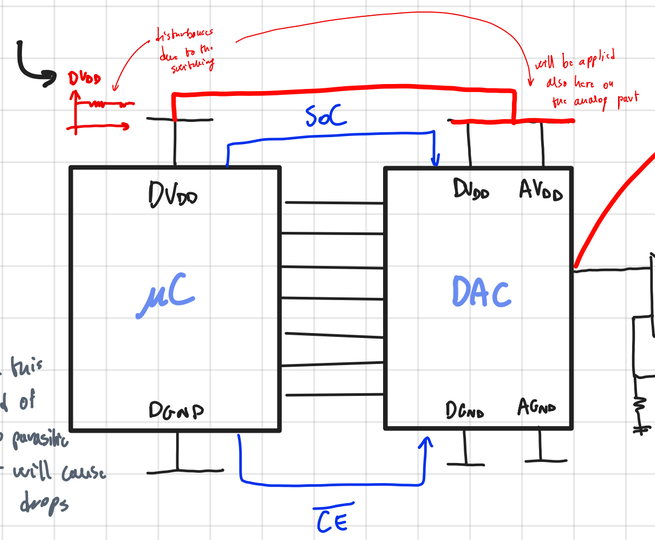


Number of bit:	n	8	16
Number of levels:	2^n	256	65536
FullScaleRange:		5V	5V
Resolution:	LeastSignificantBit = FSR/2^n	19.5mV	76µV
	1/2^n	3.9%	15ppm

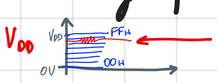
Ex.



We need 2 different P.S. because for ex. the digital part may induce disturbances, so if the P.S. is common this dist. will be applied also to the analog part.



usually V_{DD}

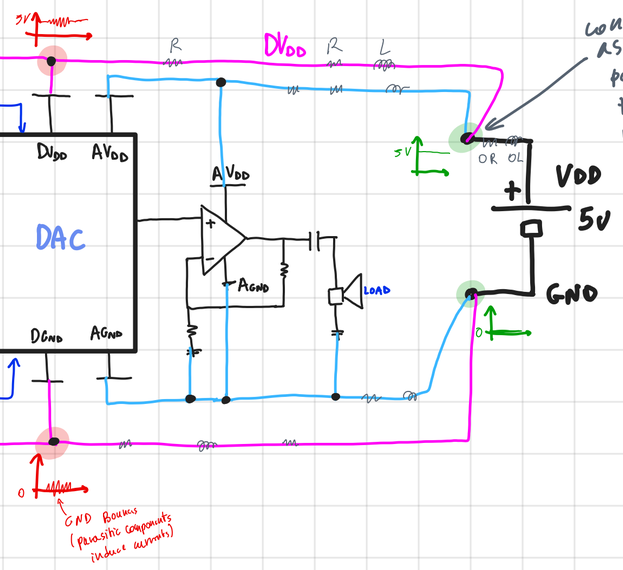


this disturbances will affect the analog output

$$A_{out}=\frac{D_{in}}{2^8}\cdot V_{ref}$$

keep the A & D P.S. separated

in practice



connect them as close as possible to the battery in this way any demand of currents (due to parasitic R or L) that will cause big voltage drops

↓

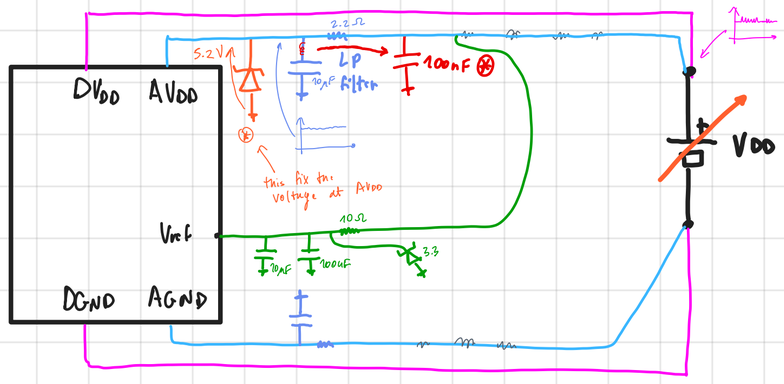
so if we connect at node ● we'll have big drops while at node ● with R≃0 L≃0 the drop will be minimal (same for A and GND case)

↳ Also V_{ref} should be as clean as possible and not directly connected to V_{DD}, we add some LP filter (RC) (it can also be used to use the same battery for A and D, but filtering out the dist.)

Note ⊛ Electrolitic capacitors (≃10µF) have a big parasitic inductance

↓

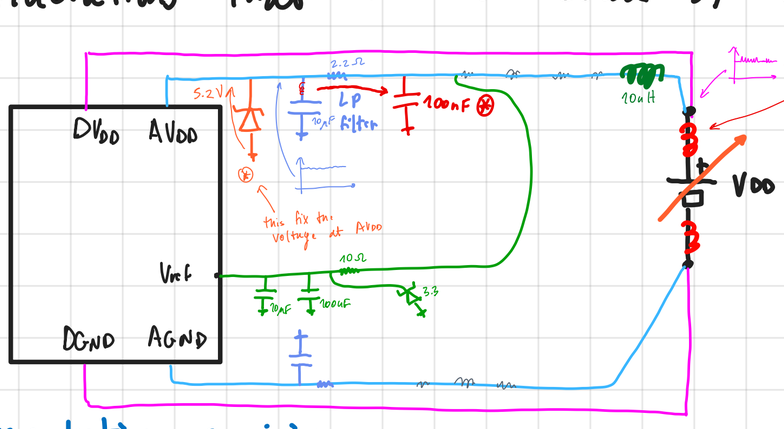
they'll be put in parallel with a tantalium capacitors (≃100 nF) that have very small parasitic inductance

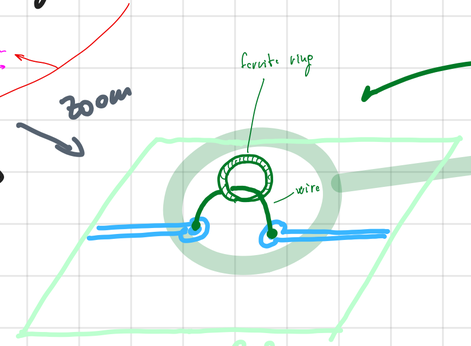


↳ the same considerations can be done for V_{ref}

↳ If we have the wish that the P.S. could be variable then we can put a Zener Diode ⊛

↳ Then to avoid the HF fluctuations that can be caused by big inductances on the P.S., a specific low value inductance (c.f. ≃10uH) is introduced





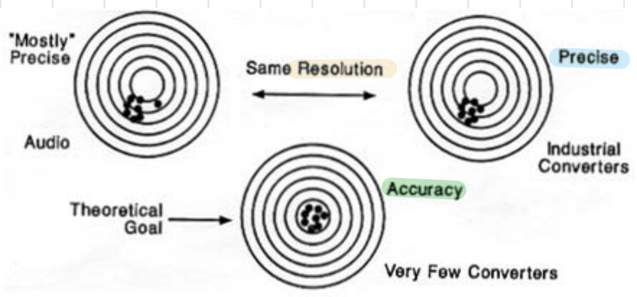
this introduces a very tiny inductance

Quality factors: accuracy, resolution, precision

Some of the main quality factors are accuracy, resolution, and precision. Since DACs are analog circuits, they suffer from inaccuracies due to mismatches between components, electronic noise, and thermal drifts, all of which can degrade the performance in terms of achievable resolution. An indicator of the DAC performance is the absolute **accuracy**, defined as the maximum error between the analog output value and the theoretical value expected. As for the other parameters of merit of the converter, even this is usually expressed in fractions of LSB. Ideally, it should never exceed ±0.5LSB, which is what would bring the only quantization noise.

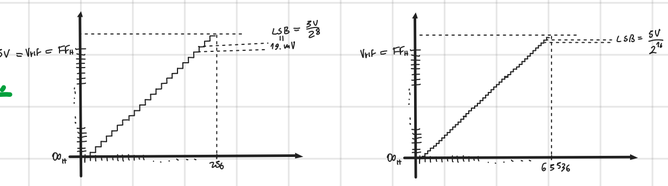
In addition to accuracy, even the other quality factors are used to define the performance of DACs, pictorially illustrated in Fig. 7.2. The **resolution** is indicated as the largest number of subdivisions in the output dynamics, which the converter is able to distinguish. Denoting n as the number of input bits, the DAC will have a resolution of 2^n levels between 0 and FSR. The **precision** is rather the quality index that grants an output value always equal to the input signal applied, i.e. it expresses the ability of the DAC to provide the same analog value (repeatable) as the same digital input applied. Note that, while accuracy implies precision, the opposite is not true.

Fig. 7.2

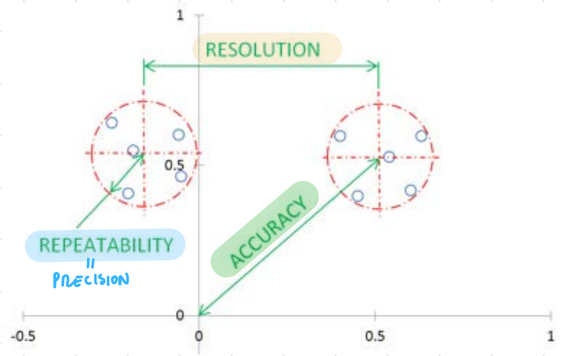


Resolution	subdivision of output dynamics that the converter is able to resolve
Precision	spread of output valuer, when the input is always the same
Accuracy	maximum error between output analog value and theoretical expected one

Ex.



8 bit 16 bit



Errors and non-linearities

(book p. 558)

The *offset error* is visible in the characteristic (Fig. 7.6) of a DAC as the distance of the first point of the decoding (all bits zero) from the ideal value equal to V_{low}; in other words, it represents the output voltage of the DAC when at its input is applied a code with all '0'. This error can be easily eliminated by subtracting it from the output voltage of the DAC and rigidly shifting all the points resulting from the conversion. This can be done by acting on the voltage V_{low} at the input of the DAC or, if the manufacturer provides it, by adjusting or compensating the potential of a dedicated pin of the integrated circuit.

The *gain error* is, instead, the error present in the code with all '1' with respect to the ideal value (FSR-LSB)=FSR·(2^n-1)/2^n, assessed after offset zeroing, also expressed in Volt or LSB. To correct this error, we need to check the slope of the DAC characteristic until it reaches the ideal value of 45°. Since, generally, it is not possible to directly operate on V_{ref} to obtain such a correction, the manufacturer provides a pin that internally acts on the DAC.

From the characteristic of Fig. 7.6, we can see how the offset error causes a shift in the characteristic upwards or downwards and how the gain error changes the slope of the real interpolating line (not 45°). Once compensated these errors, we have to address the non-linearity of the DAC, which is worse and which cannot be corrected by the user.

The integral or absolute non-linearity (INL) is represented by the maximum deviation between a real point of the characteristic of the DAC and the corresponding point on the ideal interpolation line, as shown in Fig. 7.7. The differential non-linearity (DNL) is, instead, the maximum difference of the jump in voltage between two adjacent results of the conversion and the corresponding theoretical value equal to one LSB.

Fig. 7.6

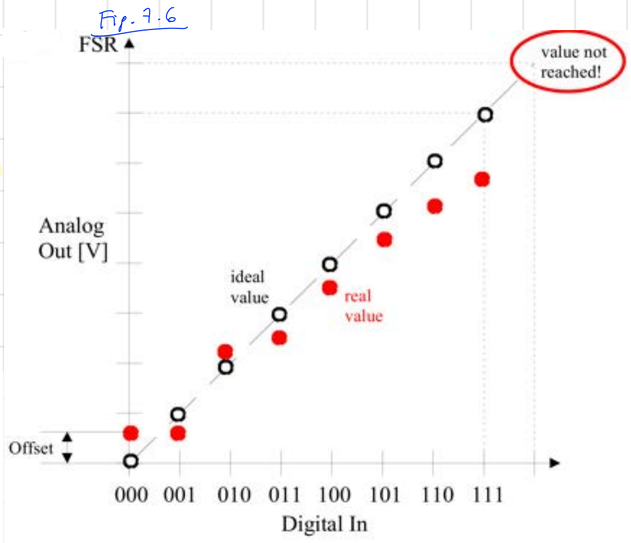
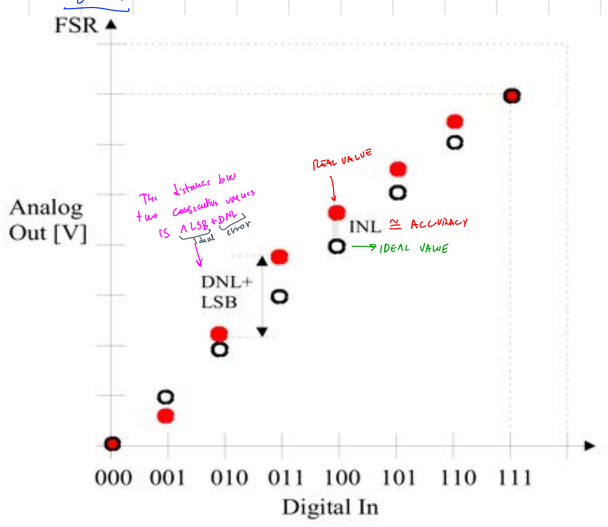


Fig. 7.7



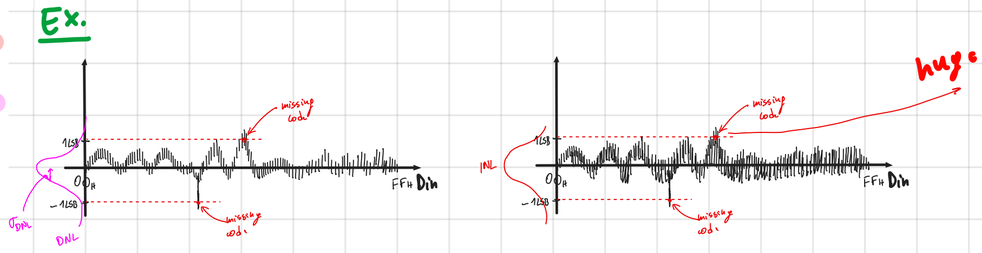
Offset: output when input is 000 } can be fixed

Gain: output when input is 111 }

INL: real-ideal distance

DNL: real-ideal step height

Ex.



Voltage-scaling DAC

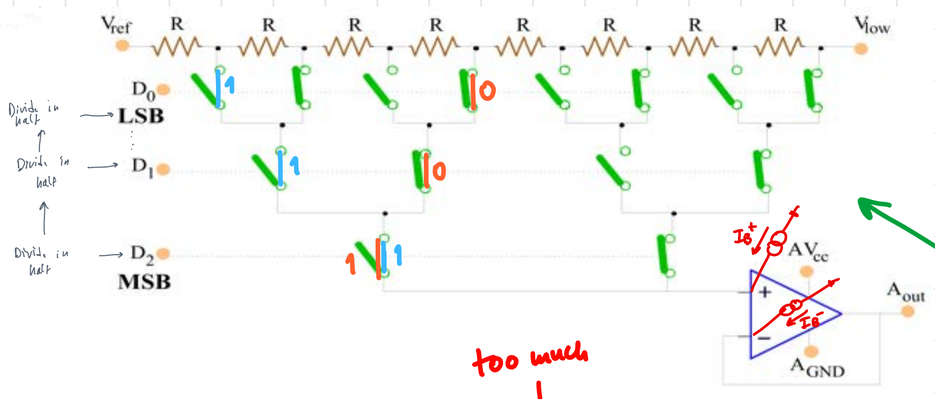
(Book p. 556)

This second simple DAC implementation, shown in Fig. 7.4, uses a resistive divider made of 2^n+1 identical to divide the entire FSR in 2^n voltage levels, each apart from their neighbor by 1LSB=(V_{ref}-V_{low})/2^n. The level corresponding to the numerical code at the input of the DAC is selected by properly closing some MOS switches controlled by the n bits of the digital word, from the least significant D_0 to the most significant D_{n-1}. The output voltage buffer ensures the possibility of providing a sufficient current to the load without worsening the accuracy of the conversion.

This circuit offers the advantage of ensuring an intrinsic monotonicity of the conversion and of needing identical resistance values, which is easily achievable with high accuracy.

However, the need for a large number of resistors and switches, equal to 2^n-1, implies that this architecture is only used for 8-bit (or less) DACs since the resistors take up a lot of space (particularly, if they are of high value). Also important are I_{bias} of the OpAmp and $I_{leakage}$ of the MOS switches that, depending on which switches are closed, change the characteristic in a non-uniform and unpredictable manner, introducing non-linearity.

Fig. 7.4



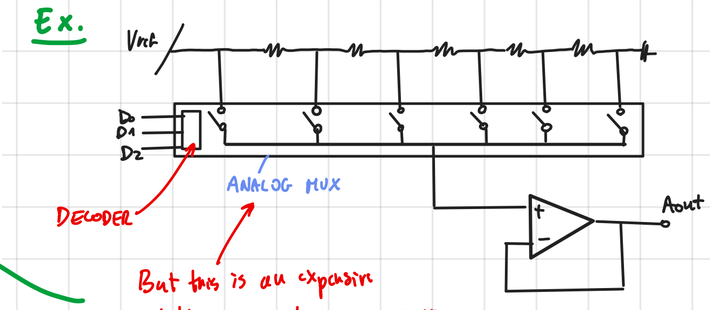
Cheaper and effective solution

too much

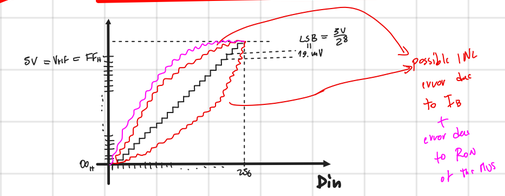
Components: 2^n resistors and 2^{n+1} pass-transistors (i.e. 2^{n+2} MOSFETs)
1 OpAmp

Advantages: easy scalability of resistors (all identical)
but OpAmp I_B current causes non-linearity issues

Ex.



Bias current error



Weighted-R DAC

(Book p. 555)

It is the simplest of converters and has the structure shown in Fig. 7.3. To determine the value of the output voltage V_{out}, we use a voltage divider with 2^n resistors (in relation to each other with an incremental ratio of the power of 2). We divide the amplitude of the signal into 2^n levels between the values V_{low} and V_{ref}, depending on how the switches are closed (or opened). The value of the output signal is given by (where V_{low}=GND):

$$V_{out} = V_{ref}\frac{R_f}{R}\left\{\frac{D_{n-1}}{1}+\frac{D_{n-2}}{2}+\ldots+\frac{D_0}{2^{n-1}}\right\} = V_{ref}\left\{\frac{D_{n-1}}{2}+\frac{D_{n-2}}{2^2}+\ldots+\frac{D_0}{2^n}\right\}$$

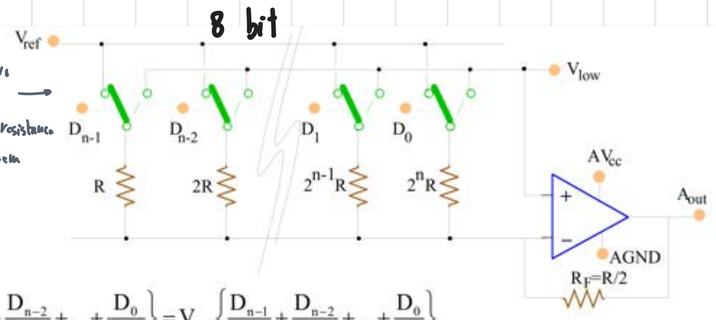
Note that the resistance at the terminal D_{n-1} gives a contribution at the output, which is equal to $1/2^{n-1}$ of the total. Furthermore, note some peculiarities of this implementation:

- variety in resistance values: starting with a low value R (a few kΩ), extremely high values 2^{n-1}·R (a few MΩ) are also necessary, which are hardly integrable except than at the expense of a large occupation of area;
- accuracies of the components: the resistance values are determined by the relation R=ρ·L/WT; generally, we have a fixed ρ/T (it depends on the manufacturing technology), and the different R values are obtained by changing the ratio L/W (at the layout level). Unfortunately, all these parameters have their tolerance levels;
- R_{ON}: each R_{ON} of the MOS switches weighs differently when in series with R rather than with 2^{n-1}·R;
- variable current consumption variable: dependent on the bits set to 1 since I_{bias} of the negative input will flow into different points of the resistive divider;
- reversed polarity: after the OpAmp voltage varies between 0V and –Vcc. Therefore, a negative power supply is needed or, at least, a charge pump internal to the integrated circuit, for the voltage reversing.

↑ lot of possible errors

Fig. 7.3

The idea is to have different currents depending on the resistance and then sum them

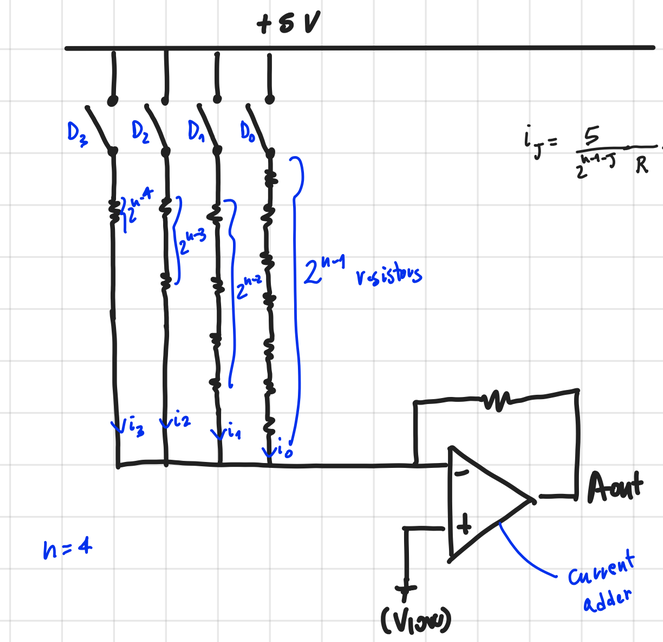


$$V_{out} = V_{ref}\frac{R_f}{R}\left\{\frac{D_{n-1}}{1}+\frac{D_{n-2}}{2}+\ldots+\frac{D_0}{2^{n-1}}\right\} = V_{ref}\left\{\frac{D_{n-1}}{2}+\frac{D_{n-2}}{2^2}+\ldots+\frac{D_0}{2^n}\right\}$$

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

Ex.

