NUMERO : 13 DATA : 26/01/2011

A P P U N T I

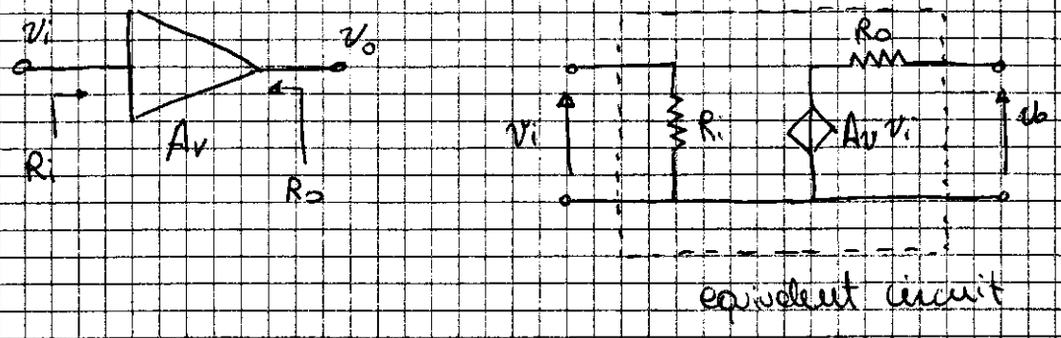
STUDENTE : CAUSA Bruno

MATERIA : Electronics For Embedded System

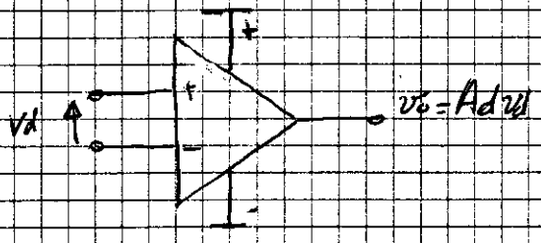
Prof. Del Corso - Gregoretti

Operational Amplifiers

- Amplifier



- Operational Amplifier



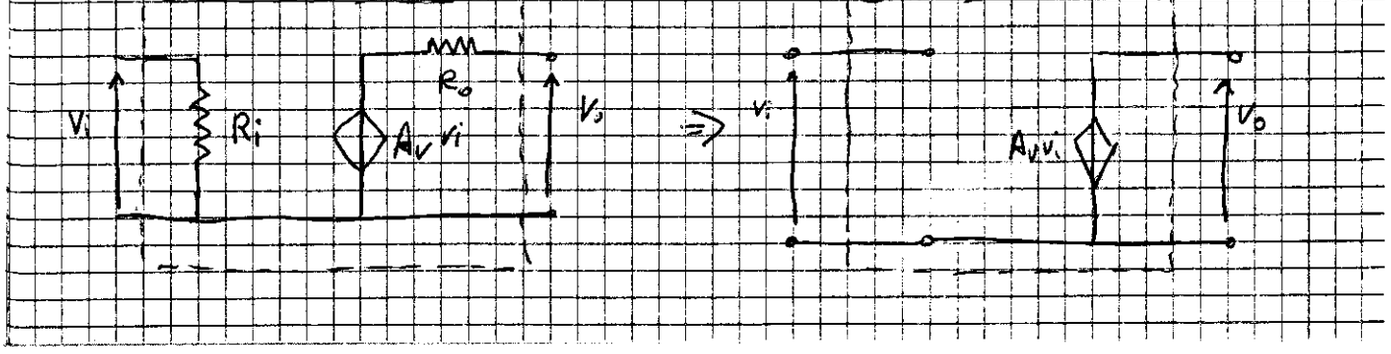
(Usually the power supply is differential (+5V, -5V for example)

Ideal assumptions

$$\boxed{\begin{aligned} A_d &\to +\infty \\ i_\oplus &= i_\ominus = 0 \end{aligned}}$$

In order to have $V_o \neq +\infty, \emptyset$ we have to feedback the op. amp.; so $V_d = 0$.

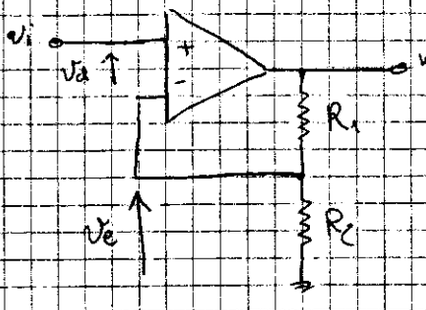
The equivalent circuit is



Real Amplifiers

The study of real operational amplifiers comes from the elimination of the two assumptions $A_d \to \infty$ and $i_\oplus = i_\ominus = 0$ initially; we can studies the effect of each of the two eliminations without considering the other, that means that the amplifier is considered ideal except from:

- differential gain: **A_d finite**



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

$$v_i = v_d + v_e = v_d + \beta v_o$$

In the ideal case $v_d = 0$ and so

$$v_o = \frac{1}{\beta} v_i = \frac{R_1 + R_2}{R_2} v_i \quad \text{, as seen before.}$$

In the real case instead we have $v_d = \frac{v_o}{A_d}$

$$v_o = \frac{1}{\frac{1}{A_d} + \beta} v_i = \frac{1}{\beta} \frac{1}{1 + \frac{1}{A_d \beta}} v_i$$

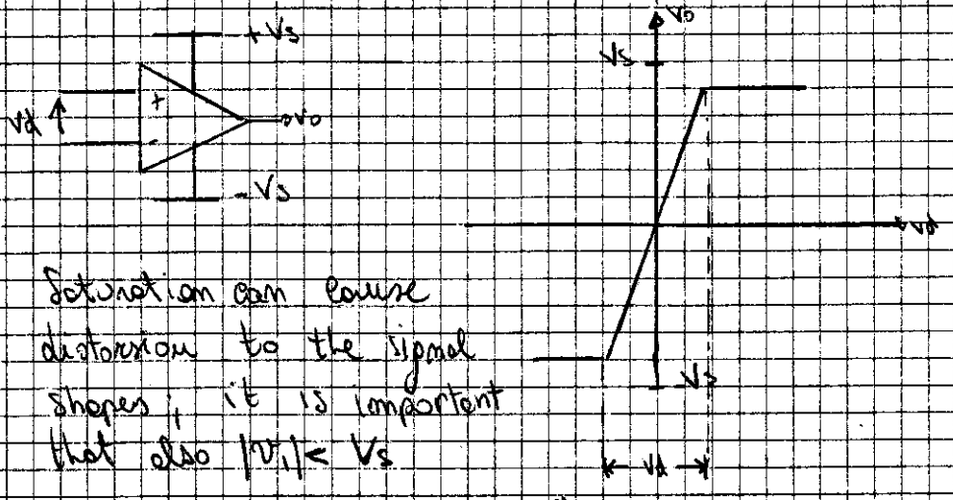
So, the higher is the product $A_d \beta = T$, less is the error that we commit assuming $A_d \to \infty$

$$A_{v\,real} = \frac{1}{1 + \frac{1}{A_d \beta}} \cdot A_{v\,ideal} \simeq A_{v\,ideal} \left(1 - \frac{1}{T} \cdots\right)$$

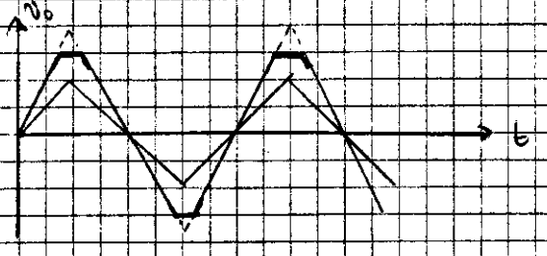
From the last formula we can evaluate the error made using the ideal formula; it is called gain error and is

$$\varepsilon_{Gain} = A_{v\,ideal} \cdot \frac{1}{T}$$

- saturation: $\boxed{-V_s < v_o < V_s \quad \forall v_d}$

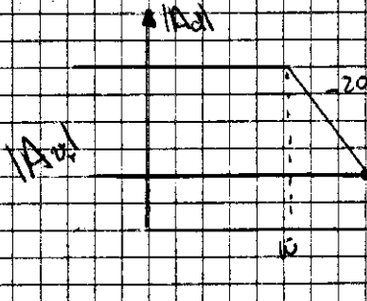


Saturation can cause distortion to the signal shapes; it is important that also $|v_i| < V_s$



The power supply can also be a simple supply with the other pin connected to ground; because in this case the excursion is lower, so it is easier to have saturation, particular amplifiers are built and they are called RAIL TO RAIL op. amps.

- frequency response: $\boxed{|A_d|_{dB} = f(\omega)}$



for $\omega < \bar{\omega}$ $A_{v_{real}} \simeq \frac{1}{\beta}$

for $\bar{\omega} \div \omega_t$ A_d is a low value

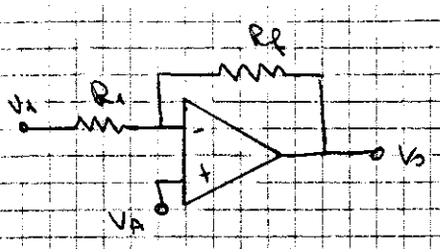
so $A_{v_{real}} = \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{A_d \beta}} \simeq$

$\simeq \frac{1}{\beta} \frac{1}{1/A_d\beta} \rightarrow A_d$

GBP (Gain × Bandwidth product) is the same for all the points in the slope segment.

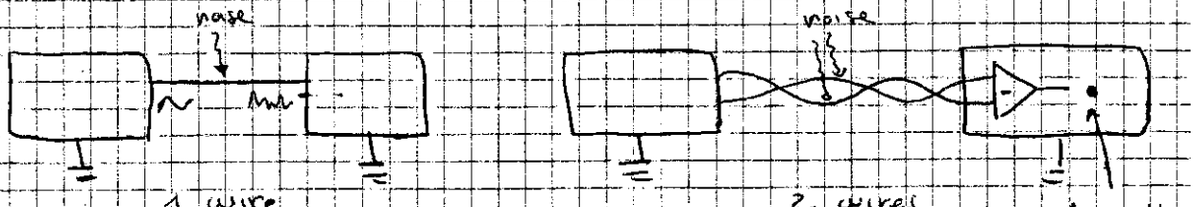
If I need more bandwidth, I have to change my operational amplifier.