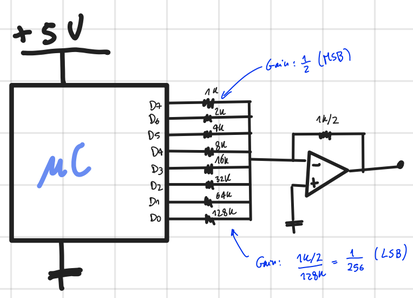


$i_J = \frac{5}{2^{n-J}R}\cdot D_J$ →

- $i_0 = \frac{5}{2^3 R} D_0$
- $i_1 = \frac{5}{2^2 \cdot R} D_1$
- $i_2 = \frac{5}{2^1 \cdot R} \cdot D_2$
- $i_3 = \frac{5}{2^0 \cdot R} D_3$

increasing current

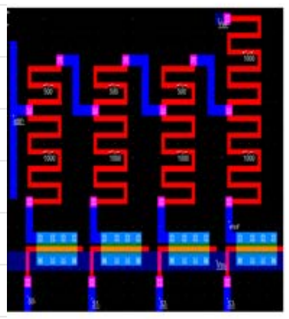
→ So coming from a μC

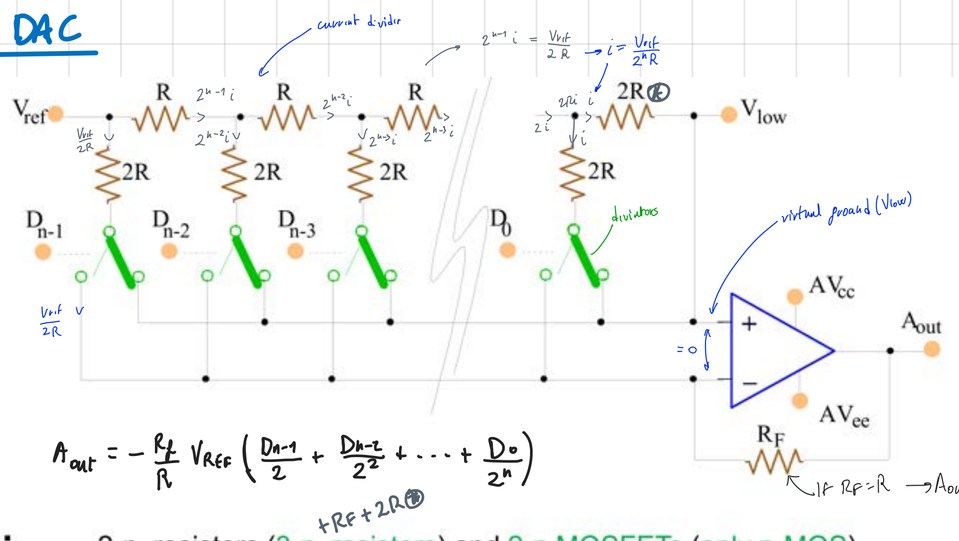


↳ Due to the fact that the offset changes with Din (due to MOS) in this type of config., because when switches are all open → $R_{eq} = \infty$, when they're all close the resistors are all in parallel $R_{eq} \simeq \frac{R}{2}$ → gain changes with Din

↳ So we use the config. in Fig. 7.3 → pmos and nmos in this way regardless of the switches position the R_{eq} is always the parallel of all resistances

Current-scaling DAC





$$A_{out} = -\frac{R_f}{R} V_{REF}\left(\frac{D_{n-1}}{2} + \frac{D_{n-2}}{2^2} + \dots + \frac{D_0}{2^n}\right)$$

+RF+2R

If $R_F = R \rightarrow A_{out} = -V_{REF}\left(\frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^n}\right)$ for $D_i = 1$ for $i = D_0, \dots, D_{n-1}$

e.g. $n=4 \rightarrow A_{out} = -V_{REF}\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16}\right) = -V_{REF}\frac{15}{16} - \left(V_{REF} - \frac{1}{16}\right)$ ✓ (LSB)

$\frac{3}{16}$, $\frac{7}{16}$

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 → R_{ON} plays no effect → const. error that can be compensated

Possible disadvantages:

- Bias current sum with the sum of the currents coming from the resistor causing an error (should be offset (my note))

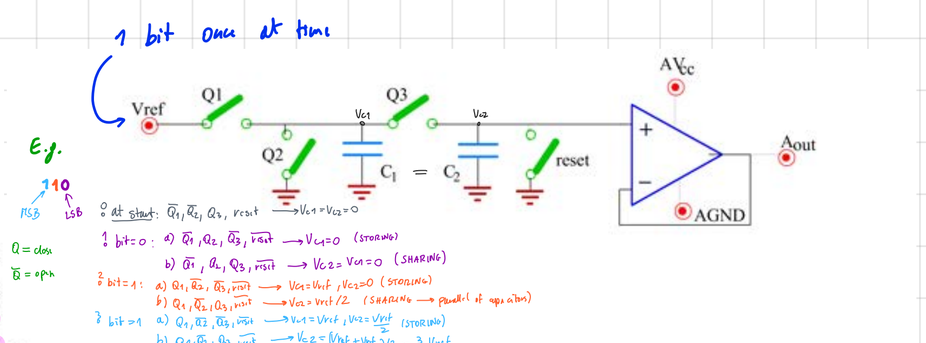
Serial-input DAC

(Book p. 557)

In this type of DAC, shown in Fig. 7.5, the individual bits are provided in sequence (serially), starting with the least significant bit, in order to sequentially control the closure of the "read" switch Q1 (if the bit is 1) or Q2 (if the bit is 0). After the closure of one of the two switches, the "share" switch Q3 is opened and closed for a short period before applying the same procedure to the next bit. Thus, if the bit is 1, the capacitor C1 is first charged to V_{ref}, and then half of the stored charge is transferred to C2 (chosen to be equal to C1), increasing its voltage. If, instead, the bit is 0, charge sharing occurs, leading to a reduction in the voltage across C2; that is, Q2 gets closed discharging C1, and the closure of Q3 halves the amount of charge present on C2. If we have a sequence of bits all equal to 1, the voltage across C2 will tend to increase by a quantity depending on the voltage value previously stored on C2 itself; otherwise, if they are equal to 0, it will continue to decrease.

As the bits come in, the two capacitors share the charge in such a way that if we get a 1 or a 0, the voltage tends to the final value. Finally, the first incoming bit is the LSB while the last (which suffers less charge transfers) is the MSB.

The system is very compact since it does not use resistors, but only MOSFET transistors, to realize switches and capacitors. This circuit has, however, the substantial disadvantage of being slower than the previous one because it requires a conversion time relatively high, which increases as the input number of bits increases:

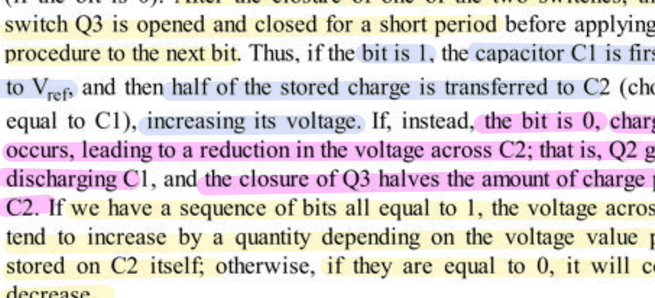


Components: only 2 capacitors and 5 MOSFETs (1 p-MOS, 2 n-MOS, 1 passtransistor)
1 OpAmp

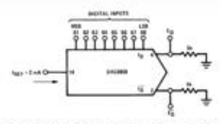
Advantages: extremely compact and easy, but needs serial input (bit by bit)
MSB first? How many bits?

Multiplying DAC

Multiplying DAC



Pin numbers represent the PDIP package. The SOIC package pin numbers differ from that of the PDIP package.



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_OmA	E_O	$\bar{E}_O$
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

TEXAS INSTRUMENTS — DAC0800, DAC0802

www.ti.com — SNAS536C – JUNE 1999 – REVISED FEBRUARY 2013

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.

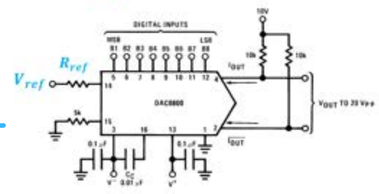
The DAC0800, DAC0802, DAC0800C and DAC0802C are a direct replacement for the DAC-08, DAC-08A, DAC-08C, and DAC-08H, respectively. For single supply operation, refer to AN-1525.

$$I_{out} = \frac{V_{ref}}{R_{ref}} \cdot \frac{D_{in}}{2^n}$$

$$I_{out} + \overline{I_{out}} = I_{FS} = \frac{V_{ref}}{R_{ref}} \cdot \frac{255}{256}$$

Here we can vary V_{ref} (in addition to Din) → $I_{out} = \frac{V_{ref}}{R_{ref}} \cdot \frac{D_{in}}{2^n}$

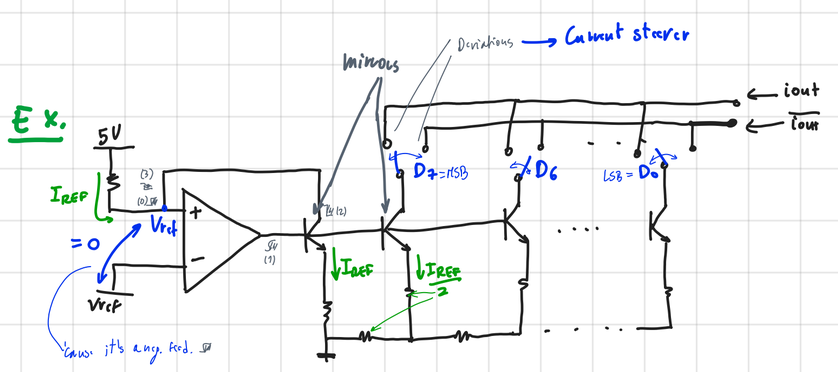
2 currents from current steering



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

Current steering



MULTIPLYING

variable X variable

$I_{out} = \frac{V_{ref}}{R_{ref}} \cdot \frac{D_{in}}{2^n}$

Dynamic Performances

(Book p. 564)

There are several types of dynamic errors: the intrinsic errors due to the fact that the analog-to-digital and digital-to-analog conversion introduce a quantization in the amplitude of the signals; those caused by the non-linear characteristic of the converter, which cause the appearance of unwanted harmonics; and errors due to the various electronic noises and interference, which are added to the useful signal. All of these will be described in detail.

Quantization error

As a sampled and quantized system, a DAC also suffers from the noise attributable to the quantization caused by the granularity of the digital code, shown in Fig. 7.11, i.e. the finite number of analog levels provided at the output. In modulus, this error is less than ½LSB. If the signal is sufficiently variable to several levels, the quantization error (deterministic) can be considered, with good approximation, a white noise not correlated with the useful signal, having a uniform distribution of probabilities, between ±½LSB. The variance of this error, with dimensions of Volt^2, is calculated by using the following expression:

$$\sigma^2 = \int_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} \varepsilon_q^2 \cdot \frac{1}{LSB} \cdot d\varepsilon_q = \frac{1}{LSB}\left[\frac{\varepsilon_q^3}{3}\right]_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} = \frac{1}{LSB}\cdot\frac{2}{3}\cdot\frac{LSB^3}{8} = \frac{LSB^2}{12}$$

The variance of an ergodic statistical process also represents its power (in our case, the quantization noise power); the square root of the variance is better known as the effective value (r.m.s., root mean square).

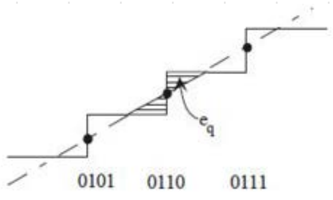


Fig. 7.11: Quantization error.

Signal to Noise Ratio (SNR)

The maximum Signal/Noise ratio of a converter is that achievable by applying a sinusoidal signal with the maximum amplitude allowed (peak to peak value equal to the FSR) with the least possible noise superimposed, i.e. the only quantization noise. Considering applying, at the input of the DAC, a signal that covers the whole permitted dynamics, for example a digital sine wave with amplitude of FSR/2, and compare it with the only quantization noise, we will obtain the following signal to noise ratio:

$$SNR|_{theoretical} = 10\log\frac{\text{Signal Power}}{\text{Quantization Noise Power}} = 10\log\frac{\left(\frac{FSR/2}{\sqrt{2}}\right)^2}{\frac{LSB^2}{12}} = 10\log\left(\frac{12}{8}\cdot 2^{2n}\right) = 6.02\cdot n + 1.76 \quad [dB]$$

where $(FSR/2\cdot\sqrt{2})^2$ is the useful power of a sinusoidal signal with maximum amplitude equal to V_p=FSR/2, and $V_{al\ eff}$=$V_p/\sqrt{2}$e 6.02·n represents the maximum dynamics.

It can be derived from the previous equation that each bit improves the signal to noise ratio of about 6dB. The meaning of the SNR is simple: given a maximum amplitude analog sine wave at the output, it describes the "granularity" of the quantized sine wave, comparing it to a corresponding "noisy" sine wave. The rms nose that provides "visually" the same resolution is just equal to the square root of $LSB^2/12$, i.e. $LSB/\sqrt{12} \approx LSB/3.5$. Thus, an 8-bit DAC has $SNR_{theoretical}$=6.02·8+1.76=50dB, therefore, for a sine wave with the maximum allowable peak to peak amplitude (equal to the FSR), it will have a quantization error that results in a "noise" which is equivalent to a real white noise, with a power equal to -50dB=1/100'000 with respect to the sine wave, i.e. with an rms value equal to 0.22% of the peak value (FSR/2) of the sine wave.

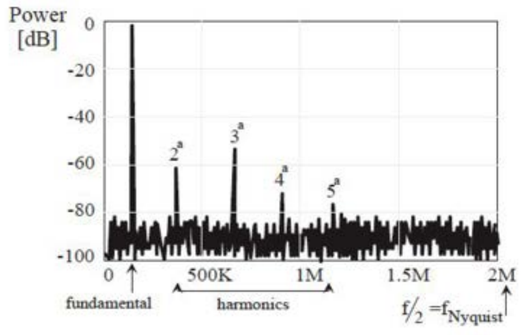
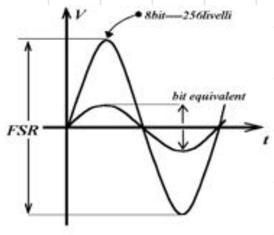


Fig. 7.12: Spectrum of the output signal of the DAC: we can see the sine wave of interest, the spurious harmonics, and the background noise; each histogram has a width that is called bin width.



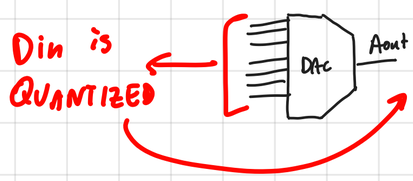
Dynamic range: $D = 20\cdot\log\frac{FSR}{LSB} = 20\cdot\log 2^n = 6.02\cdot n$

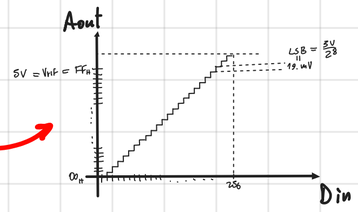
Quantization Noise: $\sigma^2 = \int_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} \varepsilon_q^2 \cdot \frac{1}{LSB} \cdot d\varepsilon_q = \frac{1}{LSB}\left[\frac{\varepsilon_q^3}{3}\right]_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} = \frac{1}{LSB}\cdot\frac{2}{3}\cdot\frac{LSB^3}{8} = \frac{LSB^2}{12}$

Ideal Signal-to-Noise Ratio: $SNR|_{max} = \frac{\text{max signal power}}{\text{min noise, just quantization error}}\Big|_{dB} = 10\cdot\log\frac{(FSR/2\sqrt{2})^2}{(LSB^2/12)} = 6.02\cdot n + 1.76$

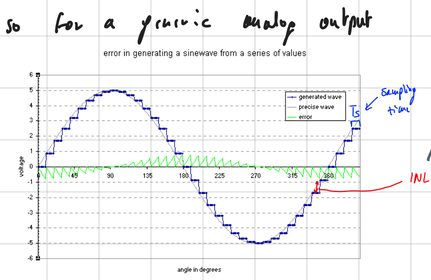
EffectiveNumberOfBit: $n_{eff} = \frac{SNR - 1.76dB}{6.02dB}$

Ex.





So for a generic analog output



If INL is very low it means that the quantization error will be at most $\pm \frac{LSB}{2}$ (random)

→ We can describe the quantization error as if it were Noise

(peak-to-peak)

↳ $\varepsilon_q = \pm \frac{1}{2} LSB$

↳ $\sigma^2 = \frac{LSB^2}{12}$ ← computed via power formula seen in the theory before

We can consider the quantized signal = ideal signal + noise

$$SNR = \frac{S}{N} = \frac{Power_{ideal\ signal}}{Power_{noise}}$$

$$Power_{ideal\ signal} = \left(\frac{V_p}{\sqrt{2}}\right)^2$$

$$Power_{quantized\ signal} = Power_{noise} = \frac{LSB^2}{12}$$

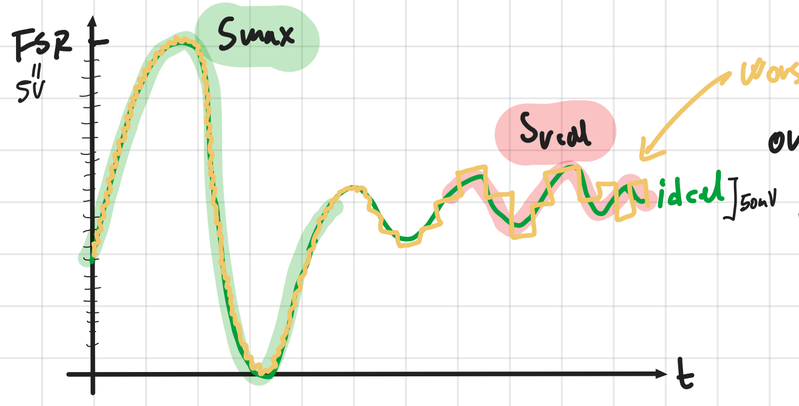
↳ In particular: SNR_{ideal} (just considering quantization noise): $SNR_{ideal} = SNR_{max} = \frac{S_{max}}{N_{min}} = \frac{\left(\frac{FSR}{2\cdot\sqrt{2}}\right)^2}{N_{quantized}} = \frac{\left(\frac{FSR}{2\cdot\sqrt{2}}\right)^2}{\frac{FSR^2}{12\cdot 2^{2n}}} = \frac{12\cdot 2^{2n}}{4\cdot 2} = \frac{3}{2}2^{2n}$

↳ In dB: $SNR_{ideal} = 10\,log_{10}\left[\frac{3}{2}2^{2n}\right] = 10\,log\frac{3}{2} + 10n\,log 2 = SNR_{max} = 6.02n + 1.76$

↳ e.g. $n = 8$ bit → $SNR_{ideal} = 50$ dB
$n = 16$ bit → $SNR_{ideal} = 98$ dB

if $n\uparrow$ quantization error $\downarrow$

↳ So if we have



worse quantization on Sreal because we didn't use the full FSR so we have not the ideal case of Smax

we didn't amplify enough to cover the FSR

e.g. n = 8 bits ΔN = 15dB

$\Delta S = \frac{50mV}{5V} = 20dB$

$SNR_{ideal} = 6.02\cdot n + 1.76 = 50dB$

$SNR_{theor.} = SNR_{ideal} - \Delta S = 50dB - 20dB = 30dB$

$SNR_{real} = SNR_{ideal} - \Delta S - \Delta N = 15dB$ ← worse

↳ $SNR_{theoretical} = \frac{S_{real}}{N_{min}} = SNR_{ideal} - \Delta S$ [in dB] → $\Delta S = \frac{S_{max}}{S_{real}}$

↳ $SNR_{real} = \frac{S_{real}}{N_{real}} = \frac{S_{ideal}\cdot\Delta S}{N_{quant} + \Delta N} = SNR_{ideal} - \Delta S - \Delta N$

noise quantization + EXTRA NOISE

↳ Equivalent Number Of Bits (ENOB): $SNR_{ideal} = 6.02\cdot ENOB + 1.76$ → $ENOB = \frac{SNR_{ideal} - 1.76}{6}$

↳ For SNR_{real} → $ENOB = \frac{SNR_{real} - 1.76}{6.02} = 2.3$ bit

8 bit

Spectral Performance

Signal to Noise and Distorsion Ratio (SiNAD)

A figure of merit to recognize the performance of a DAC converter is the SNDR or SiNAD, *Signal to (Noise+Distorsion) Ratio*, which represents the ratio between the power of the sinusoidal signal at the input and that obtained by adding the power of the quantization noise and of all the spurious harmonics generated as a result of the distortion produced by the DAC itself.

The purity of the input sine wave (digital) can be detected by performing the FFT (Fast Fourier Transform) while the spectrum of the output (analog) can be measured with a spectrum analyzer, as the one shown in Fig. 7.12. The manufacturer, to specify the performance of the converter, often provides a graph that shows the power distribution vs. frequency (limited to the Nyquist bandwidth) when the input is a sine wave at a given frequency. To be more precise, the graph provides the power of each histogram and not the power spectral density; this difference will be crucial in the reasoning below. This spectrum is the result of the analysis of the output signal of a DAC with a spectrum analyzer (a device that performs the FFT) with a filter of 1Hz, which plots the power for all frequencies.

In Fig. 7.12, we can see the power peaks of the input sine wave (at the fundamental frequency) and those due to the various harmonics resulting from the distortion (not present in an ideal DAC). The powers are normalized to 0dB; then, it is more correct to speak of dBc, where the subscript "c" states that it refers to the "carrier".

The value of about –80dB, where the average of the remaining histograms is, is called the *Noise-floor* and is essentially a white noise (constant over $f_s/2=f_c=f_{Nyquist}$). This is due to both the quantization noise power, equal to $LSB^2/12$, uniformly distributed over the various frequencies, and the other noises (thermal and shot) inside the converter. It is important to note that the noise-floor does not exactly specify the signal/noise ratio of the DAC, i.e. for a 10bit DAC (hence with $SNR_{theoretical}$=62dB), the noise-floor is not placed 62dB lower than the fundamental.

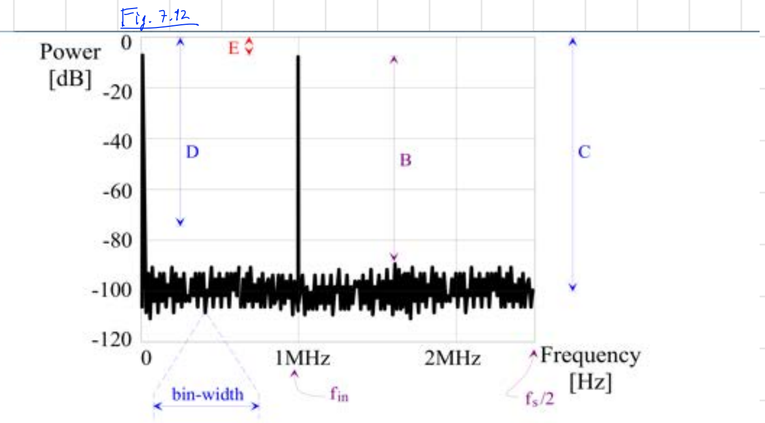
In fact, just because the spectra provide the power (and not the spectral density) contained within each histogram, the value of the Noise-floor will depend on the number N of samples used to compute the FFT. In fact, starting from N values in the time domain, the Fourier transform DFT calculates the same number in the frequency domain, distributed between 0 and f_S. Each of the N/2 samples calculated at frequencies between 0 and $f_S/2$ (Nyquist frequency) contains information related to the power that the signal takes in a histogram of width $(f_S/2)/(N/2)=f_S/N$ between $f_s(k-1)/N$ and f_sk/N (k=0… N/2-1), called *Bin Width* (width in frequency of each histogram). Dealing with a stream of digital samples sampled at f_S=500ksps and desiring to calculate the spectrum with a precision in frequency (bin-width in fact) of 0.2Hz, N=500,000/0.2=2,500,000 histograms in frequency, as many samples in time consequently, would be required. Since these follow each other every $1/f_S$=2µs, it will be necessary to acquire the signal for a period of time equal to $N\cdot 1/f_S$=2,500,000·2µs=5s.

In conclusion, in the case of a sine wave centered at a well precise frequency, the amplitude of the corresponding histogram would always remain constant since as N increases, the power of the sine wave remains the same and always centered in a single histogram. Instead, in the case of white noise, the increase in N decreases the value of the power contained in each histogram by the same amount. For this reason, the Noise-floor level is given by the following expression:

$$\text{Noise Floor} = -\{6.02\cdot n + 1.76 + 10\cdot\log N\} = -SNR\big|_{theoretical} - 10\cdot\log N$$

For every doubling of N, there is an increase of 3dB in the noise noise-floor at the same actual performance. This suggests that, with a higher number of samples, the level of noise in power goes down because the description of the signal improves.