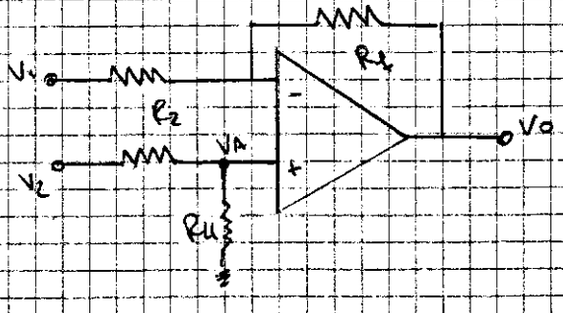
Subtractor



$$V_o = -V_1 \frac{R_f}{R_1} + \left(1 + \frac{R_f}{R_1}\right) V_A$$

But maybe we want $V_o = K(V_1 - V_2)$; differential signals are often used because they reject common mode noise.





$$V_A = \frac{R_4}{R_2 + R_4} V_2$$

$$V_o = -V_1 \frac{R_1}{R_1} + \left(1 + \frac{R_f}{R_1}\right) \frac{R_4}{R_2 + R_4} V_2$$

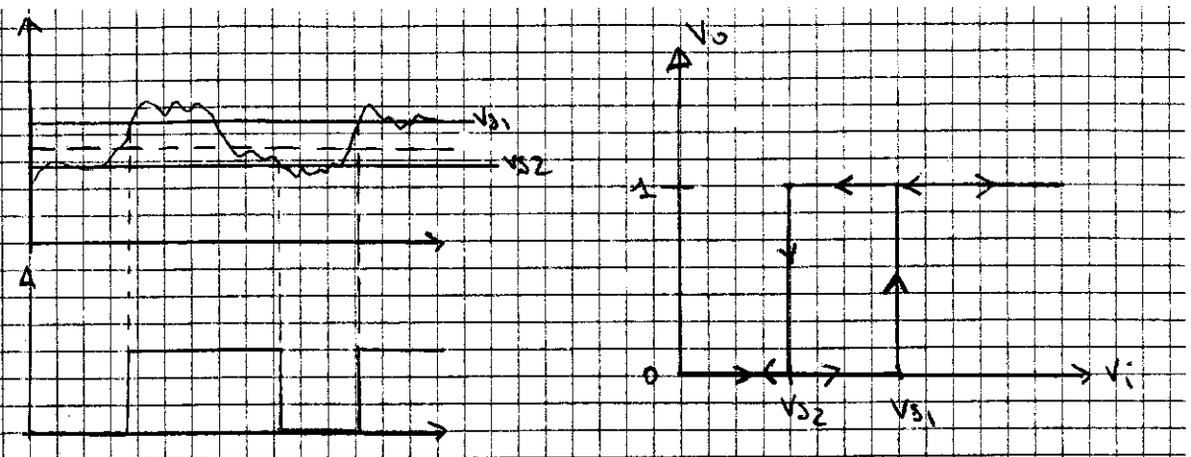
In order to have a perfect differential signal, it must be:

$$\frac{R_f}{R_1} = \left(1 + \frac{R_f}{R_1}\right) \cdot \frac{R_4}{R_2 + R_4}$$

$$\frac{R_2 + R_4}{R_4} \cdot \frac{R_f}{R_1} = \frac{R_f}{R_1}\left(1 + \frac{R_1}{R_f}\right)$$

$$\frac{R_2 + R_4}{R_4} = \frac{R_f + R_1}{R_f}$$

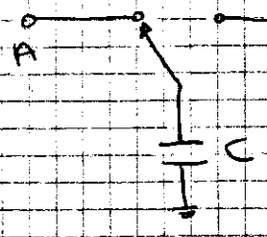
$$\longrightarrow \begin{cases} R_f = R_4 \\ R_2 = R_1 \end{cases}$$



Explanation of the hysteresis curve: let's assume that the initial value of the output is low - 0. It won't change until $v_i = V_{s1}$; when $v_i = V_{s1}$ the output will switch to logic 1, if the input is still increasing, the output will continue to be 1; if the output instead is decreasing, it has to reach the value $v_i = V_{s2}$ in order to switch to the logic 0.

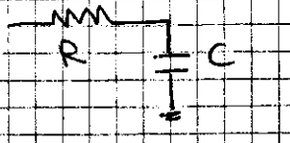
This comparator is also called SCHMITT TRIGGER: the feedback (positive!) allows the output to change value in just few nanoseconds (while the normal comparator needs μs).

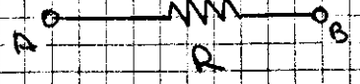
③ Switched capacitor SC

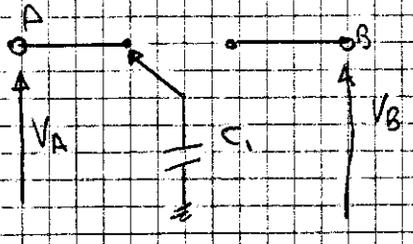


They are actually the most used, are made of a capacitor and a switch, easily implemented with a transistor.

Let's have a look at a simple 1st order filter with a resistance and a capacitor. We would like to replace the resistance with a switched capacitor.







Ⓐ Charge of the capacitor: $Q = V_A C_1$

Ⓑ Discharge of the capacitor: $Q = V_B C_1$

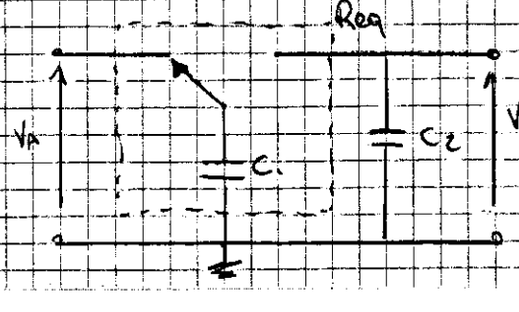
$$\Delta Q = C_1 (V_A - V_B)$$

Assuming f as the switching frequency

$$I = \Delta Q \, f = C_1 f (V_A - V_B)$$

$$\underline{\underline{R_{eq} = \frac{V_A - V_B}{I} = \frac{1}{C_1 f}}}$$

So now we can use the switched capacitor instead of the resistance.