Fig. 7.12



Total noise (only quantization error): $\sigma_q^2 = \frac{LSB^2}{12}$

Bin-width: $\text{bin} - \text{width} = \frac{f_s}{N_{samples}}$ **Noise in each bin:** $\sigma_q^2/N_{samples} = \frac{LSB^2}{12}/N_{samples}$

... hence $\text{Noise Floor} = -\left\{6.02\cdot n + 1.76 + 10\cdot\log\left(\frac{N_{samples}}{2}\right)\right\} = -SNR\big|_{teorico} - 10\cdot\log\left(\frac{N_{samples}}{2}\right)$

Harmonic Distortion → (look also at prob EX07-1 for an example)

The parameter THD (Total Harmonic Distortion) provides the ratio between the rms power of the different spurious harmonics generated by the DAC and that of the useful signal (the fundamental). The calculation should consider at least the first nine harmonics, although manufacturers consider only the first five. In the following expression, we assume that the power of the useful signal is normalized to 0dB:

$$THD = \frac{\sqrt{\sum_{i=2}^{9} V_i^2}}{V_1} = 20\cdot\log\sqrt{10^{\left(\frac{2^{a}\text{ harmonic (dB)}}{20}\right)^2} + 10^{\left(\frac{3^{a}\text{ harmonic (dB)}}{20}\right)^2} + \dots} \quad [dB]$$

V_1^2 is the power of the input fundamental while V_i^2 is the power of the i-th harmonic.

Finally, the parameter IMD (Inter Modulation Distortion) takes into account the intermodulation distortion that is created when two or more sine waves enter in a non-linear DAC. Because of non-linearity, the conversion of the sum of two sine waves at different frequencies f_1 and f_2 determines the appearance of spurious sinusoids at frequencies different from that of the fundamentals at the input (f_1 and f_2). Such a relationship gives us an indication of how non-linear the DAC is. In fact, a non-linear characteristic generates some spurious products (they are the spurious components) in the sum and difference frequencies of any linear combination of the original ones, i.e. at frequencies $m\cdot f_1 \pm n\cdot f_2$. The IMD is defined as the ratio between the power of these 'beats' and the useful power of the input signals:

$$IMD \equiv \frac{\sqrt{\sum nf_1 \pm mf}}{\sqrt{v_1^2 + v_2^2}}$$

Note that the second order components are f_1+f_2 and f_1-f_2, which does not mean that the signal is $V_1\cdot\sin(2\pi f_1 t)+V_2\cdot\sin(2\pi f_2 t)$, but just $V\cdot\sin[2\pi(f_1+f_2)t]$.

Signal to Noise Ratio: **real** ratio between signal and noise (NO distortion, NO disturbances)

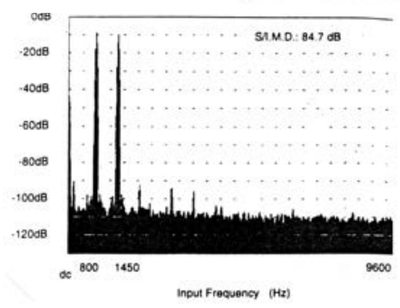
$$SiNAD = \frac{\text{power of real signal}}{\text{power of real white noise}}\Big|_{dB} = 10\cdot\log\frac{\left(Vin_{peak}/2\sqrt{2}\right)^2}{\sum\text{powers of bins at the level of NF}}$$

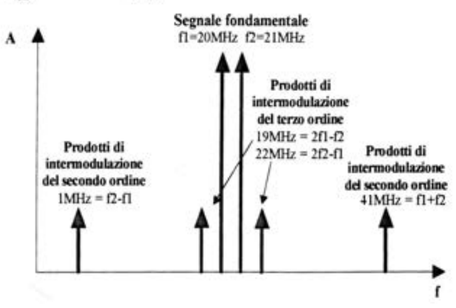
Signal to Noise And Distorsion: real ratio between signal and noise+distortion(harmonics)+disturbances

$$SiNAD = \frac{\text{power of useful signal}}{\text{total power of real noise and harmonics}}\Big|_{dB} = 10\cdot\log\frac{\left(Vin_{peak}/2\sqrt{2}\right)^2}{\sum\text{powers of ALL bins}}$$

Total Harmonic Distorsion: ratio between all harmonics and useful signal (NO disturbances)

$$THD = \frac{\text{power of all harmonics}}{\text{power of the useful signal}}\Big|_{dB} = 10\cdot\log\frac{\sum_{k=2}^{9} P_{f=k\cdot fin}}{P_{f=fin}}$$

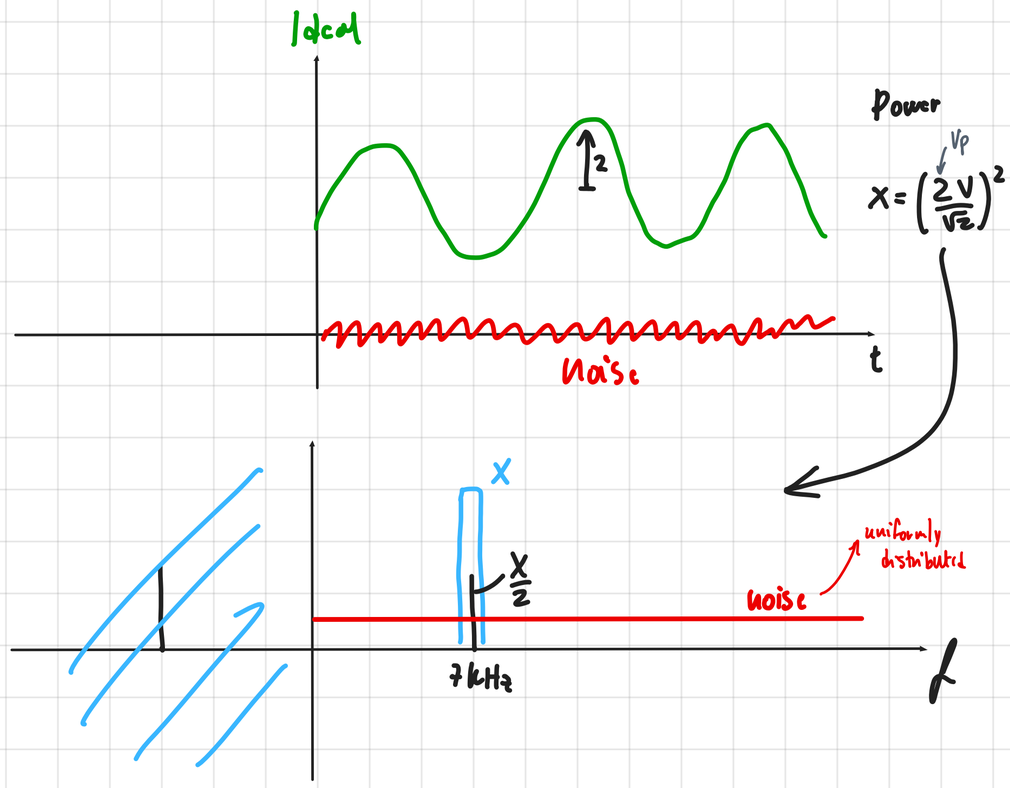


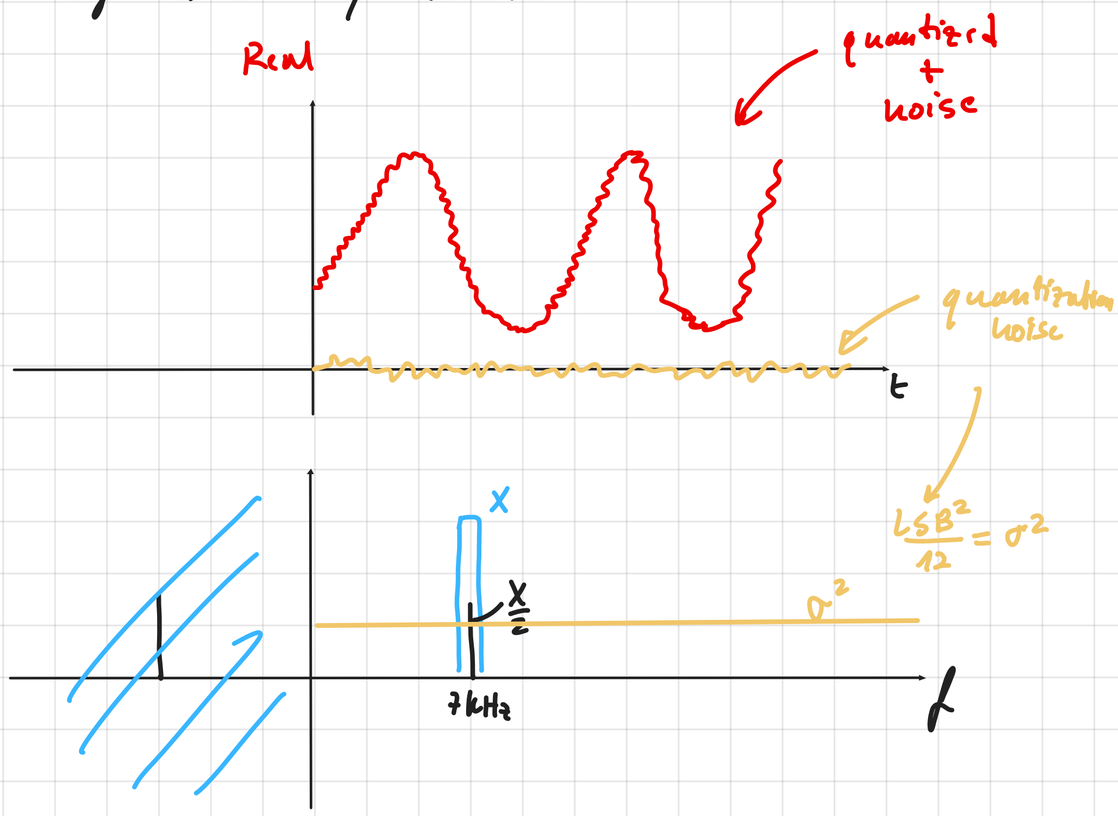


Intermodulation Distorsion: ratio among all intermodulation products and 2 useful signals

$$IMD = \frac{\text{power of all intermodulation products}}{\text{power of the 2 useful signals at 2 different frequencies}}\Big|_{dB} = \frac{\sqrt{\sum nf_1 \pm mf_2}}{\sqrt{v_1^2 + v_2^2}}$$

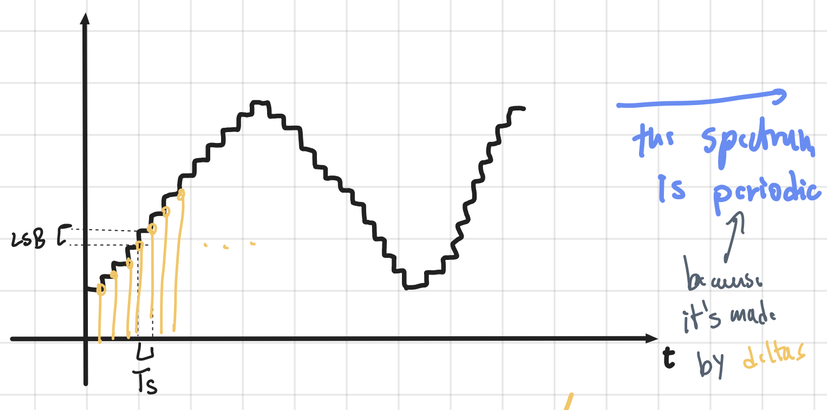
Ex. Instead of analyzing the signal in the time domain we study the signals in freq. domain





↳ we consider only the pos. spectrum

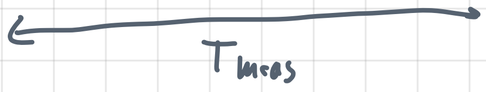
↳ In particular we know:



the spectrum is periodic

↑ because it's made by deltas

$N_{samples} = \frac{T_{meas}}{T_s} = T_{meas} \cdot f_s$

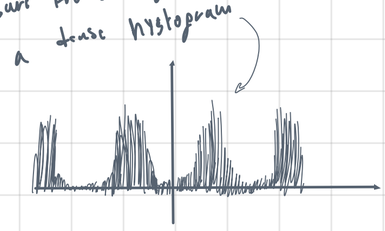


↳ we consider only the pos. spectrum

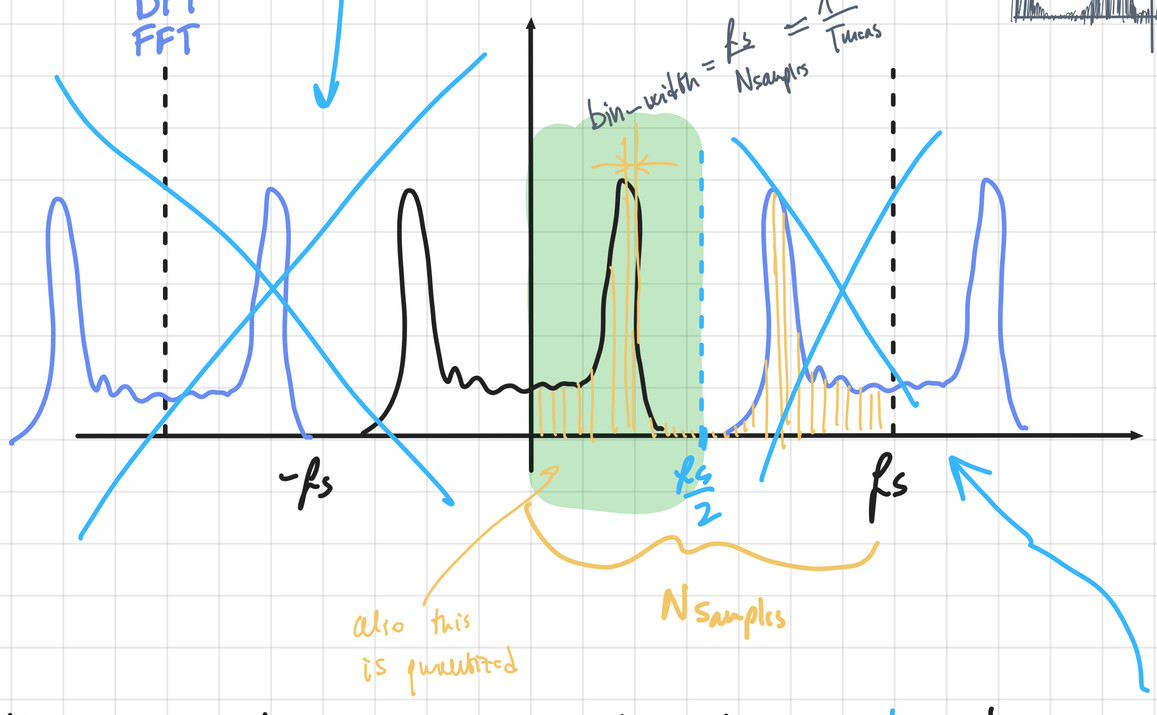
symmetric

DFT FFT

If we measure for a long period of time we get a tone histogram



$\text{bin-width} = \frac{f_s}{N_{samples}} = \frac{1}{T_{meas}}$



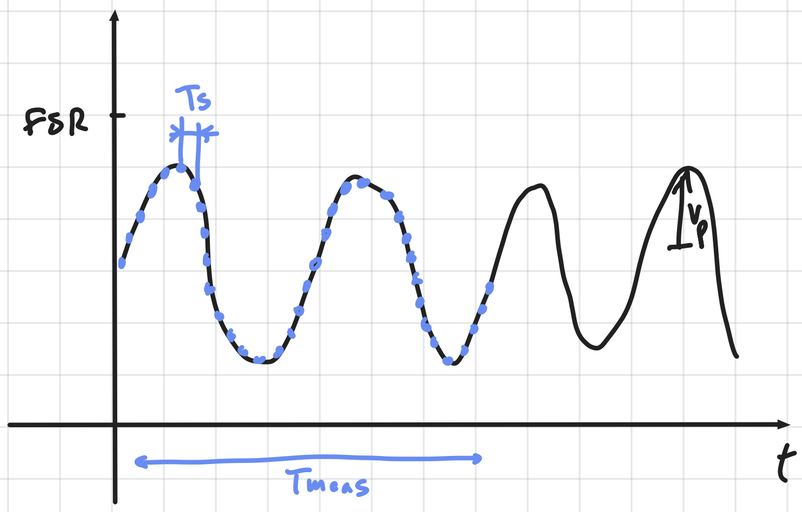
also this is quantized

$N_{samples}$

↳ If the signal is real the spectrum is symmetric also around $\frac{f_s}{2}$

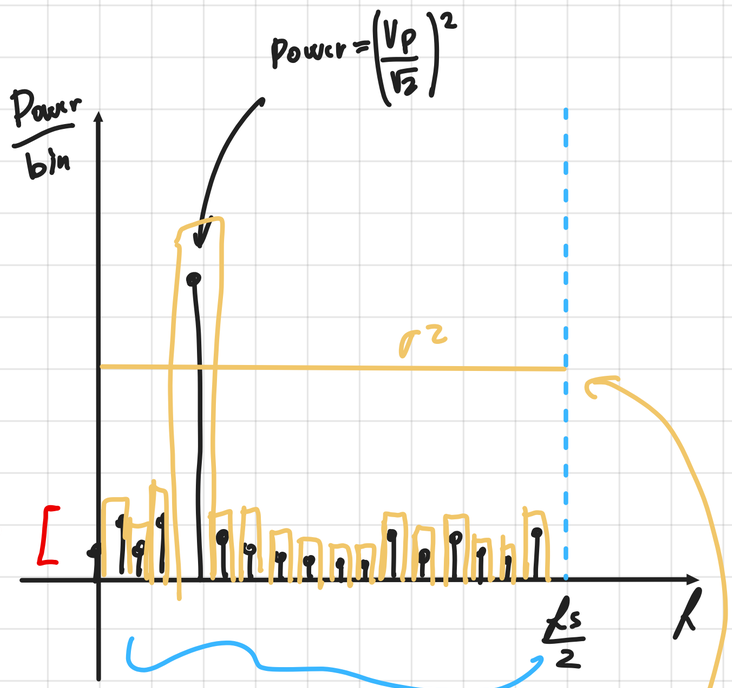
↳ we study the spectrum just in this area $\in [0, \frac{f_s}{2}]$

↳ So given a signal in time domain:



↳ $N_{samples} = \frac{T_{meas}}{T_s}$

FFT



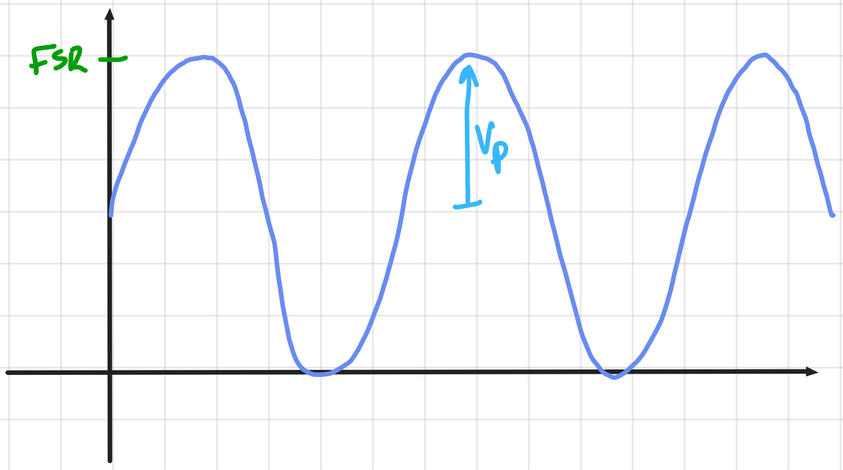
~ height of these samples

Total noise due to quantization $= \sigma^2 = \frac{LSB^2}{12}$

the quantization error spreads uniformly over the bins

$\text{Noise floor} = \frac{\frac{LSB^2}{12}}{\frac{N_{samples}}{2}}$ ← So the height of noise in the spectrum will be (devided for the # bins)

↳ Now consider the maximum possible signal (in the FSR):



→ $S_{max} = \frac{FSR}{2} = V_p$

S_{max}^2 = power of the max signal $= \left(\frac{V_p}{\sqrt{2}}\right)^2 = \left(\frac{FSR}{2\sqrt{2}}\right)^2$

↳ The idea now is to perform a kinda of normalization of everything

We put: $S_{max}^2 = 0\,dB_c$

↑ carrier

→ It means that $0dB_c$ is the typical power of a signal whose amplitude is the max achievable in the FSR of the converter

↳ So knowing that:

$$SNR_{ideal} = \frac{S_{max}}{N_{min}} \overset{(\cdot)^2}{=} \frac{S_{max}^2}{\frac{LSB^2}{12}} = 6.02 \cdot n + 1.76 \quad \rightarrow \quad N_{min}^2 = \frac{LSB^2}{12} = \frac{S_{max}^2}{SNR_{ideal}} = S_{max}^2 - (6.02\,n + 1.76)$$

↑ $0dB_c$

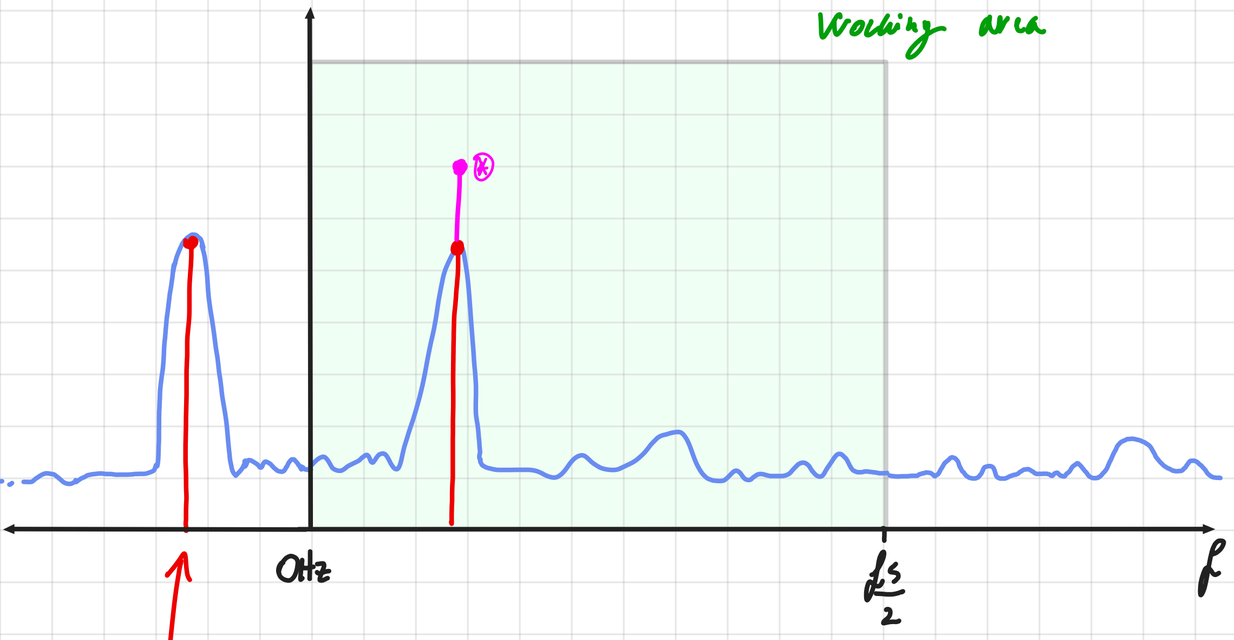
↳ $N_{min}^2 = \frac{LSB^2}{12} = -(SNR_{ideal})_{dB}$

ex. $FSR = 5V$, $n = 8$ bit

$SNR_{ideal} = 6.02n + 1.76 = 50\,dB$ $\quad S_{max_p} = \frac{5}{2} = 2.5$ $\quad S_{max_{rms}} = \frac{2.5}{\sqrt{2}}$ $\quad S_{max_{pow}}^2 = \left(\frac{2.5}{\sqrt{2}}\right)^2 = 3.125\,V^2$