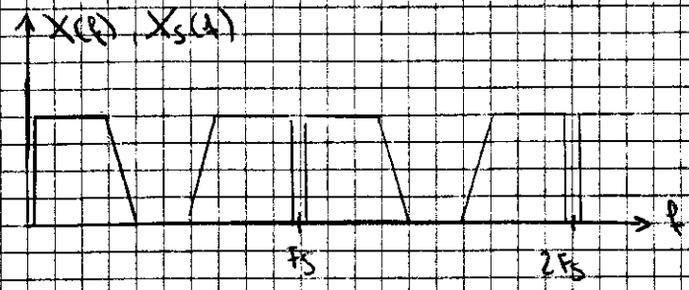


The time constant of the circuit is

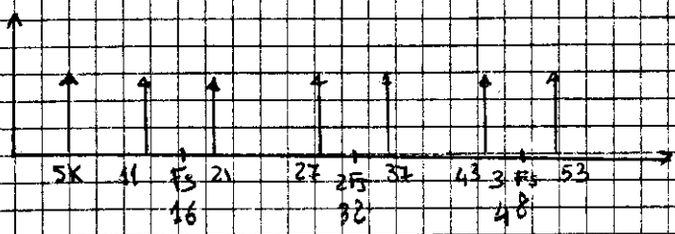
$$\tau = R_{eq} C_2 = \frac{1}{f} \frac{C_2}{C_1}$$

The ratio of the capacitors is easily precise: this is the main advantage!

In the frequency domain the sampling cause a repetion of the spectrum, so:



The sine wave of the previous page (1st picture) - f = 5KHz, with a fs = 16 KHz, would have this spectrum:



Working on the frequency domain, in order to reconstruct the original signal after a sampling, we need a pass-band filter that includes only the original spectrum. So it is necessary that the repetition of the spectrum doesn't effect the original one. This is the Shannon theorem:

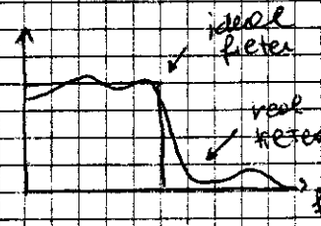
$$F_s \geq 2\,BW$$

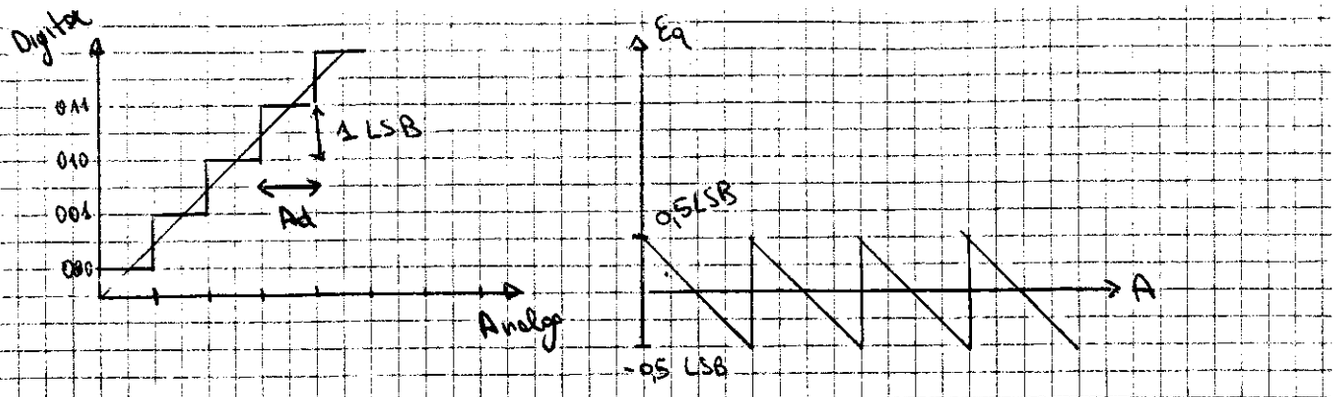
where BW is the bandwidth of the signal.

This ideal situation has two problems:

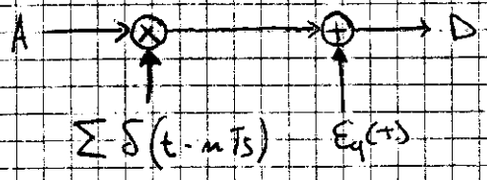
1) the filter is supposed to be ideal, while in the real case the filter cannot be instantaneously $\emptyset$.