I define it as $0\,dB_c$

↳ $N_{min}^2 = \frac{LSB^2}{12} = -SNR_{ideal} = -50\,dB_c$ → $SNR = \frac{S}{N} = \frac{0\,dB_c}{-50\,dB_c} = (0 - (-50))_{dB} = 50\,dB$

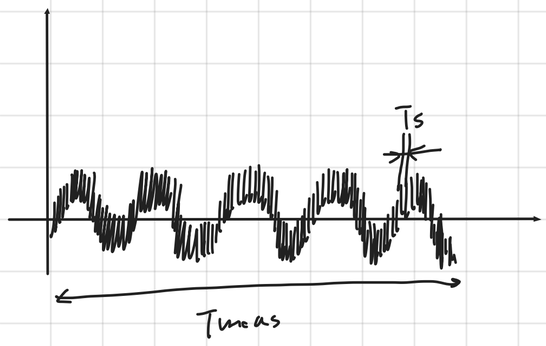


To consider also this power, but still working in the positive area we have to double the height of the power of that sinusoid (*)

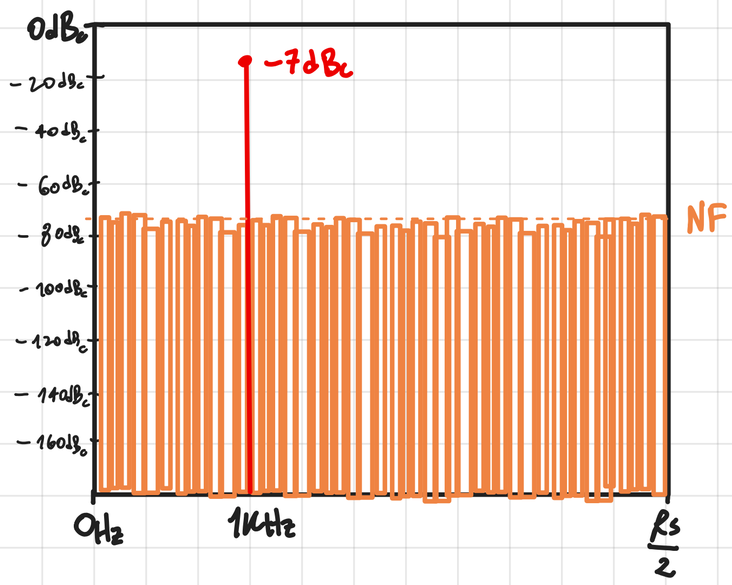
taking into account all the considerations we should plot

$f_s = \frac{1}{T_s}$

$N_{sample} = \frac{T_{meas}}{T_s}$



Power within each bin:

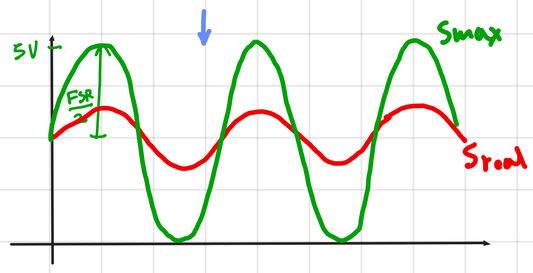


FSR = 5V

$0dB_c = S_{max} = \left(\frac{FSR}{2\sqrt{2}}\right)^2$

$S_{real} \neq S_{max} = 2.5\,V_p$

$S_{real} = 0.5\,V_p = 500\,mV_p$



$N_{min}^2 = \frac{LSB^2}{12} = \ldots \rightarrow -SNR_{ideal} = -50\,dB_c$

$SNR_{ideal} = 6.02n + 1.76 = 50\,dB$

$N_{sample} = 1024$

↳ attenuation:

$\Delta s = \frac{S_{max}}{S_{real}} = \frac{2.5\,V_p}{0.5\,V_p} = \frac{2.5}{0.5} = 5 = 7\,dB$

attenuation wrt $S_{max}(0dB_c)$

↳ power of $S_{real} = -7\,dB_c$

Noise floor: $NF = \frac{N_{min}}{\frac{N_{samples}}{2}} = -SNR_{ideal} - 10\log_{10}\left(\frac{N_{samples}}{2}\right) =$

$\overset{!}{=} -50dB_c - 10\log_{10}\left(\frac{1024}{2}\right) = -50\,dB_c - 27\,dB = -77\,dB_c$

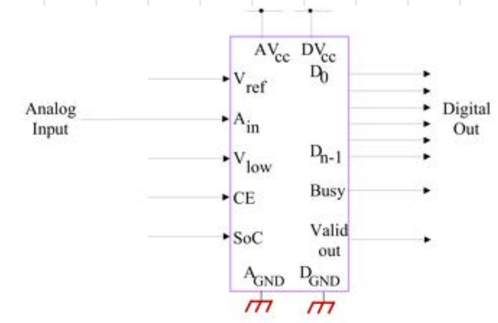
$N_{real} = NF \cdot \frac{N_{samples}}{2} = NF_{dBc} + 10\log\frac{N_{samples}}{2}$

Introduction

(Book p. 579)

Analog to digital converters are circuits that convert the analog signal provided at their input into the corresponding binary word during a time period ranging from µs to ms. In Fig. 8.1 is shown the schematic of a generic ADC converter. It gets, at the input entitled 'Analog Input', an analog voltage and converts that voltage into the corresponding output binary word. The output bits, indicated as D_0-D_{n-1}, can be provided either in parallel or in series. Moreover, many ADCs have additional pins, such as 'Valid out' to indicate the completion of conversion, 'ChipEnable' to enable or disable the whole ADC (by turning off the internal circuitry and putting the outputs to high impedance), 'Start of Conversion' (SoC) to begin the conversion, and 'Busy' (set by the ADC during conversion process).

Fig. 8.1



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

The device requires an external reference voltage V_{ref} (the maximum value corresponding to an output code with all 1s) and V_{low} (the minimum value, usually the ground, corresponding to a code with all 0s). Two models are commercially available: the so-called 'single quadrant', which converts voltages between 0V and FSR, and the 'two quadrants' with input dynamics which symmetrically extend to negative voltages (between $-V_{ref}$ and $+V_{ref}$). Moreover, also the digital code may vary, as shown in Tab. 8.1.

Non-linear errors

(Book p. 585)

ADCs are also affected by *non-linearity* errors due to the fact that the characteristics do not have the midpoints of the steps lying along a straight line, as would be the bisector in the case of an ideal ADC. The deviation from the ideal case can be assessed by comparing the real characteristics of the converter either with the line linking the values corresponding to the extreme codes '000' and '111' or with the best straight interpolating line of the real ADC characteristics ('*best-linear fitting*'), as shown in Fig. 8.6.

It is possible to express non-linearity in two ways (Fig. 8.7). The first way is entitled the *Differential Non-Linearity* (DNL) and is evaluated by considering the difference between the width of the n-th step and its ideal value of 1LSB. The second way is entitled the *Integral Non-Linearity* (INL), equal to the distance between the centre of the real step and that of the theoretical one (laying on the bisector or on the best-linear fitting). Manufacturers usually provide the maximum for both values. Actually, it can be seen how the INL relative to a code is equal to the sum of DNL of all of the previous codes.

When the DNL becomes equal to –1LSB, the real step is shorter than the ideal one by -1LSB, i.e. its width is zero: this means that that code is missing and that the ADC will never be able to provide it at its output, as shown in Fig. 8.7, on the right. Manufacturers state whether or not their ADCs suffer from this problem (*no missing code*), even if this specification is often referred to a number of bits lower than those actually available for the ADC. For example, a 10bit ADC can be guaranteed as 'no missing code' over 9 bits; this means that there may be missing codes, but not adjacent to each other, so that we will skip from a code to another distant 2 LSB and not more.

Fig. 8.7

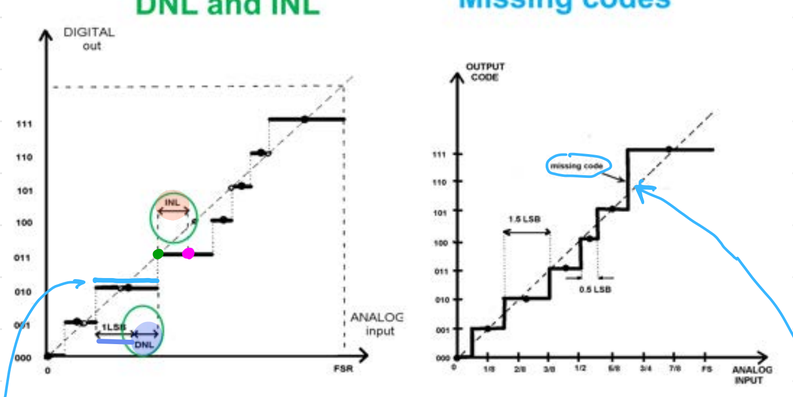
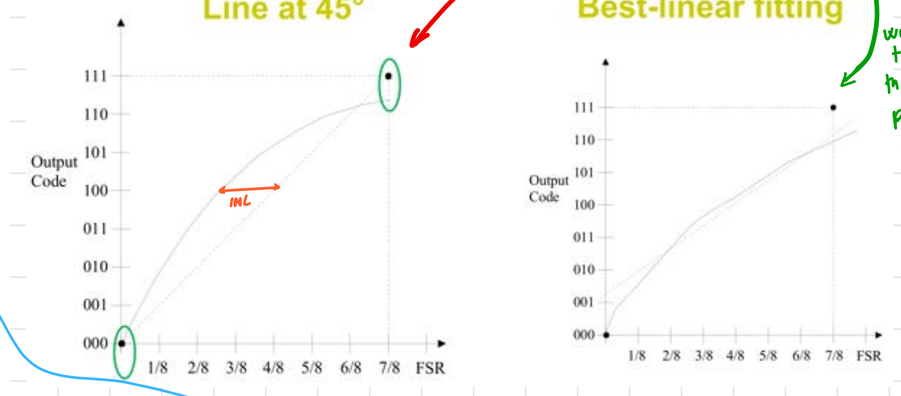


Fig. 8.6



Note: In ADC the analog input is NOT quantized so it's represent with continuous flat regions, while in DAC the input (D) was quantized so the output (A) was quantized as well.

this width = 0

Signal conditioning

→ We have to specify V_{ref}, V_{low} so we should adjust amplitudes and impedances to coherently specify them based on Vin

Signal conditioning: adjust amplitudes and impedances

Product: temperature meter

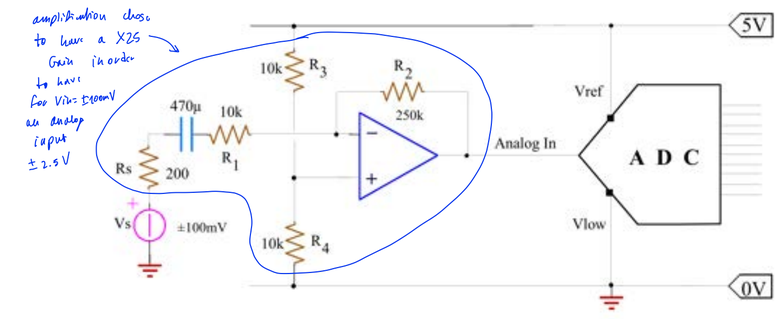
Specs:

	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	

pretty linear characteristic of the thermo resistance

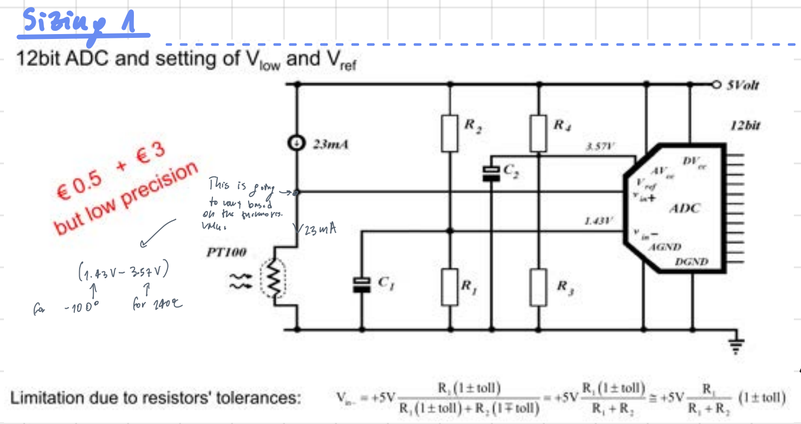
we need a 12-bit ADC to have 2400 steps

will discretized actually with 4096 steps



Sizing 1

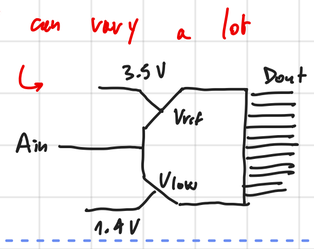
12bit ADC and setting of V_{low} and V_{ref}

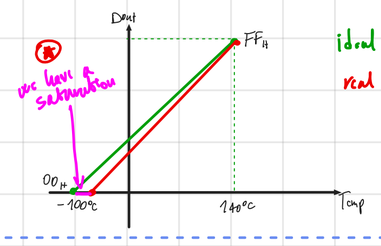


Limitation due to resistors' tolerances: $V_{in-} = +5V\frac{R_1(1\pm toll)}{R_1(1\pm toll)+R_2(1\mp toll)} = +5V\frac{R_1(1\pm toll)}{R_1+R_2} \cong +5V\frac{R_1}{R_1+R_2}(1\pm toll)$

→ To set R_1, R_2, R_3, R_4 we consider the voltage partition in order to set $V_{ref} = 3.57V$ $V_{low} = 1.43V$

If the resistors have bad tollerances it means that $V_{low} = V_{in}^-$ changes with the tollerances, so for bad components, it can vary a lot



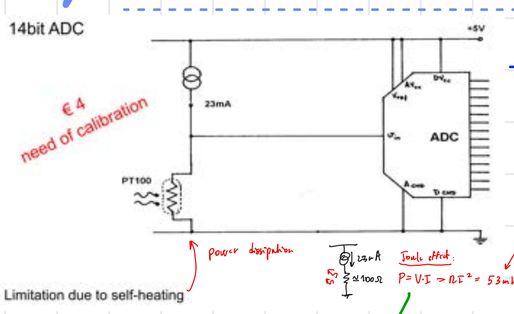


ideal characteristic

real characteristic (due to toll) ⊗

Sizing 2

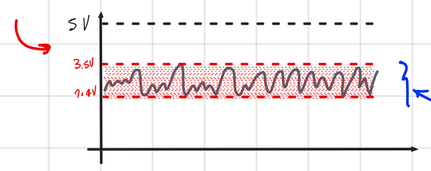
14bit ADC



Limitation due to self-heating

→ We directly connect V_{ref} to the max P.S. (5V) and V_{low} to the min (0V)

↳ But in this case we cannot use anymore the whole precision given by the 12 bits



because now we're NOT exploiting the FSR = 5V

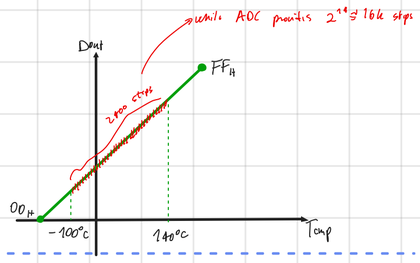
We wanted 2400 levels here and not in a FSR=5V

$5 : 2^n = (3.5 - 1.4) : 2400$

we have to buy a ADC with n>12 bit to have 2400 in this range

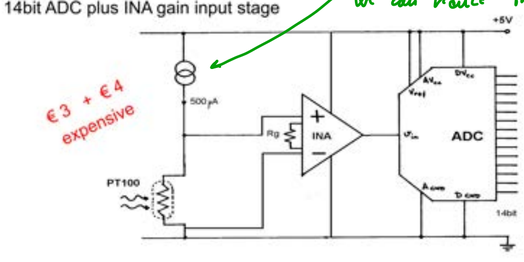
↳ ... n=14!

$LSB = \frac{5}{2^{14}} \cong \frac{(3.5-1.4)}{2400}$



Sizing 3

14bit ADC plus INA gain input stage



Expensive

We can reduce the power dissipation by reducing the current

↳ but having reduced the current also the input signal reduces → we have to amplify it

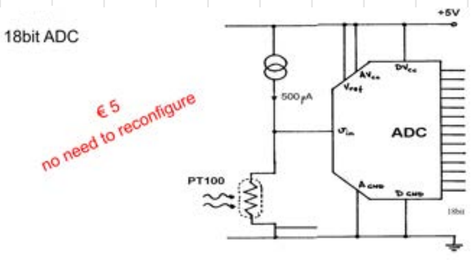
↑

Buy an INA for ex.

↳ EXPENSIVE!

Sizing 4

18bit ADC

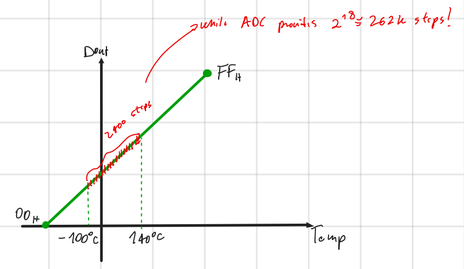


Cheap and no need for calibration or re-redign in case of different sensor, temperature range, resolution

↳ We can avoid to buy an INA

↳ but since the signal is attenuated we need a better resolution

↳ buy an ADC with a higher number of bits (18)

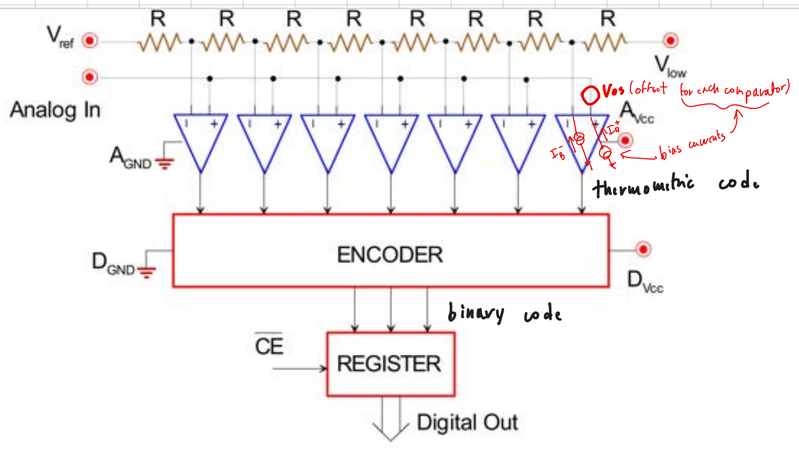


Flash ADC

The Flash ADC makes a parallel conversion of the input signal, ensuring extremely low conversion times (of the order of tens of nanoseconds). It is based on a resistive divider with 2^n resistors to create the corresponding 2^n–1 quantization levels and to be compared with the input voltage through as many comparators (Fig. 8.8). From the thermometric code of 2^n–1 bits at the output of the comparators (comparators switch in order with increasing input signal), we get to the desired binary code of n bits by simply using a digital encoder. The output register allows storing the last datum while the ADC proceeds with a new conversion.

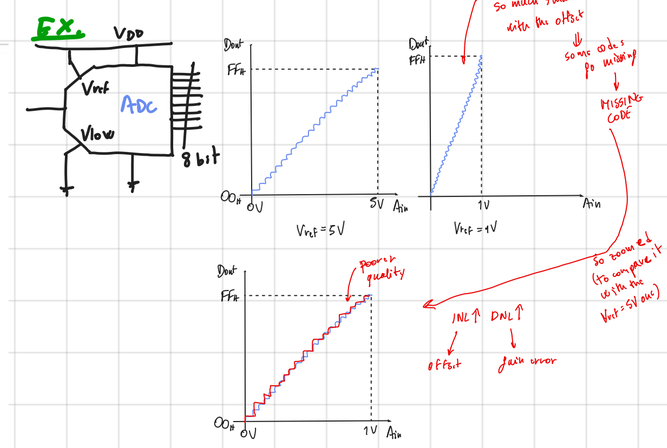
The extreme simplicity and speed of conversion of Flash converters are obtained at the cost of a high silicon area required and power dissipation rapidly increasing with the number of desired bits; in fact, already for an ADC with only 8 bits, 2^n–1=255 comparators are needed! In addition, the bias and leakage currents and the parasitic capacitors of the inputs of the comparators determine an alteration in the input voltage divider, with consequences on the linearity of the converter, which are more and more pronounced as the number of bits increases. The offset of the different components could lead to incorrect switching, resulting in missing codes, or even non-monotonicity of the ADC characteristics. It is definitely the faster ADC (conversion time T_C<50ns, i.e. $f_S=1/T_C$>20MHz), but with few bits (n≤10).

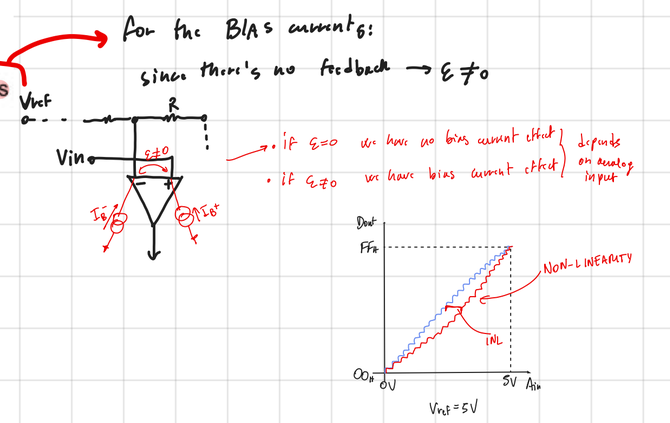
Fig. 8.8



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





Staircase ADC

Some ADCs need a DAC inside. The idea behind these configurations is to change, with subsequent adjustments, the binary code at the DAC input until its output voltage reaches the input signal value; the code that gives this condition will be the output of the ADC. This requires a comparator with precision (and hence offset) lower than ±½LSB.

One method to perform such a code search is to implement a simple sequential search feeding the DAC with a binary counter (Fig. 8.9), which forces a staircase at the output of the DAC. As long as the command /SoC (Start of Conversion) is disabled (i.e. at level High, being active low since it has a line over it), the flip/flop is held with Q low and /Q high, that is, the clock can get to the counter through the AND, but the counter is kept at zero by the reset pin (active high). Just activated the /SoC (making it low), the counter starts its counting, which will be stopped when V_{DAC} passes V_{in}.

At this point, the flip-flop S/R will set the 'End Of Conversion', and will simultaneously disable the clock to reach the counter (resetting /Q to low level). The value in the counter will be the desired Digital Out.