2) a finite-frequency spectrum means an infinite time signal; in the real case instead the time signal is limited and so the spectrum is unlimited, that means that there is always a little quantity of aliasing.
In order to reduce this second effect usually a



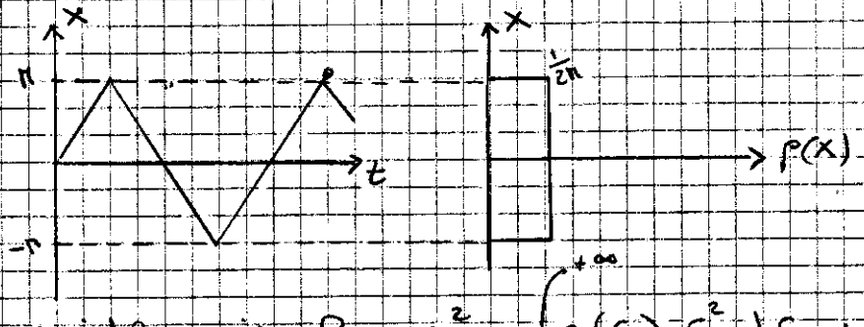


Quantization can be seen as a noise that is added to the sampled signal



We can define the SNR_q due to the quantization error:

$$SNR_q = \frac{P_s}{P_{\varepsilon_q}}$$



The power of noise, that is a random variable, is $P_{\varepsilon_q} = \sigma^2_{\varepsilon_q} = \int_{-\infty}^{+\infty} p(\varepsilon_q)\, \varepsilon_q^2\, d\varepsilon_q$ [illegible]

In the picture is shown the Probability Density Function (PDF) of a triangular wave, like $\varepsilon_q(A)$ is, assuming $H = Ad/2$ we have:

$$P_{\varepsilon_q} = \sigma^2_{\varepsilon_q} = \int_{-\infty}^{+\infty} p(\varepsilon_q)\, \varepsilon_q^2\, d\varepsilon_q = \int_{-\frac{Ad}{2}}^{+\frac{Ad}{2}} \frac{1}{Ad}\, \varepsilon_q^2\, d\varepsilon_q =$$

$$= \frac{1}{Ad}\left[\frac{\varepsilon_q^3}{3}\right]_{-\frac{Ad}{2}}^{\frac{Ad}{2}} = \frac{1}{3Ad}\left(\frac{Ad^3}{8} + \frac{Ad^3}{8}\right) =$$

$$= \frac{1}{3Ad}\,\frac{Ad^3}{4} = \frac{Ad^2}{12}$$

Errors due to: amplifier, filter, S&H, A/D; the total error is the sum of this errors, SNR_T

Convention (due to sine wave):

$$SNR_q = N \cdot 6\ dB + 1{,}76\ dB$$

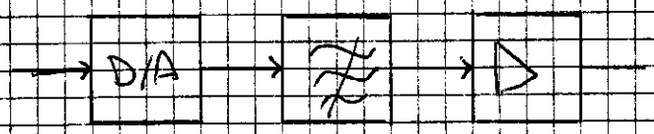
$$N = \frac{SNR_q - 1{,}76}{6}$$

If we substitute SNR_q with SNR_T we obtain the ENOB, equivalent number of useful bits:

$$ENOB = \frac{SNR_T - 1{,}76}{6}$$

D/A converters

The chain for a D/A conversion is:



Let's concentrate on the first block - D/A: we will see initially the other errors that occur in a DAC; they are linear errors, gain and offset, integral and differential nonlinearity, and dynamic parameters.

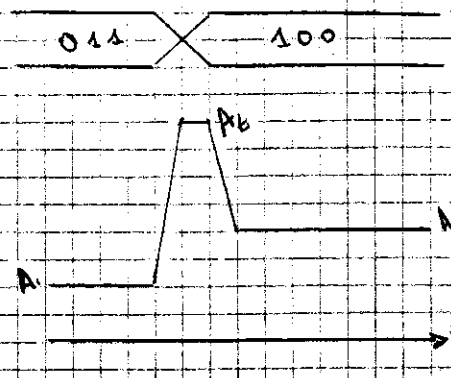
many close points forming a line.

A(D) is a sequence of dots

$N \to \infty$ becomes a line

$E_o \to$ OFFSET, compensated adding a constant

$E_g \to$ GAIN ERROR, difference among ideal and best approximation slopes, compensated by a gain correction

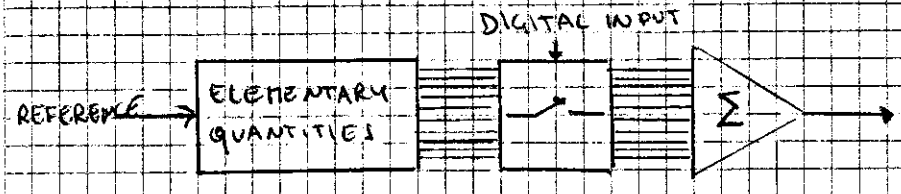


This phenomenon is called GLITCH.

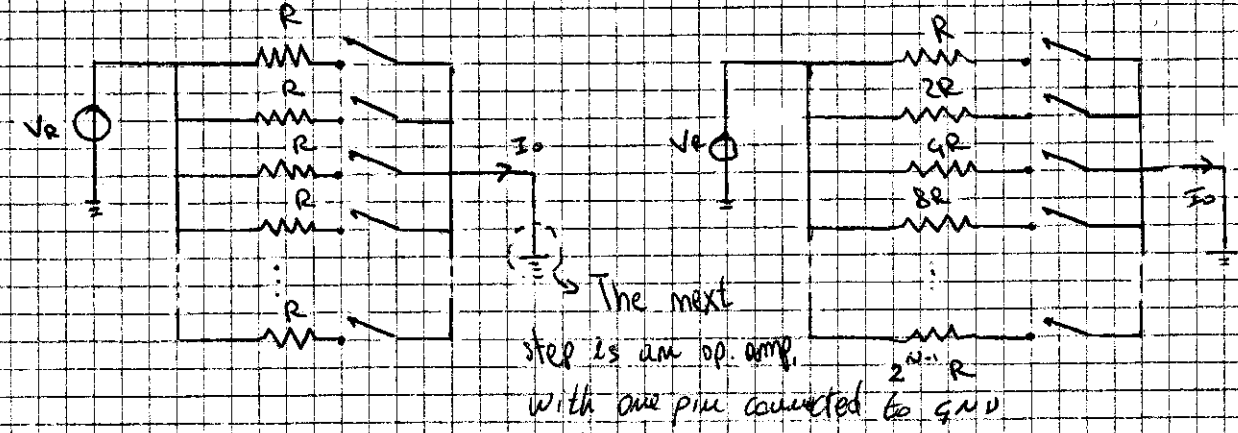
DAC error summary:

- linear { gain E_G ; offset E_O
- nonlinear { integral E_{INL} ; differential E_{DNL}
- dynamic { settling time T_S ; glitches

Let's see now how to build a DAC: it can be seen as the sum of elementary quantities controlled by D



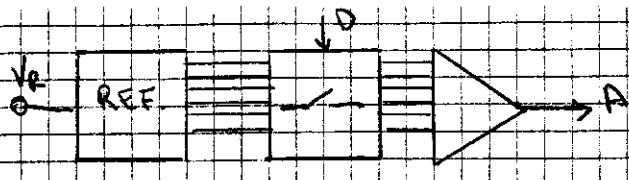
The elementary quantities can be all 1, and it is called UNIFORM CONVERSION, or powers of 2, and it is called WEIGHTED CONVERSION

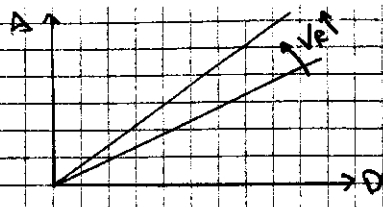


The next step is an op. amp. with one pin connected to GND

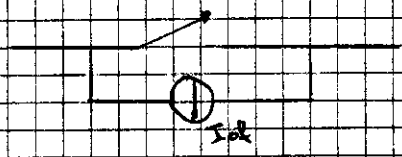
Assuming a current analog output, these are the electrical circuit of the two converters; the first one is the uniform converter: its problem is that if the digital value is N-bits, the uniform converter needs 2^N resistances, so it becomes very big! The second is the weighted converter: for an N-bits digital value it needs N weighted branches, that means N different values of resistors that must be precise!

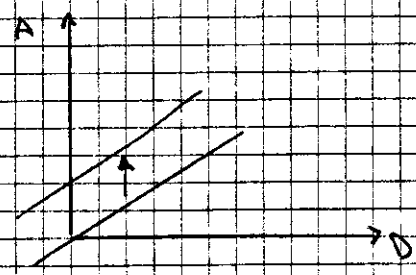
The circuits we draw have a switch that can switch between V_R and GND; this is called voltage switch; another opport





(1) The first error is a GAIN error that can occur when the value of V_R is not exactly the ideal value but has some percentage of error.

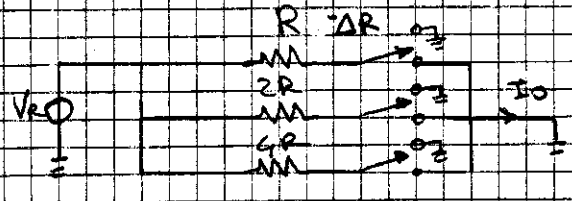


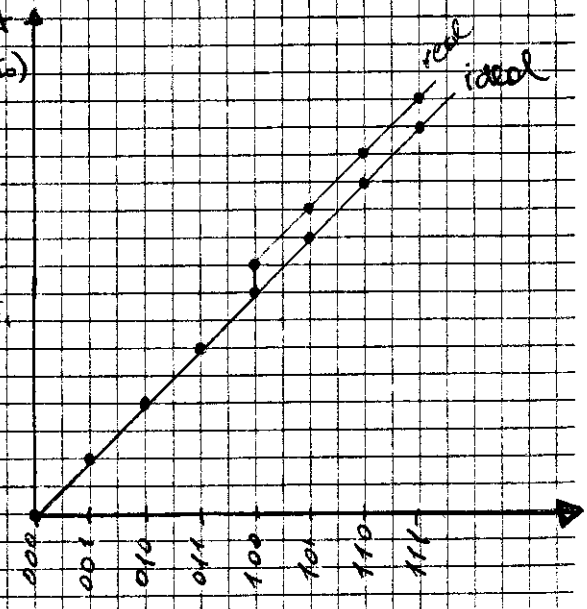


(2) OFFSET error is due to the offset of an eventual op. amp, but mostly the cause is the Leakage current of switches, that means that $I_{SWITCH} \neq 0$ even if it is OFF. This current depends mostly on temperature.