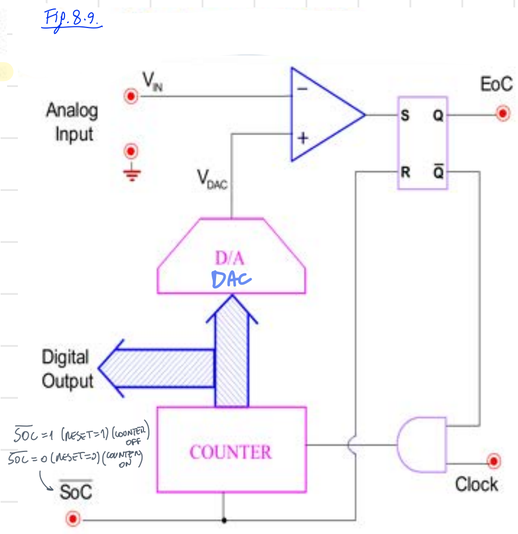
The precision of the conversion depends on that of the DAC, which must have monotonic characteristics, small offset and non-linearity errors. The conversion time T_C, which depends on the amplitude of V_{in}, may require up to 2^n clock periods (when v_{in} is the maximum, that is, equal to the FSR, and all the bits have to be set to 1), therefore:

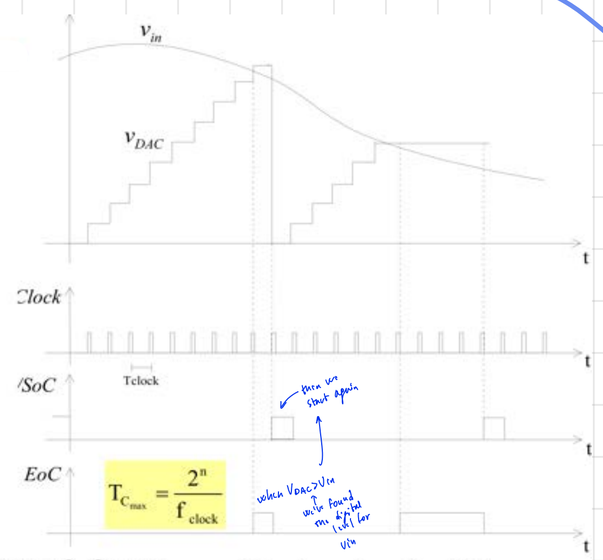
$$T_{C_{max}} = \frac{2^n}{f_{clock}}$$

A 10bit converter with f_{clock}=10MHz will have a maximum T_C equal to 102.4µs, namely f_S<10kHz; according to the Sampling Theorem, this limits the maximum frequency of the input signal to only about $f_{in,max}=f_S/2$=5kHz. The maximum clock frequency is directly related to the speed of the DAC and that of the comparator; for instance, with f_{clock}=10MHz, in less than 100ns, the counter has to increase, the DAC settles its output, and the comparator eventually switches.

Another disadvantage of the staircase architecture is represented by the fact that the sampling comb is not constant. In fact, providing the signal at regular periods for the start of the conversion SoC, the ADC will finish after a period of time dependent on the value of the input, therefore not granting a sampling with regular steps. The consequences would be high non-linearities. To avoid this problem, it is necessary to introduce a Sample & Hold at the input, which freezes the datum all the times the command SoC is set; in this case, when /SoC goes low, the S&H has to go in the hold phase.

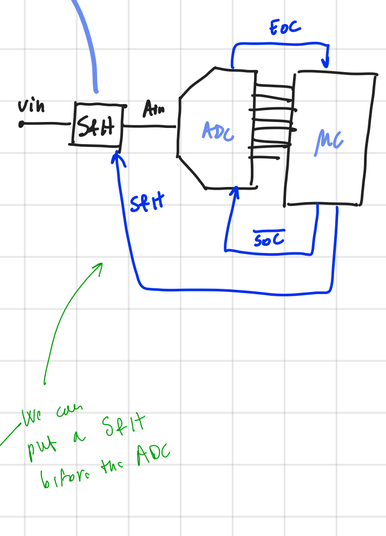
Fig. 8.9





Components: 1 DAC and 1 counter
1 comparator

Pros & Cons: precision depends on the DAC
not very fast (T_C>100µs)
↖ depends on Vin amplitude



We can put a S&H before the ADC

Tracking ADC

In the tracking ADC (Fig. 8.10), at every new conversion, the counter is not simply reset and progressively increased as in the previous case, but it is made to increase or decrease, depending on whether the input voltage is greater or lower than that given by the DAC. In this way, after the first phase of staircase (similar to the previous case), if the input signal is slowly variable in time, the DAC is able to follow it in few clock periods and eventually remains attached to it at every clock period. An up-down counter properly driven by a simple logic that exploits the output of the comparator is needed.

For the ADC to be able to remain 'hooked', it is necessary to limit the maximum slope of the input signal so that $dV/dt|_{max}$<LSB/T_{clock}, i.e. limit the sinusoidal frequency below:

$$f_{max} \le \frac{f_{clock}}{2^n \pi}$$

$\left|\frac{dV}{dt}\right|_{max} = 2\pi f_{max} \cdot V_{p\,max} < \frac{LSB}{T_{clock}}$

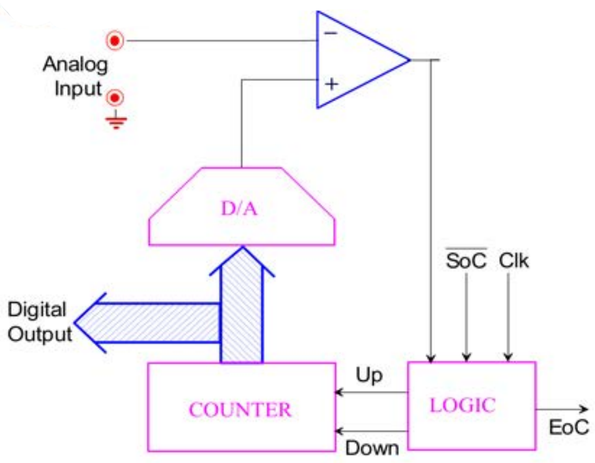
$2\pi f_{max} \frac{FSR}{2} \le \frac{FSR}{2^n} f_{clk}$

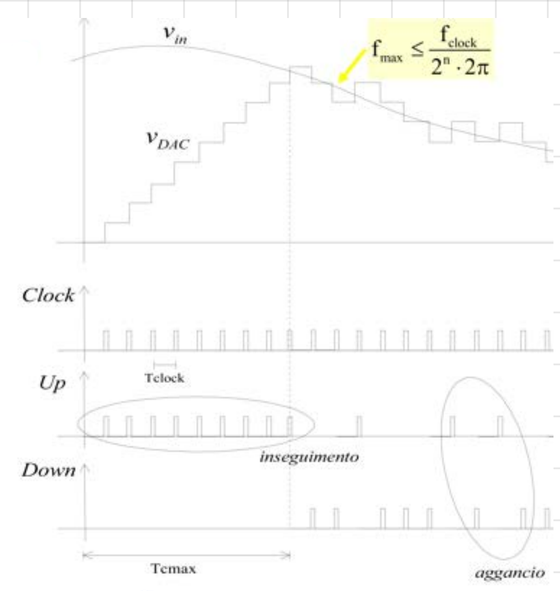
$f_{clk} \ge \pi \cdot 2^n f_{max}$

For example, a 10bit converter with f_{clock}=10MHz could follow a sine wave at full dynamics with f_{max}≤3kHz. Note, however, that this ADC is much faster than the previous one (although the previous example limited the maximum input frequency to 5kHz) because now the ADC provides a correct sample at every clock pulse, namely every 100ns (and not every T_C=0.1ms, as before)!

The signal of End-of-Conversion, to make sure that the digital code at the output is correct, i.e. the converter has effectively hooked the input signal, should be obtained through a simple logic network that checks the alternation of Up and Down signals, as shown in Fig. 8.10, on the right. This will require a frequency of the signal much greater than the f_{max} found so that, even in situations of high slope, there is the 'alternation' mentioned above; in other words, with f=f_{max}, the signal is effectively 'hooked', but we will have only 'up' and 'down' pulses.

Fig. 8.10

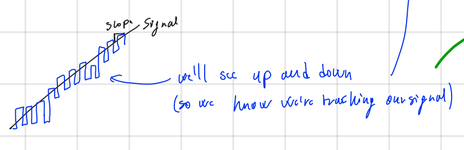




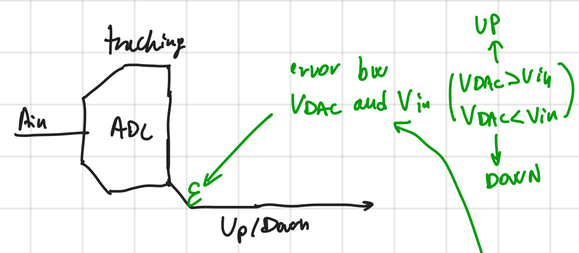
Components: 1 DAC, 1 counter, 1 comparator,
1 up/down logic

Advantages: precision depends on DAC fast (T_C>100ns) possible just ONE BIT (*oversampling*)

use $f_{clk} >> \pi 2^n f_{max}$
(like 4 times)
↓
Oversampling



we could just use the UP/DOWN bit because when we're "hooked to the signal" we know that
• UP: VDAC > Vin → decrease output value
• DOWN: VDAC < Vin → increase output value



this type of ADC is also called Δ modulator

Single-slope ADC

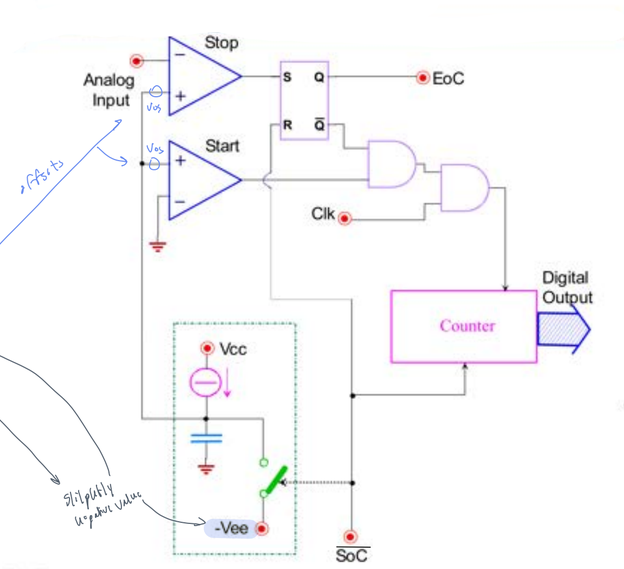
This converter is based on a working principle similar to that of the staircase ADC, but uses an analog ramp (and not a digital staircase) to make the comparison, as clearly shown in Fig. 8.11, thanks to a constant current source and a capacitor (instead of a clock and a counter). Until the command SoC (Start of Conversion) is set, the ramp does not start, and the counter is kept to zero. Across the capacitor a linear charge will develop, just like a ramp. The 'Stop' comparator will give the End of Conversion (setting EoC) when the ramp will reach the value V_{in}.

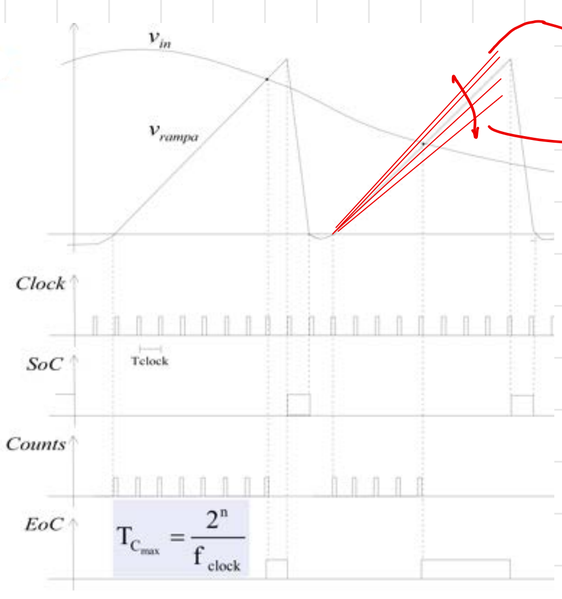
The closure of the switch that enables the capacitor charge is not instantaneous (due to the resistance of the MOS switch and parasitic). For this reason, a second 'Start' comparator is introduced, which enables the counting only when the ramp crosses a certain threshold. To avoid problems due to the offset voltages of the comparator, it is preferred to make the ramp start from slightly negative voltages.

The maximum conversion time is high (equal to that of the staircase ADC): $T_{Cmax}=2^n/f_{clock}$. Also in this case, we have the problem of the irregular sampling comb which can have high non-linearity effects if not treated in the right way, for instance introducing an S&H before the ADC.

Unfortunately, the circuit remains very sensitive to the tolerance degrees of the capacitor, of the current source I_{ref}, and of the clock period. Tolerance degrees or thermal drifts on the parameters of the charge cause a ramp more or less suddenly that, ultimately, will cause a different reading at the output of the counter and hence a low conversion accuracy. To appreciably reduce these problems, there is 'double ramp' architecture.

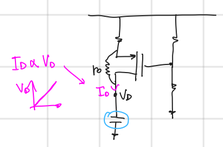
Fig. 8.11

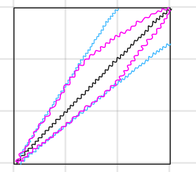




Advantages: precision dependents on dV/dt = I/C many bit (n>16) but slow (T_C>1ms)
possible enhancement: double-ramp, dual-slope ADC

$\frac{dV}{dt} = \frac{I(1 \pm tol_I)}{C(1 \pm tol_C)}$
↳ variations due to tolerance
↳ and also current source usually depends on the voltage → lead to non-lin.
↳ e.g. possible current source schemes





this errors cause distortions in the output

Dual-slope ADC

The feature of this ADC is to use a first charge ramp with a variable slope, proportional to the analog input signal (obtained by a simple integration of the signal itself), for a fixed time, followed by a discharge ramp with a constant slope (obtained by integrating a constant reference voltage). The scheme is shown in Fig. 8.12. Once the SoC command has been received, a suitable control logic internal to the ADC resets the integration capacitor (through the closure of the MOS M3) and the digital counter and then proceeds with the integration of the input signal (closure of M1) for all the time necessary for the counter to overflow, i.e. 2^n clock pulses. Once the overflow bit of the counter (its n+1th output bit) has been set, the control logic opens M1 and closes M2 (there is no need to reset the counter because, after the overflow, it automatically restarts from 0) to begin the phase of discharge (note the opposite signs of V_{in} and V_{ref}) with a constant slope.

When the voltage across the capacitor (and hence the output of the first OpAmp) is back to 0V, the conversion is over, as shown in Fig. 8.13; the number N of clock pulses required to bring the voltage V_X of the integrator back to zero is proportional to the value of the input V_{in}; therefore, the following is the expected digital result:

$$N = 2^n \cdot \frac{V_{in}}{V_{ref}}$$

Note how in the first ramp the amplitude achieved depends on R, C, and V_{in} despite the current $I=V_{in}/R$. The discharge with a constant slope will last for a period of time $T_2=T_1 \cdot V_{in}/V_{ref}$. However, since both charge (first ramp) and discharge (second ramp) of the capacitor take place with the same integration constant R and C, the value N does not depend on R, C, and f_{clock} (at least as long as these quantities remain stable throughout all the conversion time). In fact, possible tolerance degrees of the parameters R and C equally act on the two integration phases with no effect on the conversion accuracy, as shown in Fig. 8.14. Therefore, the dual slope ADC offers good performance in terms of linearity and accuracy (as long as the static and dynamic performance of the integrator and of the comparator are good), making it possible to achieve resolution even higher than 20 bits. The maximum conversion time, equal to $T_{cmax}=T_1+T_{2max}=2 \cdot 2^n/f_c$ (Fig. 8.14), is unfortunately high, equal to twice the corresponding single ramp ADC.

Finally, a very interesting feature of this architecture, as that of any other architecture that integrates the input signal, is the possibility to reject disturbances, with frequencies multiple of the integration period, superimposed on the useful signal V_{in}. In fact, a generic disturbance (both bump and ripple) superimposed on the input voltage can produce a change in the slope of the charge of C, altering the voltage reached at the end of the phase T1 and, thus, the result N of the conversion. Instead, any fluctuation in V_{in}, which is likely to end the phase T1 with the same amplitude V_X stored in the capacitor, will not give any error on the output N. This peculiarity is specified with the Normal Mode Rejection (NMR), shown in Fig. 8.15, for two different clock frequencies, properly chosen (depending on the number n of bits) to be able to completely reject disturbances at 50Hz or 60Hz from the network, along with all their harmonics.

Fig. 8.12.

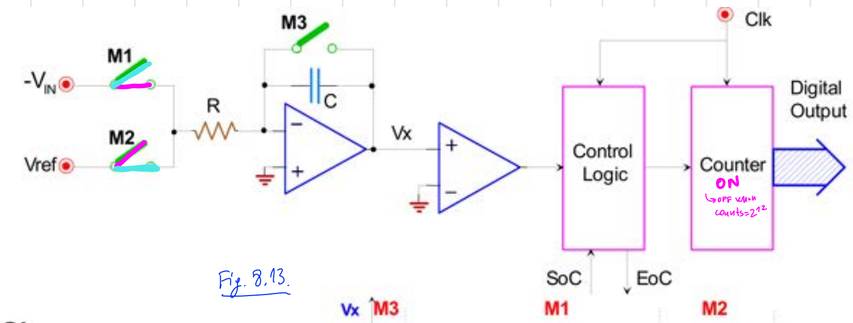
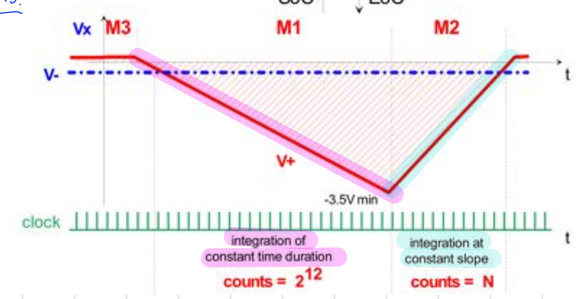


Fig. 8.13.



Advantages:

precision independent of R and C

$$T_{C_{max}} = 2 \cdot \frac{2^n}{f_{clock}}$$

Fig. 8.14.

1st advantage of dual-slope ADC: rejection of tolerances of R and C

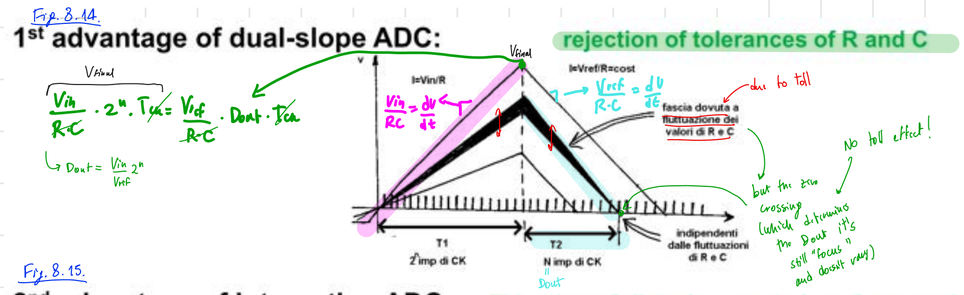
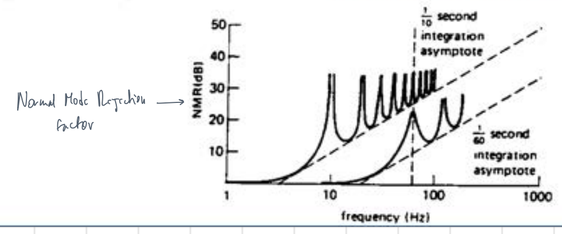
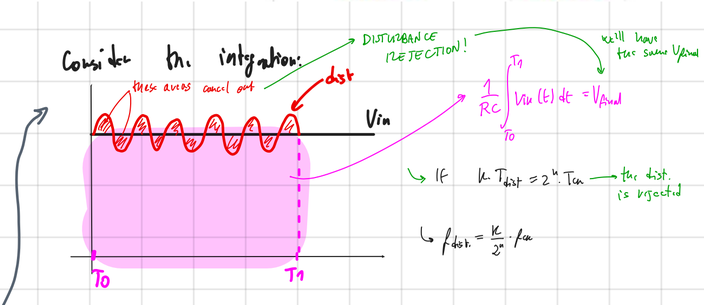


Fig. 8.15.

2nd advantage of integration ADCs: rejection of disturbance at given frequencies



Possible questions:
- Is the bias current going to 'cause any error?
- Does the offset cause any error?
- Is charge injection relevant? Should we compensate it with a dummy cell?



Successive approximation ADC (SAR)

This type of analog to digital converter exploits a very efficient code search technique, entitled *binary search,* which requires a number of clock cycles equal only to the number of bits to be provided at the output (Fig. 8.16). Following the command Start Of Conversion (SoC), the sequential SAR network begins by asserting the Most Significant Bit (MSB) of the digital code and comparing the corresponding analog value (equal to FSR/2), generated by the internal DAC, with the voltage at the input: in the case V_{in} is still greater than V_{DAC}, the level of the bit is kept; otherwise, it is lowered to 0. At every subsequent clock stroke, the SAR sets a bit at a time and then decides whether to keep it that way or to reset the level, until the LSB is reached.

A conversion requires only n+1 clock pulses instead of 2^n of the ADCs seen so far:

$$T_{Cmax}=(n+1)/f_{clock}$$

For example, a 10bit ADC with f_{clock} equal to 10MHz has $T_{Cmax}=1.1\mu s$ (instead of 100µs of a ramp ADC), and it becomes possible to handle signals with a maximum frequency of approximately 450kHz (a factor of 100 compared to the single-slope ADC). The improvement in conversion speed is more and more pronounced at increasing resolution. For example, for a 14bit ADC, the conversion would be $2^n/(n+1)=1092$ times faster.

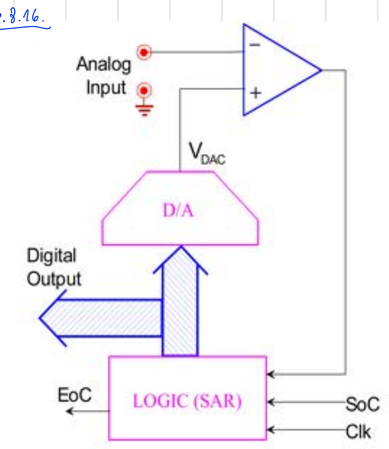
In order for the conversion to be successful, it is important that the input signal V_{in} is constant within ½LSB during the entire search for the correct code, i.e.:

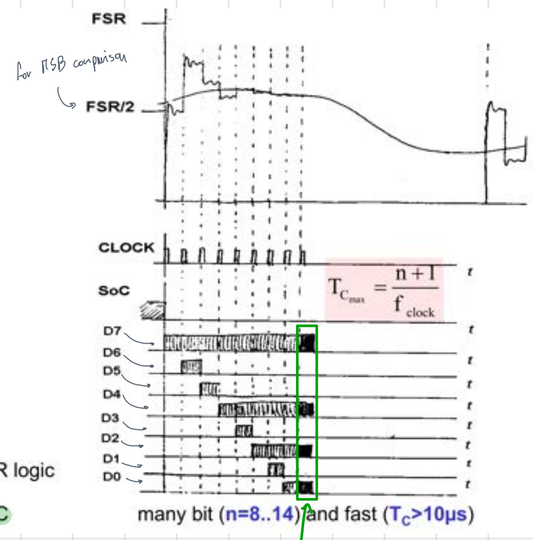
$$f_{in,max} \le \frac{f_{clock}}{2\pi \cdot 2^{n+1} \cdot (n+1)}$$

in the example above, that would result in a limiting frequency of only 155Hz! If the input signal is too variable during the conversion time, this would lead to an error (see Fig. 8.17).

We must now resolve the question: "How can it be possible that an ADC capable of handling signals at 450kHz (according to the Sampling Theorem) actually fails to convert sine waves with a maximum frequency of less than a few hundred Hz?" To overcome this obstacle, just ensure that V_{in} cannot change during the time T_C, not to distort the conversion: this can be obtained simply by putting an S&H before the ADC.

Fig. 8.16.





$$T_{C_{max}} = \frac{n+1}{f_{clock}}$$

Components: 1 DAC and 1 sequential SAR logic

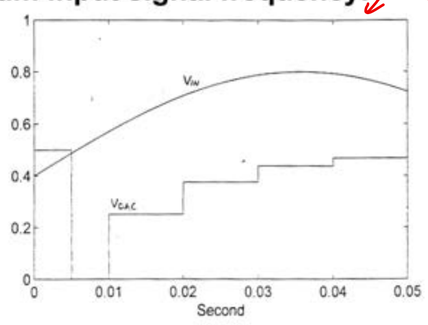
Advantages: precision depends on the DAC

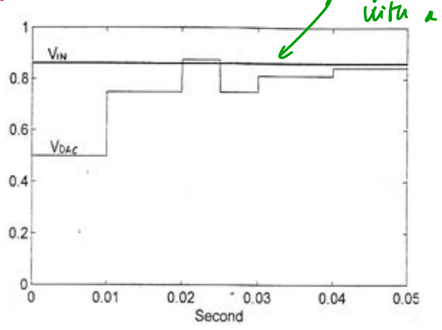
many bit (n=8..14) and fast (T_C>10µs)

after 8 comparison we have the Dout = 10010101

Fig. 8.17.

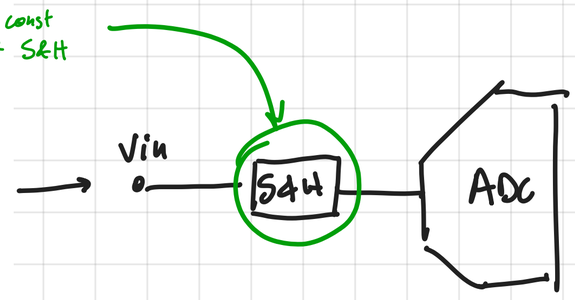
Maximum input signal frequency:





When Vin is variable

keep it const with a S&H



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:

$$f_{in,max} \le \frac{f_{sampling}}{2} = \frac{f_{clock}}{2 \cdot (n+1)}$$

$$\left|\frac{dV}{dt}\right|_{max} = 2\pi f_{max} \cdot V_{pmax} \le \frac{dV}{dt} = \frac{1LSB}{n\,T_{ck}} = \frac{FSR}{2^n} \frac{f_{clock}}{n} \longrightarrow 2\pi f_{max} \frac{FSR}{2} \le \frac{FSR}{2^n \cdot n} f_{ck} \longrightarrow f_{ck} \ge \pi \, 2^n \cdot n \, f_{max}$$

$$T_{conv} = (n+1)\,T_{ck}$$

$$\frac{1}{T_{conv}} = \frac{1}{T_S} = f_{sampling} = \frac{1}{(n+1)} f_{ck} \longrightarrow f_{max} \le \frac{f_{sampling}}{2} \longrightarrow f_{ck} \ge 2 \cdot (n+1) f_{max}$$

Shannon

Timings

To fully understand how to drive an ADC, it is necessary to study its appropriate timing, supplied by manufacturers themselves. In addition to the tables summarizing the features of the ADC (such as that in Tab. 8.2), these precise 'conversion timings' for measuring the speed of the converter are important. In fact, the *conversion time* T_C described so far is important when we want to make the conversion in a very specific instant of time, indicated by the activation of the SoC: this type of event is called the *single-shot*. In other applications, however, it is important to make a continuous series of conversions at the highest rate possible; this type of use is called the *free-running*, and it is important to quantify the number of *Samples per Second* (SpS), which the converter can provide.

In general, the SpS are different from the *sampling rate* ($1/T_S$) or the *sampling frequency* f_S of the ADC. The latter is generally less than the inverse of the conversion time T_C due to the presence of some incidental timing and various delays (settling time, hold time, reset delay …).

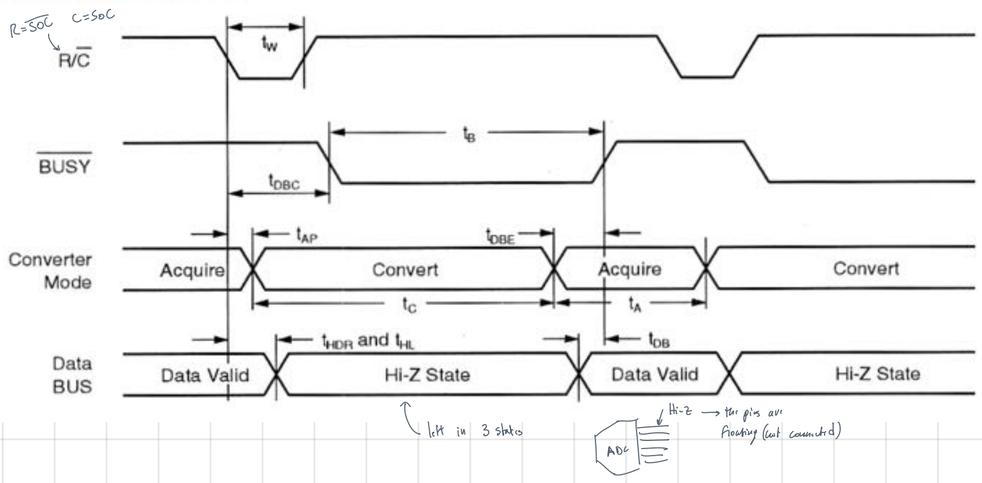
Single-shot ADC

The typical timing diagram of a non-pipelined ADC is shown in Fig. 8.19. The request for a new conversion is done by lowering the R/C signal (Ready/Convert, homologous of the SoC); the ADC will then convert the analog input stored during the previous 'Ready' phase (R/C high). In the datasheet of the component, the manufacturer specifies the time required to transit from the 'Acquire' phase to the 'Convert' phase (t_{AP}), coincident with the opening delay of the internal S&H. However, it is only after a time t_{DBC} starting from the falling edge of R/C when the conversion actually begins, and the 'busy' signal is set (low). From the conversion end, it may take a delay t_{DBE} before the ADC communicates with the outside world that the conversion is finished, raising again the pin of busy. That delay also includes the settling time that the latches on the output bus need to provide the correct data. Simultaneously with these operations, the ADC starts a new phase of acquisition and tracking of the input signal so as to be ready for the next conversion.

The total time for the conversion of the data is hence equivalent to the sum $t_{DBC}+t_B$, not only to $T_C=t_c$ as described so far, precisely because of the additional 'overheads' that add up to the actual conversion. Finally, the delay t_{HDR}, required a new conversion, with which the ADC will take the result of the previous conversion from the bus is shown, putting it into a high impedance state. In conclusion, for this ADC, we have: $SpS<1/T_C$.

Fig. 8.19

single-shot ADCs:



Free-running pipelined ADCs (not really seen at lesson)

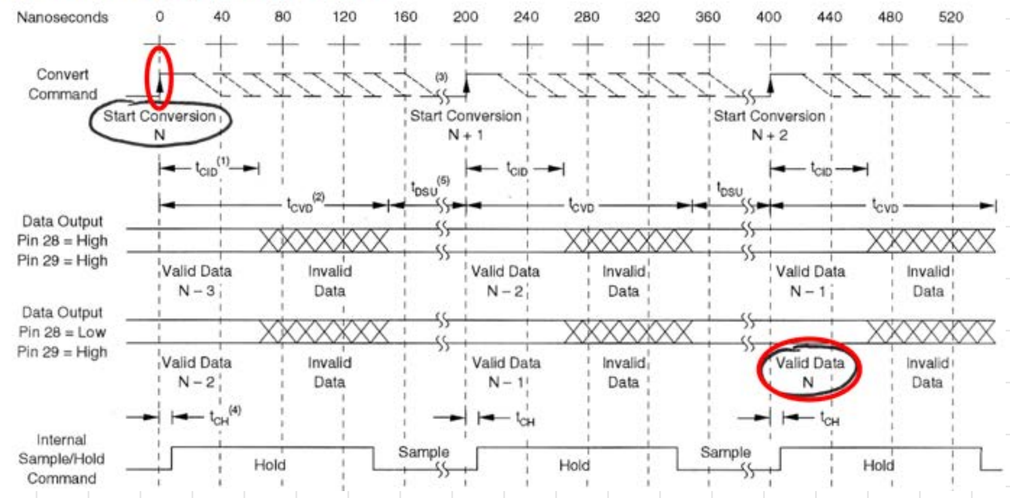
There are pipelined type structures in which it is possible to start a new conversion even if the conversion corresponding to the previously acquired data is not yet completed, thanks to more cascaded blocks that work as in an assembly line (or bucket-brigade style), each performing a simple task (such as the S&H, the SAR, the comparison logic, the latch of the output data, …), but on different samples. In this way, it is possible to maximize the *throughput*, although the single conversion can be slow.

Fig. 8.20 shows the timing of a pipelined ADC that accepts input data (and therefore provides the conversion at the output) at a frequency two or three times higher than that of the conversion, i.e. SpS=2/T_C or 2/T_C (depending on how the pins 28 and 29 of the ADC are set). The advantage of this configuration is to have a high *throughput* compared to the rate of conversion of each sample.

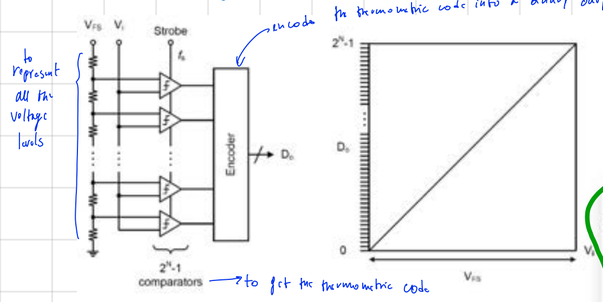
The disadvantage is the *latency* with which the converted data will appear at the output; in this case, it may be $2 \cdot t_{CVD}$ or $3 \cdot t_{CVD}$, depending on how the ADC is set (as shown in Fig. 8.20). In particular applications, such as single-shot or not regular conversions, it is preferable to use non-pipelined ADCs with no latency.

Fig. 8.20.

free-running pipelined ADCs:



Subranging ADC



to represent all the voltage levels

encode the thermometric code into a binary output code

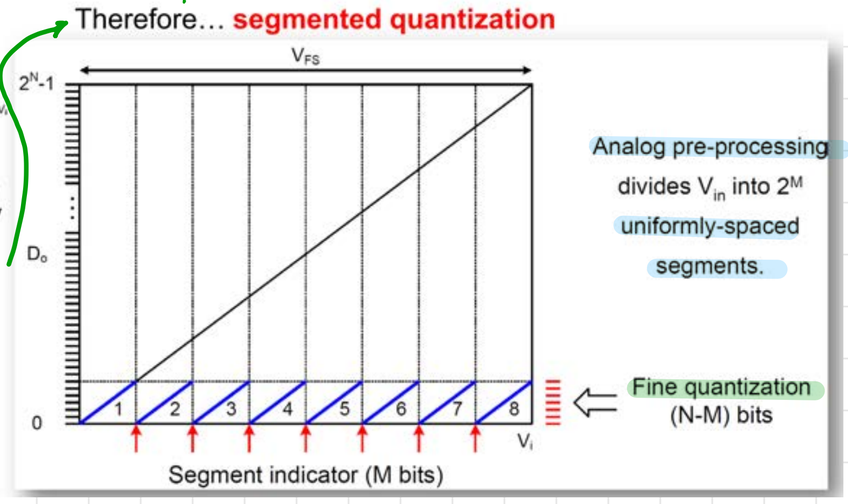
to get the thermometric code

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

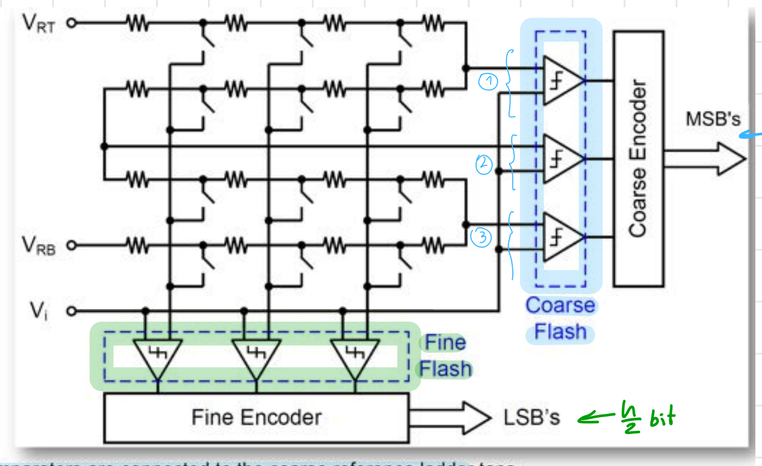
↳ e.g. comparators keep consuming power even if they're not triggered

We can perform the coarse quantization with less level and then the fine quantization with more levels (more precise)

Therefore… **segmented quantization**



ranges



- Coarse comparators are connected to the coarse reference ladder taps
- After getting MSBs, the fine ladder taps are enabled and fine LSB are computed
  - ↳ to perform on the (range) identified by the coarse flash
    - ↳ fine flash only on a subrange

↳ consumes less power and silicon area

Interpolation Flash-ADC

(Book p. 643)

When the required resolution reaches 8÷10bit, the technological efforts to realize a classical flash ADC do not suffice to adopt the classical scheme (Fig. 9.1, on the left). The area occupation, the power consumption, the resistive network tolerance, the comparator offset, and these uncertainties on the signal propagation would be unsustainable. In such a case, we could use "interpolation" and "folding" techniques.

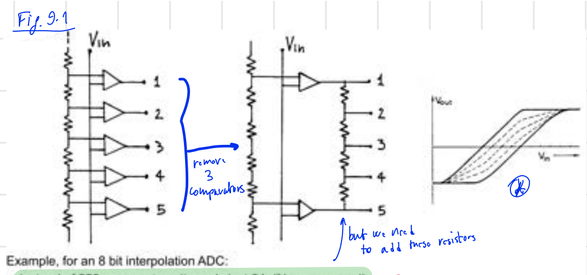
The *interpolation* concept is shown in Fig. 9.1, in the center: for an 8bit ADC; instead of using 256 comparators, we could use only 64 comparators, removing three comparators from every four comparators. The three missing outputs are reconstructed with a network of resistors between the remaining outputs. We can assume the contact points between the resistors ("taps") as the removed comparator outputs.

The Fig. 9.1 (on the right) shows the characteristic V_{out}/V_{in} for the proposed structure; the bold lines refer to the remaining comparators, and the dotted lines are for the interpolation taps. Because the comparators are real, the characteristics are not steps, i.e. V_{out} does not commute from the low value to the high value when V_{in} exceeds the reference voltage for an infinitesimal value. Because of the comparator's finite gain, we have a region with a vertical slope, in which the comparators have a quasi-linear behavior around the reference voltage. Therefore, V_{out} is an index of similarity between V_{in} and V_{ref}. We can exploit this information to reconstruct the output signal for the removed comparators. The reconstructed voltages have a high distortion at large values of V_{out} (i.e. in case of comparators close to the saturation). Nevertheless, the information that reaches the latches connected downstream the comparators is not the value V_{out}, but the sign of V_{out} (i.e. above or below 0V).

This technique allows considerably reducing the area occupation and the power consumption while the reduction in the input capacitance and the layout dimensions significantly improve the dynamic performance. A commercial example is the AD9060 by Analog Devices, which obtains 10bit at 75Msps with only 512 comparators rather than 1024; these comparators encode 9bit while the remaining bit is obtained by means of interpolation. A similar approach, *interpolated-parallel*, allows considerably reducing the effect of the offset on the differential non-linearity. For example, we can obtain 10bit at 300Msps by means of a network with 128 OpAmps with differential output, whose outputs are connected, pair-wise, to the interpolation network. These interpolation networks are constituted by eight resistors that are connected to eight comparators, as shown in Fig. 9.2. The number of comparators is unchanged (128·8=1024), so the area and the power consumption are high; however, the advantages are significant. Considering two consecutive OpAmps, N-1 and N, we can see in the figure that the interpolation is reached connecting four resistors between the non-inverted output voltages V_{n-1} and V_n and the other four between the inverted voltages CV_{n-1} and CV_n. At this step, the eight comparators can be connected to have eight reference voltages indicated by the white circles.

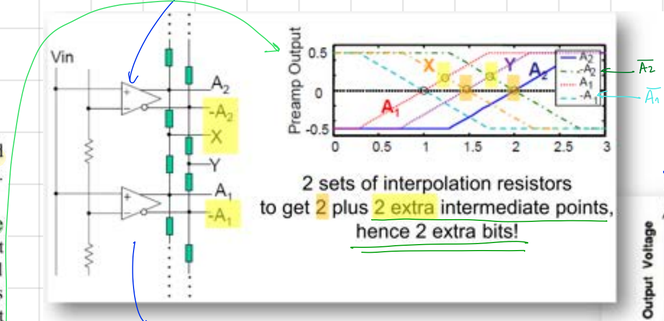
The first result is the reduction in the input capacitance in respect of the *traditional* scheme. In fact, there are only 128 OpAmps instead of the 1024 comparators; a value shown in literature is only 8pF. The second result is the DNL reduction due to the offset voltages V_{ov}. Particularly, the DNL due to the OpAmp offset is reduced by a factor equal to the interpolation factor M (in this case equal to 8); in fact, the amplitude error for every interval is limited to $V_{ov}/8$. Moreover, the DNL due to the comparator offset is reduced by a factor equal to the gain G of the OpAmps (in this case equal to 10). Indeed, the amplification reduces the comparator offset voltage at the input by a factor of G. Therefore, to obtain a DNL=±½LSB, while the conventional flash 10bit ADCs requires an offset of 150μV (not simply reachable), the interpolated-parallel scheme requires a value of 0.8mV, simply reachable.

Fig. 9.1



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, …)



differential comparator

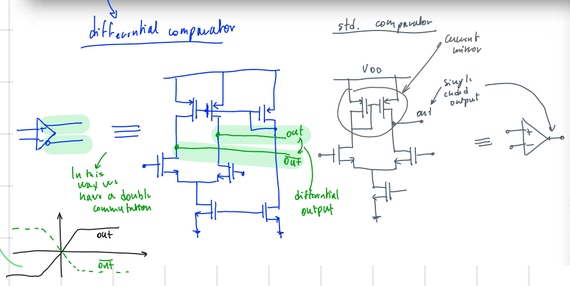
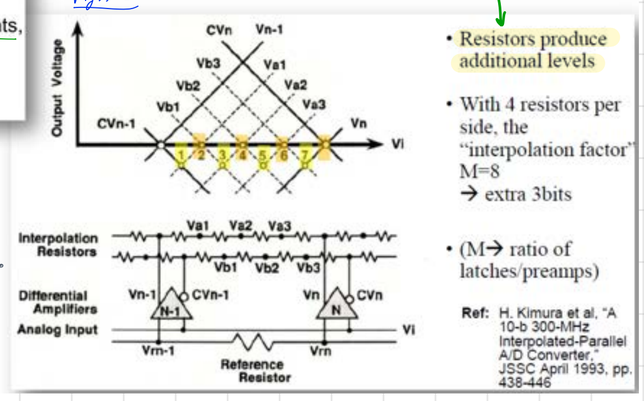


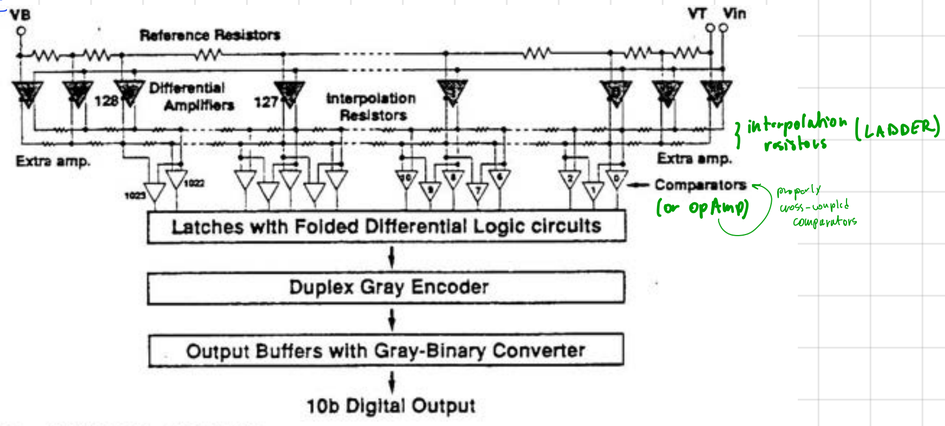
Fig. 9.2

We can use more than one resistor ladder to add other interpolating levels



- Resistors produce additional levels
- With 4 resistors per side, the "interpolation factor" M=8 → extra 3bits
- (M→ ratio of latches/preamps)

Ref: H. Kimura et al, "A 10-b 300-MHz Interpolated-Parallel A/D Converter," JSSC April 1993, pp. 438-446



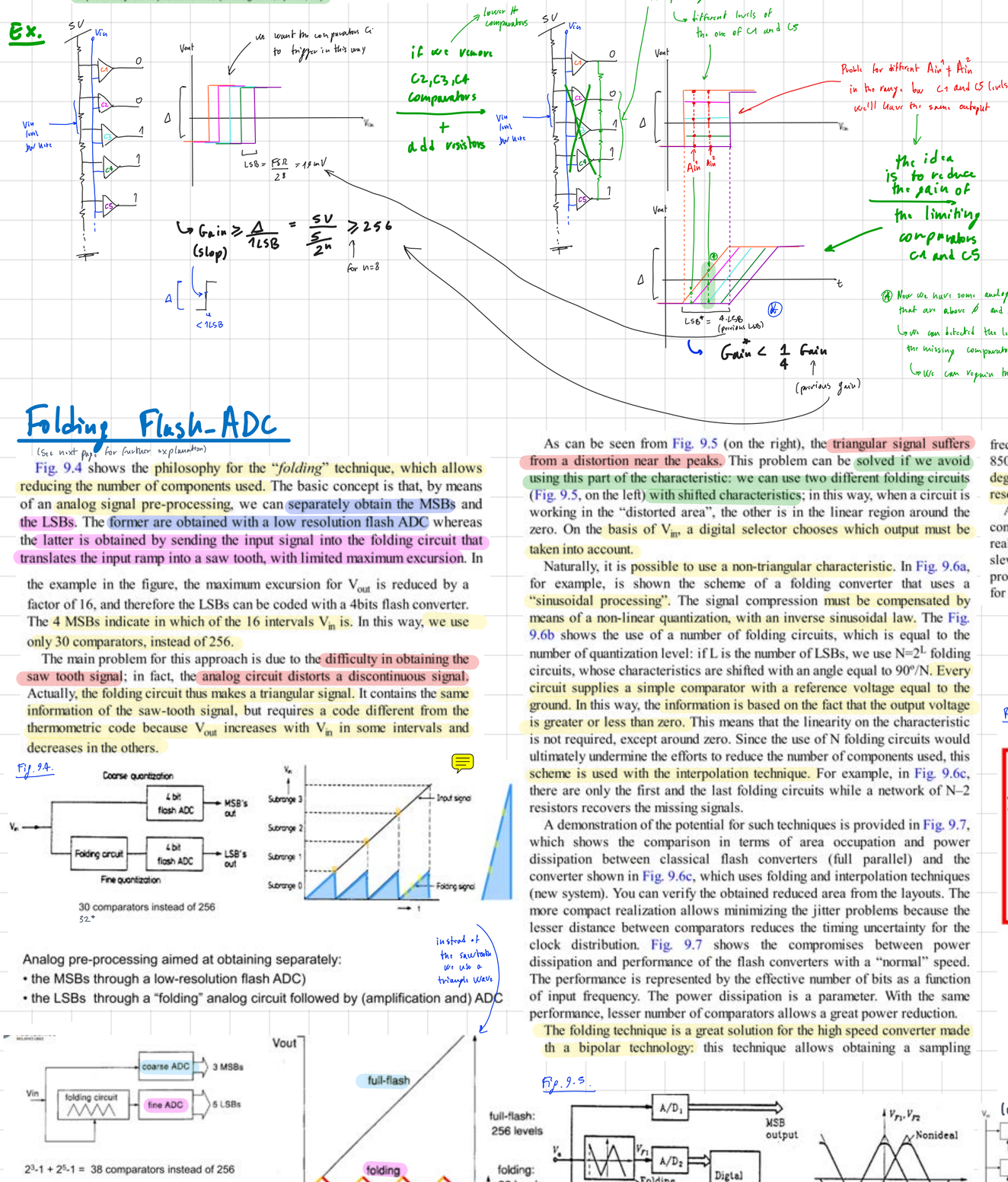
interpolation resistors (LADDER)

properly cross-coupled comparators

Example of a 10bit ADC with 300Msps:

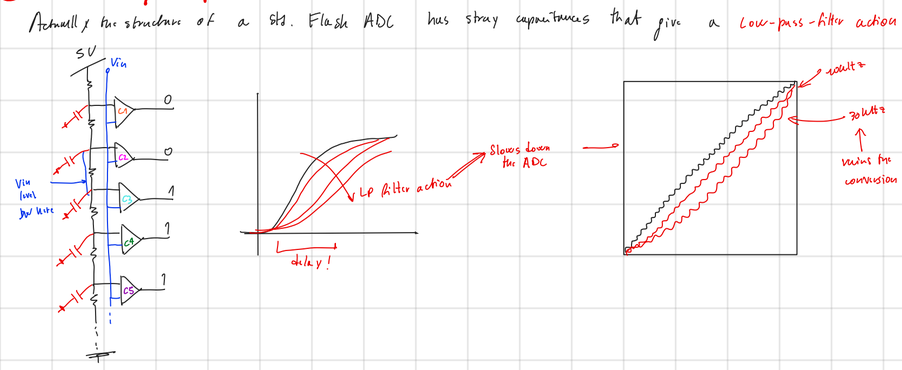
- 128 differential OpAmps (G=10), each with 4+4 resistors and 8 comparators
- same total number of comparators (128·8=1024), hence same area and dissipation
- but advantages for a reduced C_{in} (128/1024), better DNL ($V_{osOpAmp}/8$ e V_{osComp}/Gain)

Ex.