(1) Both GAIN and OFFSET don't depend on D value, so they can easily be corrected, once they've been detected

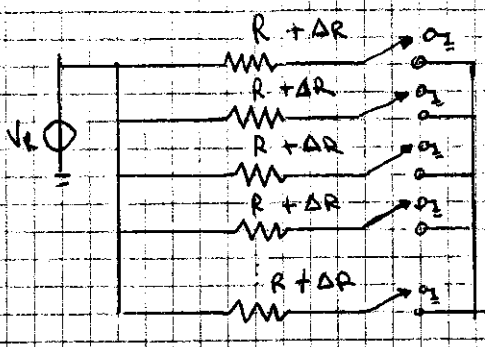
(3) Consider a weighted converter (current switches) - 3 bits:

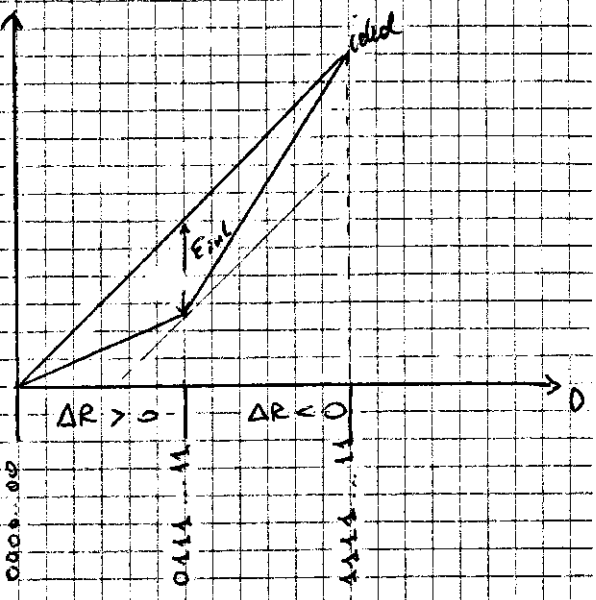




remember that the lower is the value of the resistor, the higher is the current, the more is significant the bit.
In this first example imagine that we have an error on R of $-\Delta R$; when the switch of that branch is on, that means when MSB is 1, the output will be higher.
Let's see what happens if the error is in the second OR

④ Consider now a uniform converter; imagine to have N resistors and that each resistor has an error because they are same resistors in the same circuit it is very likely that the error will be $\pm\Delta R$ for all of them. Imagine the first $N/2$ resistors have $\Delta R > 0$, the others $\Delta R < 0$





For the first half, the higher resistances will allow less current to pass, so lower values; in the second half the lower values will equilibrate. We can so find the E_{INL}.

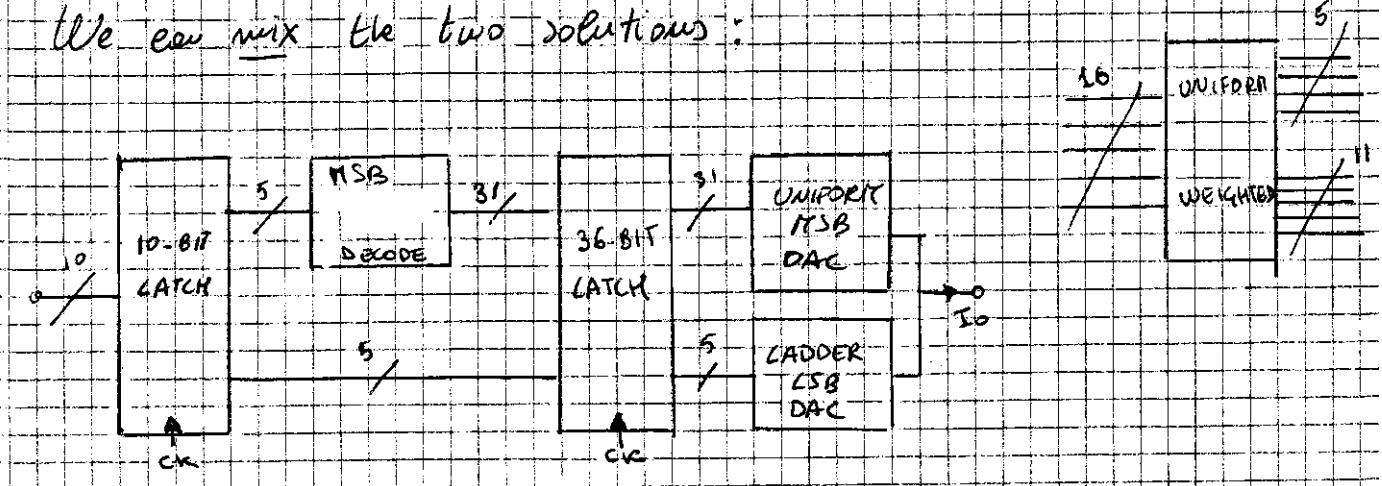
As the operation in an uniform DAC is always an adding of current (D increasing) it is impossible to have non-monotonicity.

So:

WEIGHTED: N bits → N branches
NON MONOTONICITY (HIGH PRECISION NEEDED)

UNIFORM: N bits → $2^N - 1$ branches
Always MONOTONE

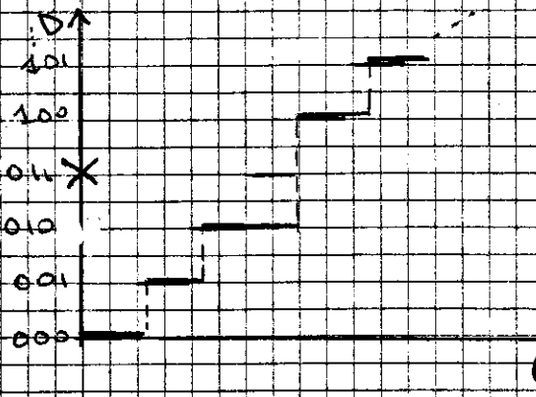
We can mix the two solutions:



The two latches are needed in order to synchronize all the bits and avoid GLITCHES. The MSB DECODE converts the value of the 5 bit in a 31-bit code for the switches of the uniform DAC.