Ex. (B) → stray capacitance

Actually the structure of a std. Flash ADC has stray capacitances that give a low-pass-filter action



Folding Flash-ADC

(See next page for further explanation)

Fig. 9.4 shows the philosophy for the "*folding*" technique, which allows reducing the number of components used. The basic concept is that, by means of an analog signal pre-processing, we can separately obtain the MSBs and the LSBs. The former are obtained with a low resolution flash ADC whereas the latter is obtained by sending the input signal into the folding circuit that translates the input ramp into a saw tooth, with limited maximum excursion. In the example in the figure, the maximum excursion for V_{out} is reduced by a factor of 16, and therefore the LSBs can be coded with a 4bits flash converter. The 4 MSBs indicate in which of the 16 intervals V_{in} is. In this way, we use only 30 comparators, instead of 256.

The main problem for this approach is due to the difficulty in obtaining the saw tooth signal; in fact, the analog circuit distorts a discontinuous signal. Actually, the folding circuit thus makes a triangular signal. It contains the same information of the saw-tooth signal, but requires a code different from the thermometric code because V_{out} increases with V_{in} in some intervals and decreases in the others.

Fig. 9.4.

30 comparators instead of 256

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

instead of the saw-tooth we use a triangle wave

Different ways of folding

As can be seen from Fig. 9.5 (on the right), the triangular signal suffers from a distortion near the peaks. This problem can be solved if we avoid using this part of the characteristic: we can use two different folding circuits (Fig. 9.5, on the left) with shifted characteristics; in this way, when a circuit is working in the "distorted area", the other is in the linear region around the zero. On the basis of V_{in}, a digital selector chooses which output must be taken into account.

Naturally, it is possible to use a non-triangular characteristic. In Fig. 9.6a, for example, is shown the scheme of a folding converter that uses a "sinusoidal processing". The signal compression must be compensated by means of a non-linear quantization, with an inverse sinusoidal law. The Fig. 9.6b shows the use of a number of folding circuits, which is equal to the number of quantization level: if L is the number of LSBs, we use $N=2^L$ folding circuits, whose characteristics are shifted with an angle equal to 90°/N. Every circuit supplies a simple comparator with a reference voltage equal to the ground. In this way, the information is based on the fact that the output voltage is greater or less than zero. This means that the linearity on the characteristic is not required, except around zero. Since the use of N folding circuits would ultimately undermine the efforts to reduce the number of components used, this scheme is used with the interpolation technique. For example, in Fig. 9.6c, there are only the first and the last folding circuits while a network of N–2 resistors recovers the missing signals.

A demonstration of the potential for such techniques is provided in Fig. 9.7, which shows the comparison in terms of area occupation and power dissipation between classical flash converters (full parallel) and the converter shown in Fig. 9.6c, which uses folding and interpolation techniques (new system). You can verify the obtained reduced area from the layouts. The more compact realization allows minimizing the jitter problems because the lesser distance between comparators reduces the timing uncertainty for the clock distribution. Fig. 9.7 shows the compromises between power dissipation and performance of the flash converters with a "normal" speed. The performance is represented by the effective number of bits as a function of input frequency. The power dissipation is a parameter. With the same performance, lesser number of comparators allows a great power reduction.

The folding technique is a great solution for the high speed converter made th a bipolar technology: this technique allows obtaining a sampling frequency equal to 650MHz with a resolution of 8bit and power dissipation of 850mW. Statistical surveys show that the current manufacturing tolerance degrees for the transistor preclude the possibility of obtaining more effective resolution of 10bit with the folding technique.

About the GaAs technology, it is particularly useful to make a flash converter with wide bandwidth: the MESFET technology allows the realization of S&H stages with very high performance, for example with a slew-rate of 4.5kV/μs and jitter limited to only 3ps. About the resolution, the problem related to the GaAs technology does not allow overcoming the 5bit for high speed converters.

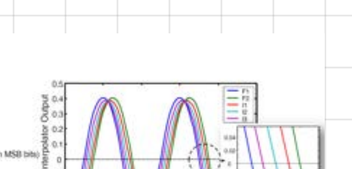
Fig. 9.5.

…ple of a double folding circuit, in order to avoid non idealities of folded edges

… other folding shapes can be used…

(not asked at the exam)

(B) implementation of triangle wave →



Example of how to generate folds:

- via source-coupled pairs
- with $V_{ref1} < V_{ref2} < V_{ref3} < V_{ref4}$
- as V_{in} changes, only one of M1, M3, M5, M7 is on

- Actually, curves are rounded, ← due to saturation
- therefore, accurate only at zero-crossings
- In fact, most folding ADCs do not use the folds, but only the zero-crossings!

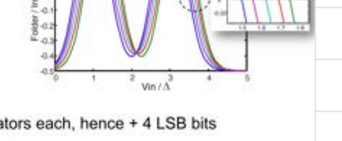
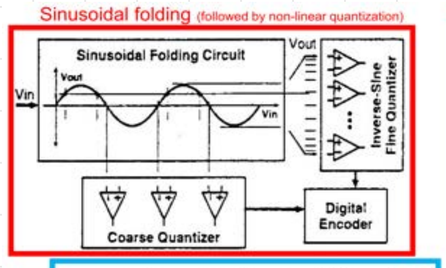
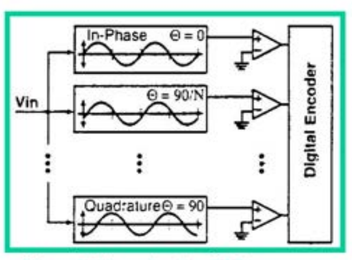


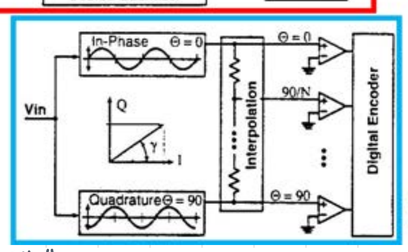
Fig. 9.6 (not asked)

Sinusoidal folding (followed by non-linear quantization)



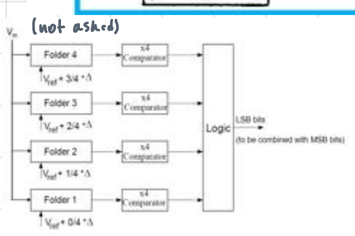


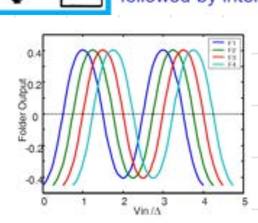
Many folding circuits (256), one for each quantization level



Few (2) folding circuits, followed by interpolation

(not asked)





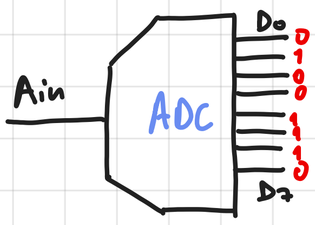
(not asked)

Example of 4 folders with 4 folds each, so 16 zero-crossings, hence + 4 LSB bits … anywat upper limit… due to added complexity.

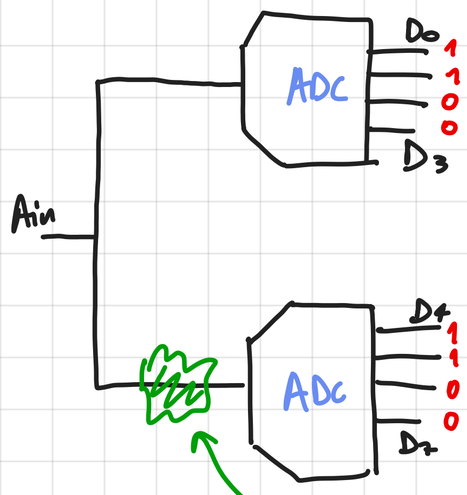
Example of 4 folders with 4 resistor interpolators each, hence + 4 LSB bits … other ideas?

Ex. Folding ADC

- The idea: Imagine to have:



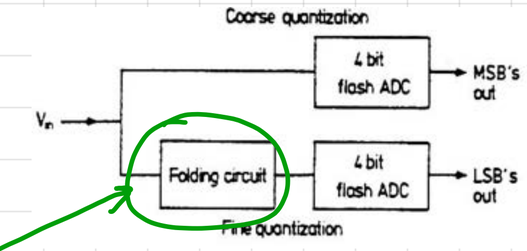
vs.

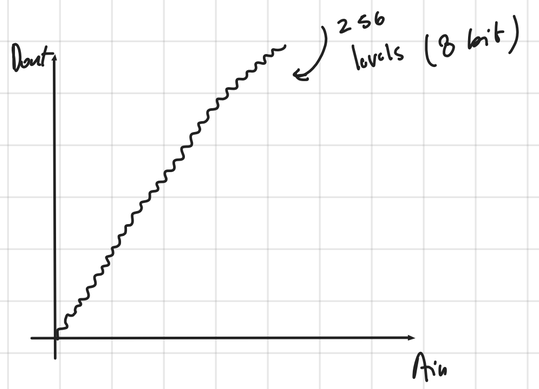


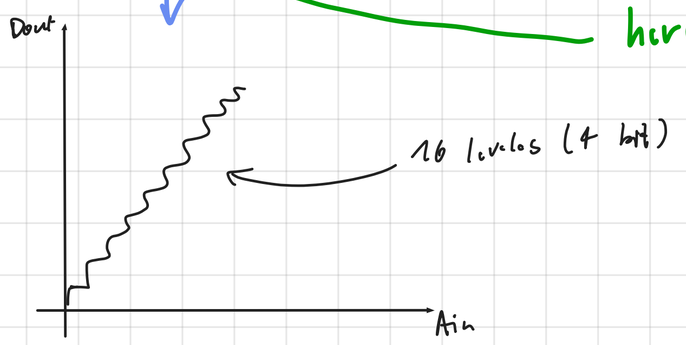
BUT different from this

Identical conversion

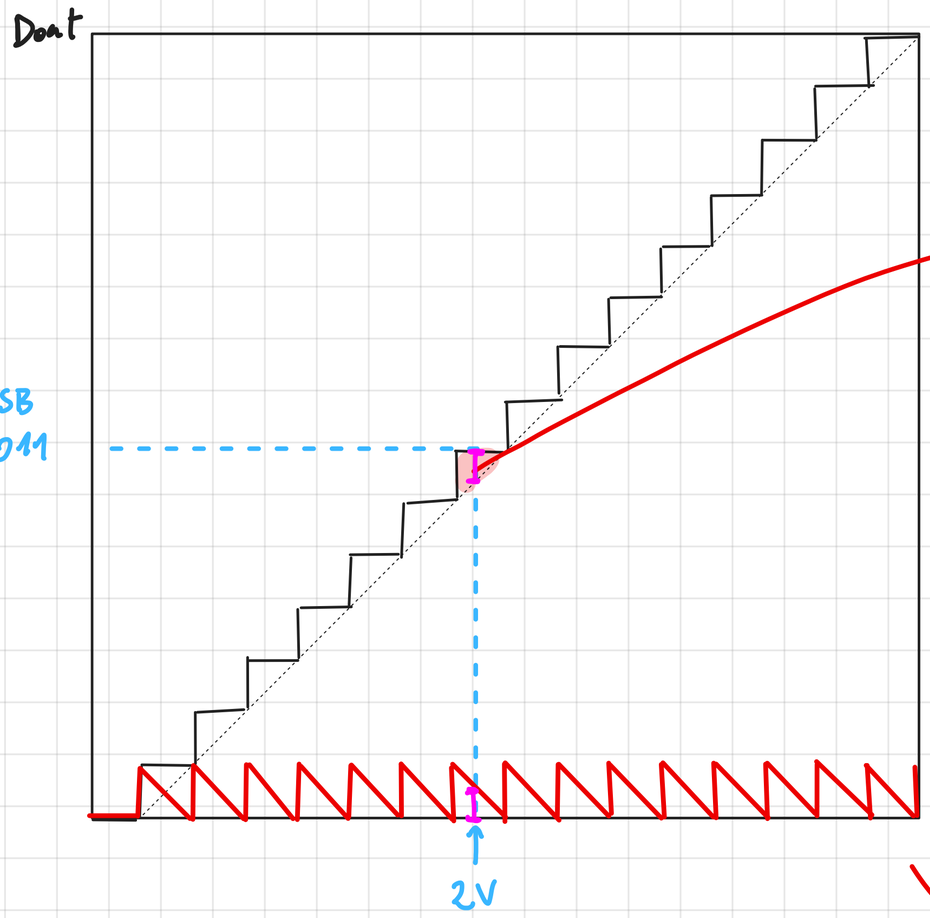
We should add something here







↳ We should consider the all range of the ideal I/O curve with 16 levels:

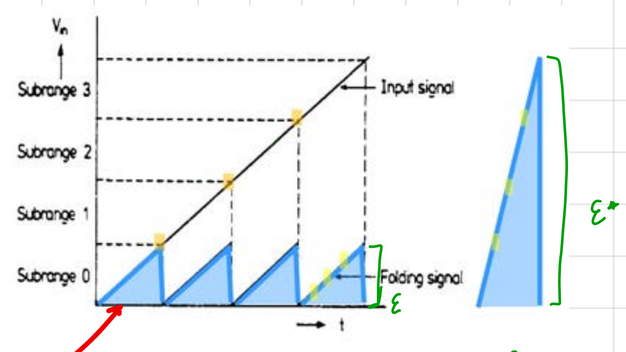


error due to quantization

↳ we need a circuit that predict it ← folding circuit

returns the residual error due to quantization

error curve



amplifying the error range $\varepsilon \rightarrow \varepsilon^*$

↳ Now consider:

We saw that the folding circuit returns the residual error

We can propose the following Folding ADC:

Half-Flash-ADC

In the preceding paragraph, it is clear that the flash technique is very prohibitive for resolution above 10bit. In many applications, we need 10÷12bit of resolution with a sampling frequency above Msps, ensuring low power consumption. These specifications are not simply reconcilable with flash architecture. Moreover, we should prefer the CMOS technology, instead of the bipolar, to integrate the ADC with the signal processing stages, for example for video applications.

The "half-flash" technology (noticed as "two-step flash") allows significantly reducing the number of components, the power dissipation, and the input capacitance. In half-flash converters, the conversion is made with a cascade of two flash steps; a complete conversion therefore requires two steps.

Fig. 9.8 is a schematization of a common half-flash N bits converter. During the first step, the signal is applied to the first flash converter, which made a *low resolution conversion* (coarse quantization) obtaining the M MSBs. These are converted by a DAC, whose accuracy must be at least equal to that of the whole ADC, i.e. N bits. This value, which is an approximation of the input signal, is subtracted from the same input signal, and the result is the *residual*, i.e. the error committed approximating the signal with M bits. The residual is amplified and sent to the second flash that determines the L LSBs making the *fine quantization*. The output of the converters are summed, and the result is the N bits digital output. By and large, we desire to satisfy M+L>N in order to perform the correction of any conversion error; the additional bits are named "overlap bits". The S&H stage is needed because a significant input variation between a conversion phase and another could bring non-reliable results.

The advantage of this configuration is evident. Consider a 12bits ADC: the flash realization should require 4096 comparators; instead, with the half-flash technique (as for example those of Fig. 9.9), we use $2^7+2^6=192$ comparators, with an overlap of 1bit. With this resolution, we obtain a great reduction in terms of use of area and power dissipation; moreover, the input signal sees capacitance very much lower than that of an equal resolution flash ADC. The main drawback is the half conversion time because a complete conversion requires two clock cycles.

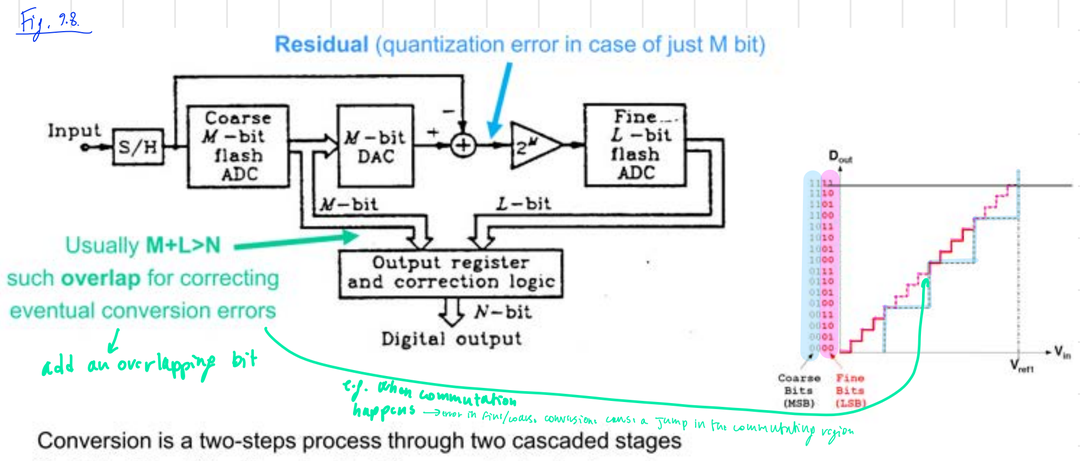
For example, we can mention two Analog Devices 10bits ADCs, fabricated in TTl bipolar technology: the AD9020 allows a sampling frequency of 60Msps with power consumption of 3W while the half-flash AD9040 can sample up to 40Msps with consumption of only 900mW.

As a typical half-flash example, consider the ADC depicted in Fig. 9.9 with the clock signal temporization; it is made in CMOS technology and has a resolution of 12bit, a sampling frequency of 5Msps, and power dissipation of only 200mW. The operation is controlled by two clock signals, Φ_1 and Φ_2. During the phase "sampling/offset cancellation", Φ_1 and Φ_2 are both high: an S&H samples the analog signal, the comparators for both the flash stages are in the "offset storage" phase, and the DAC and the adder are reset. When Φ_1 goes down, the first stage comparators are "strobed", i.e. produce the corresponding thermometric code for the MSBs while the second stage comparators are in the "tracking" phase, i.e. have the output of the adder as an input. At this point, the conversion is made. When Φ_2 goes down, also the second stage comparators are in the "strobe" phase and made the fine quantization.

Such an operation is typical of the CMOS switched capacitors (SC) implementation. Fig. 9.10 shows the basic structure of a "charge balancing" SC comparator. The clocks Φ_1 and Φ_2 work in a complementary way: when Φ_1 is high, the capacitance C_1 is charged with the tension $V_{i2}-V_{ov}$, and the parasitic capacitance C_p is charged to $V_{ov}\cdot A/(A_0+1)\approx V_{ov}$. During the second phase, Φ_2 is high, and the comparator works in an open-loop: the capacitance C_p is charged with the tension $V_{ov}-(V_{i2}-V_{i1})\cdot C_1/(C_1+C_p)$, the output voltage therefore depends only on $(V_{i1}-V_{i2})$, not on V_{ov}.

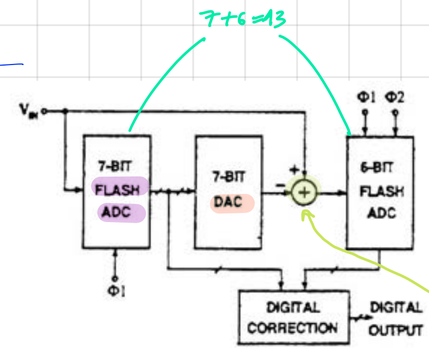
The Fig. 9.12 shows the simplified structure of a part of the ADC. The thermometric code produced by the first flash directly controls the DAC that is constituted by an equal value capacitance network. During the reset phase, the switch Sp is ON while the output of the comparators are held to zero. In the following phase, Sp is OFF while, if the thermometric code has a value n (i.e. if it has a number n of "1"), the first n capacitors will receive a voltage V_{ref}, and the remaining N–n capacitors will be held to the ground, so the DAC output voltage will be $V_o=V_{ref}\cdot n/N$. Because the errors of the first flash are minimized by the digital error correction, the whole ADC linearity is determined firstly by the DAC, which must be designed accurately. In this case, the DAC configuration ensures the monotonicity while the measured accuracy is effectively 12bit. The voltage V_o reaches the adder amplifier, made with the SC technique, whose structure is similar to that of the comparator shown in Fig. 9.10. During the reset phase, the node D0 (Fig. 9.12) is discharged to the ground while S_1 and S_3 are ON. During the conversion phase, S_1 and S_3 are open while S_2 is closed, so the adder will produce a voltage depending on the difference between V_{in} and V_o. The amplification of this difference depends on the ratio between the capacitance values: in this case, $C_1=C_2=C_3$ holds, and therefore the residual $V_{in}-V_o$ is not amplified. In fact, a typical residual amplification of a factor 2 or 4 will proportionally decrease the resolution required by the second flash, but will involve a reduction in terms of speed and linearity of the adder. In this realization, we privilege the speed because this stage is slower than the whole converter.

Fig. 9.8



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

Fig. 9.9



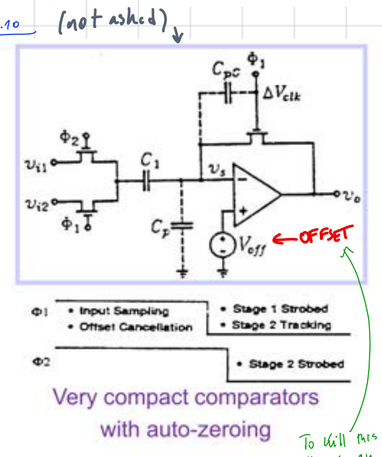
PERFORMANCE OF A/D CONVERTER

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)

the switch Sp is ON while the output of the comparators are held to zero. In the following phase, Sp is OFF while, if the thermometric code has a value n (i.e. if it has a number n of "1"), the first n capacitors will receive a voltage V_{ref}, and the remaining N–n capacitors will be held to the ground, so the DAC output voltage will be $V_o=V_{ref}\cdot n/N$. Because the errors of the first flash are minimized by the digital error correction, the whole ADC linearity is determined firstly by the DAC, which must be designed accurately. In this case, the DAC configuration ensures the monotonicity while the measured accuracy is effectively 12bit. The voltage V_o reaches the adder amplifier, made with the SC technique, whose structure is similar to that of the comparator shown in Fig. 9.10. During the reset phase, the node D0 (Fig. 9.12) is discharged to the ground while S_1 and S_3 are ON. During the conversion phase, S_1 and S_3 are open while S_2 is closed, so the adder will produce a voltage depending on the difference between V_{in} and V_o. The amplification of this difference depends on the ratio between the capacitance values: in this case, $C_1=C_2=C_3$ holds, and therefore the residual $V_{in}-V_o$ is not amplified. In fact, a typical residual amplification of a factor 2 or 4 will proportionally decrease the resolution required by the second flash, but will involve a reduction in terms of speed and linearity of the adder. In this realization, we privilege the speed because this stage is slower than the whole converter.

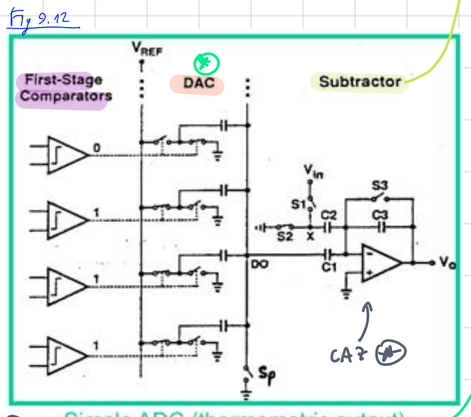
Fig. 9.10 (not asked)



Very compact comparators with auto-zeroing

To this one we use an autozeroing Amplifier (CA2) (seen at lesson ESO6)

Fig. 9.12



Simple ADC (thermometric output) and minimal DAC (charge sharing)

We directly use the thermometric code (without using an encoder to convert it to binary) to drive the following DAC

Compared to flash ADCs, half-flash ADCs minimize area occupation and power dissipation, though with slower conversion time

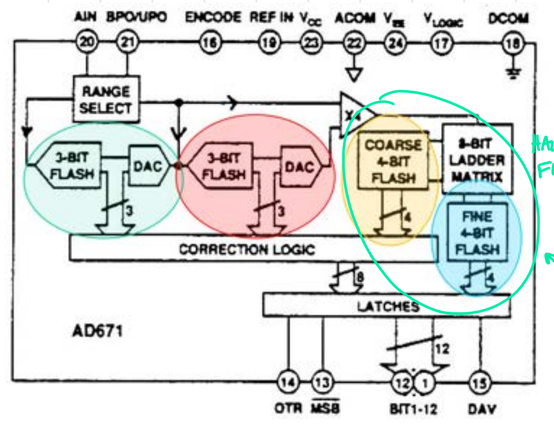
→ these solutions take a lot of conversion time

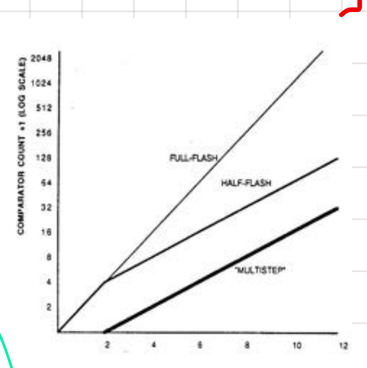
Multistep ADC

To complete the scene, we can say that there are solutions with a cascade of more than two flash stages, in those that is indicated as "multistep" conversion. Fig. 9.16 shows the scheme for a typical commercial product of a 12bit multistep converter, 2Msps in BiCMOS (AD671). In this case, the conversion is made in 4 steps: the first and second flash have 3bits while 8bits are produced by means of a half-flash which has 2 4bits stages; in this way, we employ only 48 comparators.

Fig. 9.17 makes a comparison as regards the architecture treated until now, showing the number of comparators needed as a function of the desired resolution. About the conversion time, we can assume that it is directly proportional to the number of stages.

Fig. 9.16





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 !

we can improve it using pipeline ADC
=
multistep + S&H before each stage

(conversion time improves of a factor m = # stages (ts&H))

Pipelined ADC

We see that the half-flash and multi-steps configurations allow, in respect of flash ADCs, greatly minimizing the occupation of area and the power dissipation, at the expense of the conversion speed. Some applications require, at the same time, high resolution and a great sampling frequency. For example, in the field of video applications, the HDTV devices need converters with 10bits and a sampling frequency up to 75Msps. A quick comparison of the architecture allows making the following comments:

- 10bit and 75Msps with flash technique means great area consumption, very high power dissipation (2.8W), and absolute need for bipolar technology;
- 10bit with half-flash technique brings to a maximum frequency of 40Msps and consumption of 1W using the bipolar technology while, with CMOS, we obtain a lower frequency, in the order of few Msps.

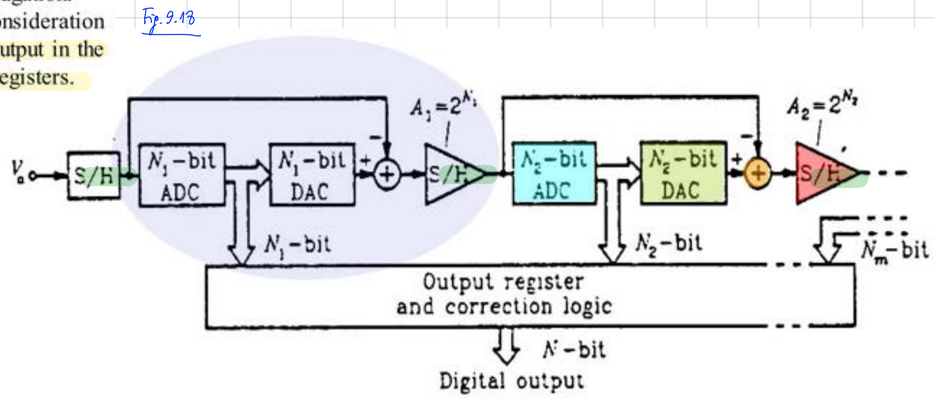
We have to consider that, in some applications, the resolution required by the ADC tends to increase; for example, the technological improvements in the field of the image sensors (i.e. CCD) will lead to design a converter with 12÷13bits with low power dissipation. The demand for a high sampling frequency, low area occupation, and moderate power consumption is fully satisfied by using a *pipelined* converter.

Fig. 9.18 shows the structure of a typical pipelined ADC. It is composed of *m* stages, each of which contains an S&H amplifier, a low resolution flash ADC (typically with resolution of 1÷4bit), a DAC, and an analog adder. In respect of the multistage structure, in the pipelined structure, there is an S&H between a stage and the following; in this way, after the conversion of the i-th stage, the residual which reaches the (i+1)-th stage is maintained by the S&H.

In doing so, the i-th stage can reach a new sample. In a pipelined ADC with *m* stages, the sampling frequency is equal to the clock frequency while every conversion requires *m* clock cycles, and therefore the code for a certain sample appears at the output after a certain delay named the *latency* or *pipeline delay*. This delay, in many applications, is not a problem. For example, in television or telephonic broadcasts, an additional delay is not important if compared with the delay attributable to the signal propagation.

About the error correction and the overlap bits, the same consideration made in the preceding paragraph is valid. To obtain a parallel output in the pipelined ADC, delay lines are needed; these are made with shift registers.

Fig. 9.18



m-stages each one composed by:

ADC Flash (low resolution 1÷4bit), DAC, analog adder and amplifying S&H

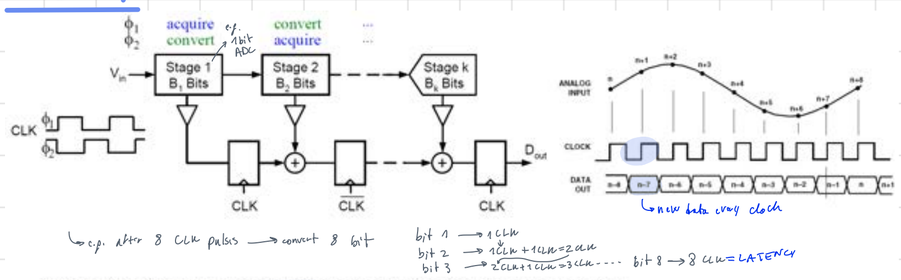
Compared to multistep ADC, now S&Hs allow parallel-pipelined processing (like bucket-brigade)

During the design phase, we need to properly choose the number of stages, and it is made on the basis of different considerations. As an initial analysis, we can say that increasing the number of stages there means:

- minimizing the hardware: to this end, the maximum saving is obtained with *n* stages with 1bit, which means that only *n* comparators are needed;
- greatly reducing the input capacitance, which allows reducing the load of the analog circuits, firstly of the S&H and, therefore, increasing the speed;
- reducing the amplification of the residual between one stage and the following one and, thus making the amplifier's response faster because the GBWP is set by the technology;
- deteriorating the accuracy because, stage after stage, the noise increases.

The compromises between speed and accuracy and the considerations on area occupation and power consumption determine the choice of the number of stages.

Latency



New output data every clock, but a latency of 8 clock cycles

Need for data alignment

Digital self-calibration techniques are suitable to the CMOS and BiCMOS technologies, which are more complex correction algorithms. For example, in a commercial 15bits pipelined ADC, made by 17 1bit stages, the DAC and residual amplifier errors (due to the charge injection, the comparator offset, and the capacitance mismatches) determine a non-ideal relationship between the input voltage and the output voltage for each stage, i.e. the received residual and the residual produced by the stage. To self-calibrate, during the calibration period (which requires, in this case, around 70ms), on-chip electronics analyze the in/out characteristic of each stage, and the parameters are stored in a small memory to be used by the calibration algorithm during the normal operations. The converter has a DNL limited to ±¼LSB.

Time-interleaved ADC

To obtain a high sampling frequency, we can use two or more ADCs connected in parallel, obtaining a structure named the *"time interleaved array"*. As shown in Fig. 9.25, this structure is made of C identical converters with N bits, each of which proceeded by an S&H. In every "channel", the input signal is sampled with a frequency equal to 1/(C·T), where T is the clock period and C is the number of channels. The channels work sequentially, thus the sampling frequency of the whole ADC is n times greater than that of the single converter whereas resolution and the maximum frequency of the input signal are unchanged.

This technique can be used to attain a conversion speed non-achievable with single converters, or when it is advantageous to obtain conversion speed by employing a series of slower converters rather than use a more complex single converter.

The main problems that the time interleaved technique has are the different channels behavior and the non-regularity of the sampling intervals. For example, if the channels have a different offset, it can happen that, with a constant input, every channel produces a different code.

In the frequency domain, this offset error occurs with the presence of spurious frequencies at the multiples of f_S/C (Fig. 9.26, at the top). If, instead, the channels have a different gain, the spurious frequencies will appear:

$$f_S/C \pm f_{in}\ 2\cdot f_S/C \pm f_{in} \ldots (C-1)\cdot f_S/C \pm f_{in}$$

as shown in Fig. 9.26 (in the center). The irregularities of the sampling intervals are due to systematic timing errors between channels (skew), which produces spurious components whose frequencies are the same as those due to the offset error, but its amplitude grows with the input signal frequency (Fig. 9.26, at the bottom).

Consider that the 8 bit have all different gains.

So consider the input signal and its conversion phases:

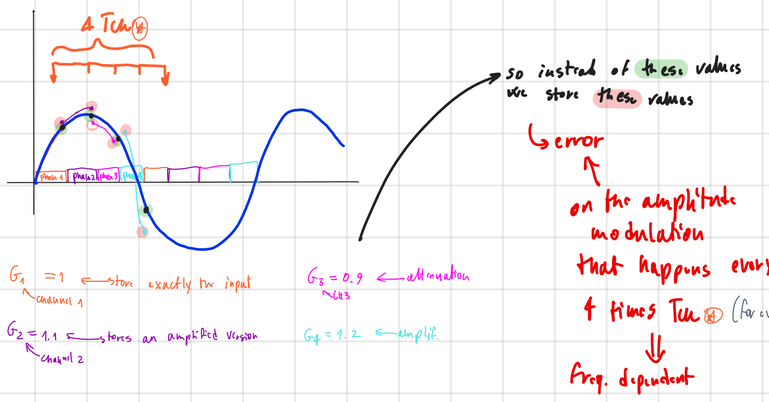


Fig. 9.20 shows the timing graph for the converter. When T&H1 tracks the input signal V_{in}, the comparators of ADC1 are in the tracking phase, too; therefore, their outputs, which control the DAC switches, change the input signal response, and, consequently, also the adder output (i.e. the residual) is modified. The residual variation is neglected by T&H2, which is in the hold phase. When the clock is low, T&H1 switch is in the hold phase, and, because ADC1 determines an approximation limited to only 4bits of the sampled signal, the first stage comparators receive the latch signal (indicated with, "regenerate") after only 1ns; i.e. when the T&H1 output reaches the final value with an approximation of 4bits and therefore long before it reaches an 10bits approximation. The DAC and T&H1 outputs continue the adjustments during the following 4ns while T&H2 is in the track phase, for which it follows the residual up to the final value. When the clock switches, T&H2 is in the hold phase while ADC2 tracks the signal.

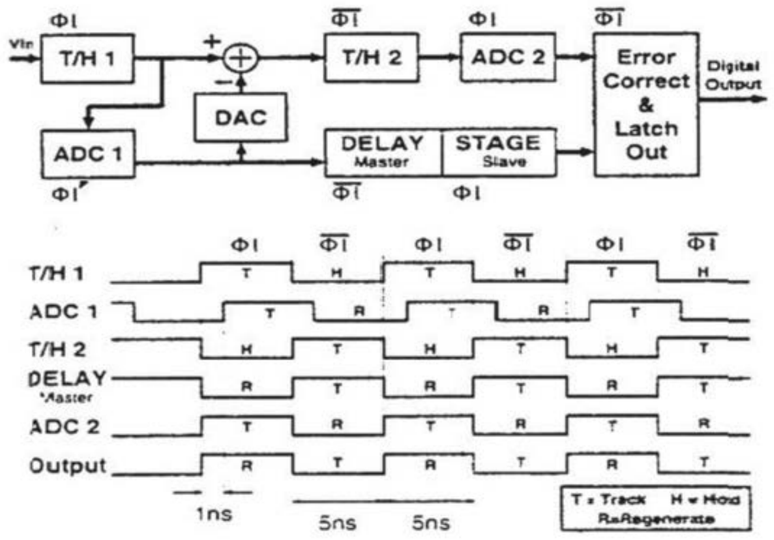


Fig. 9.20: *Pipelined ADC timing graph.*

For example:
Nbit pipelined ADC
with N stages of 1bit each

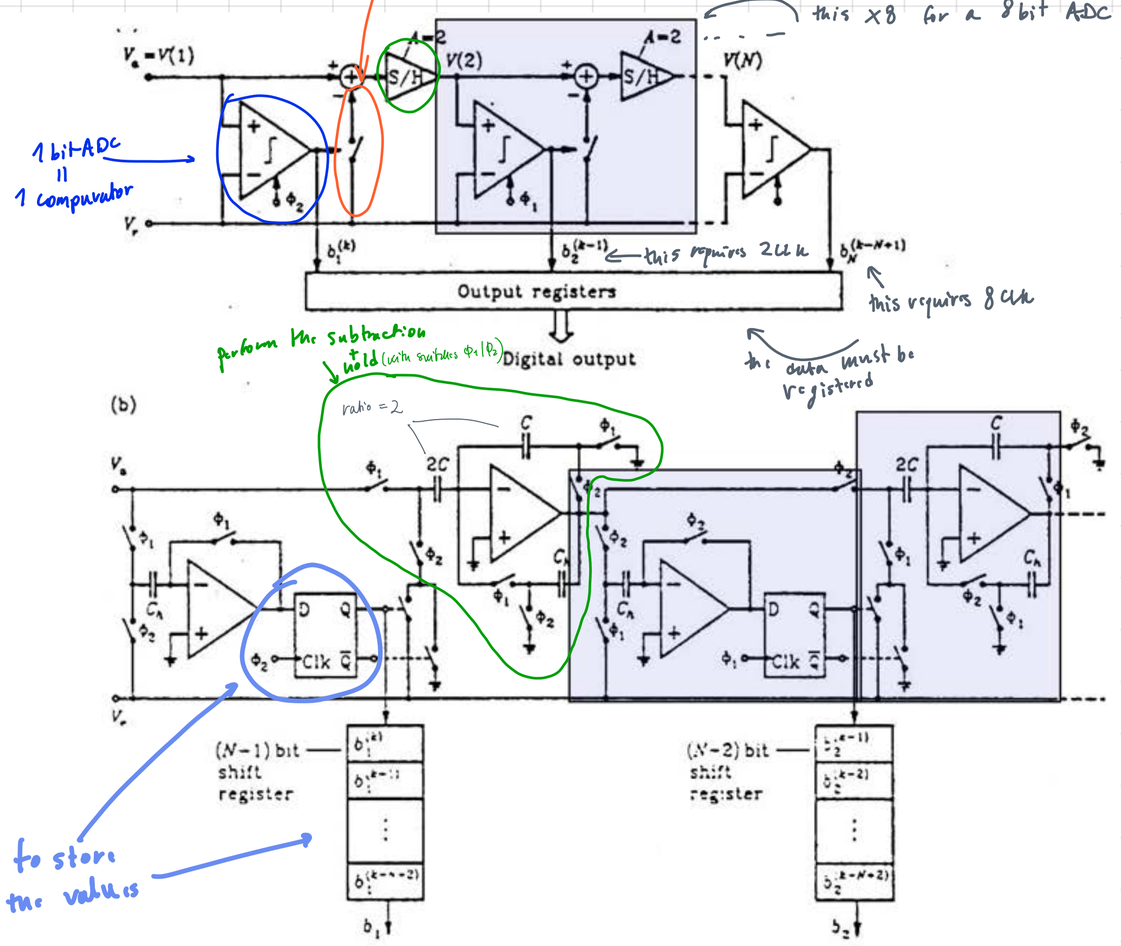
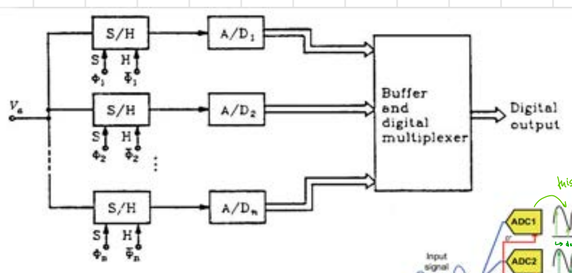
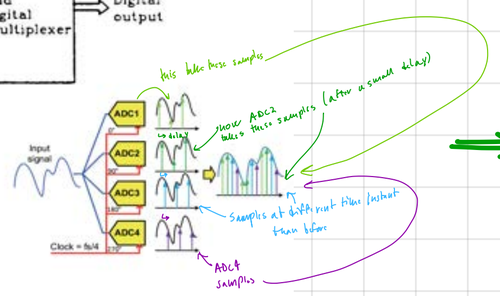


Fig. 9.25



For high f_S more ADC can run in parallel on different samples:

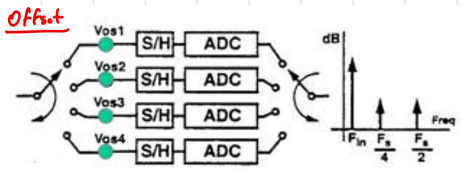
- C channels with identical Nbit ADCs each, with individual S&H
- Each channel samples the input signal at a rate $(C\cdot T_{clock})^{-1}$
- The channel operate in sequence
- Overall sampling rate is N-fold that of a single channel



⇒ higher throughput (faster)

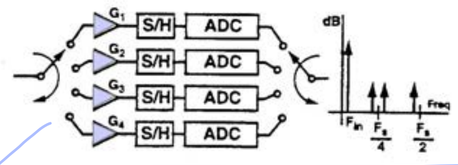
Fig. 9.26. **Errors**

- Offset

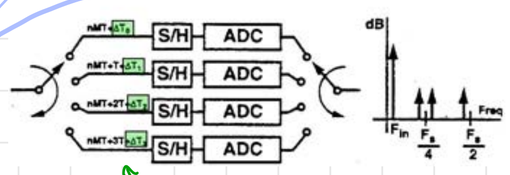


Offsets:
spurious tones at multiples of f_S/C

↳ the offset are DC errors
↳ but we have freq. dep. errors on the output



Gains:
spurious tones at frequencies :
$f_S/C \pm f_{in}$ $2\cdot f_S/C \pm f_{in}$ … $(C-1)\cdot f_S/C \pm f_{in}$

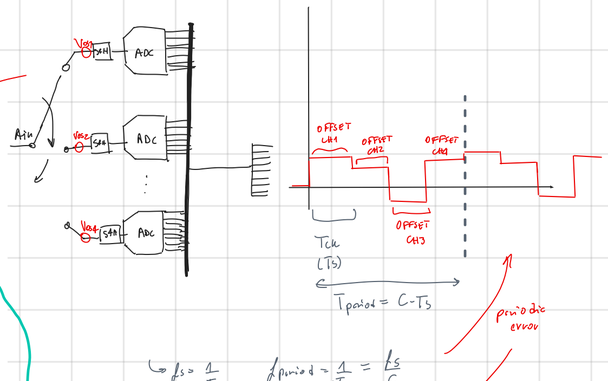


Sampling time:
spurious tones of intensity increasing with f_{in}

when we open/close

also this cause an error that is freq. dep.

↳ all these errors are usually not possible to filter out



↳ If the number of channel C is high the fundamental tones f_S/C is very low, otherwise the fundamental freq. is high and can be witheld by the error filter

(END OF THEORY!)

Fig. 9.21a shows the conceptual scheme for a pipelined ADC with N bits, made with N stages with 1bit. Its operation can be described by the following algorithm:

$$V(i+1)=2\cdot[V(i)-b_i\cdot V_r] \quad \text{for } i=1, 2, \ldots N$$

with: $V(1)=V_a>0$; $V_r=(V_a)_{max}/2$; $b_i=1$ if $V(i)>0$; $b_i=0$ if $V(i)<0$.

Fig. 9.21b shows the practical realization which uses the switched capacitor technique, whereby the switching operations of comparators and OpAmps implicitly made also the signal sampling and, therefore, contains the S&H, which cannot be implemented as, instead, is necessary with the bipolar technology.

The main limitation of the pipeline ADC accuracy made with MOS technology consists of capacitance values' coupling error (mismatches) and charge injection due to parasitic capacitance, as well as the comparators offset. Fig. 9.22 shows the possible problems due to the non-ideality of the block which made the pipelined ADC. For example, the comparators offset reduction techniques cannot remove the DAC non-linearity effects or the gain error in the residual amplification; these errors are the main contributions to the ADC non-linearity, supposing that they are not corrected by means of appropriate calibration techniques.

The current trend is to use cheaper and more effective calibration techniques particularly because the trimming process or the use of sophisticated technologies with high precision significantly increases the cost. For example, the realization of capacitors with high precision can be nullified by the parasitic capacitors introduced by the package in which the chip is enclosed.

In CMOS technology, it is easy to use self-calibration techniques, which act when the device works with the actual conditions. We can, for example, use adjustable capacitors, whose adjustment is entrusted to internal calibration logic. The adjustment occurs when the device is on; in such a case, the parameters computed by the calibration logic will suffer from some variations during the device operations. In other cases, the calibration is periodically made during the normal operations: it is the case of video application converters, in which the calibration logic intervenes in between a frame and the following.