→ E_{dnl} DIFFERENTIAL NONLINEARITY

In the ideal case each step is Δd long; in the real case we have $\Delta d \neq \Delta d'$, so the differential nonlinearity error is really $E_{dnl} = \Delta d - \Delta d'$

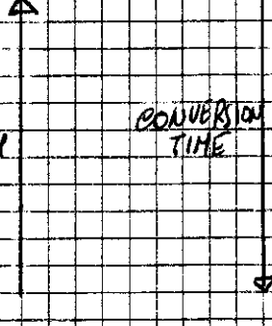


As is shown in the picture, if $E_{dnl} > 1$ LSB one code (011) will be lost. This error is called MISSING CODE ERROR and is the corresponding of the non-monotonicity error of the DAC.

In general the integral nonlinearity E_{inl} is how much the real function deviates from the ideal function, and is a unique value; the differential nonlinearity E_{dnl} instead is the difference point-to-point between the ideal behaviour and the real one, so is specific of each interval (but it has a max). The relation is:

$$E_{inl} = \sum E_{dnl}$$

As regards the dynamic parameters, just remember the CONVERSION TIME, that is the time from when the input is applyed, to when the output is correct, T_c; the conversion frequency is obviously $F_c = 1/T_c$.

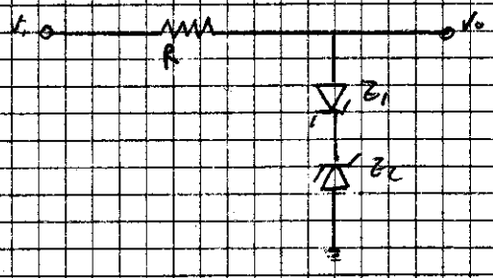


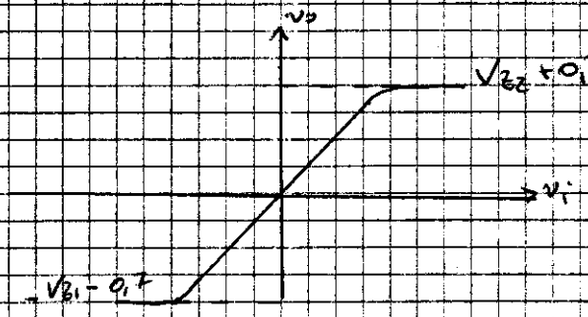
An ADC cannot be both simple and fast! To have a fast ADC is necessary a complex design, and if we want a simple ADC, it will be slower.

The complexity of an ADC is due to the number of voltage comparators; this is directly connected to the area and to the cost of the circuit. Remember that the cost of

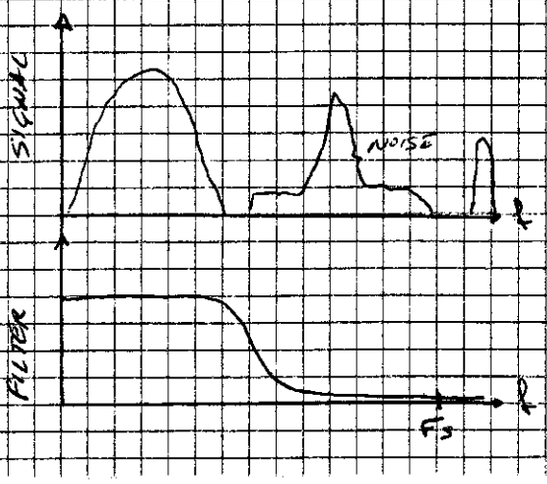
Signal conditioning

Here we have the conversion chain from analog to digital; in respect to the previous representations there is one more block that is the PROTECTION CIRCUIT; we will look first at this protection circuit and then to the part made of filters and amplifiers.

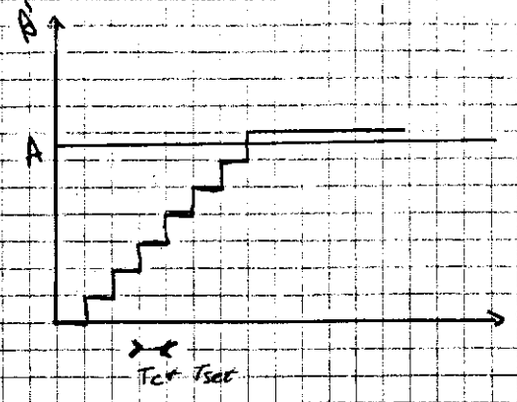




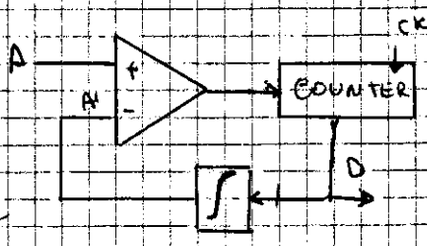
Here there is an example of protection called DOUBLE ANODE ZENER; the resistor is needed to limit input current (otherwise it would destroy also the diodes); remember that a zener diode has a voltage of 0,7 if is conducting in normal mode, and a voltage of $-V_z$ if an inverse current flow in it, when V_i is high z_1 is ON and z_2 is in the break down, vice versa if V_i is low.
The operation range depends always on the power supply; so with power off no signals can be inputed (or we will damage the circuit)



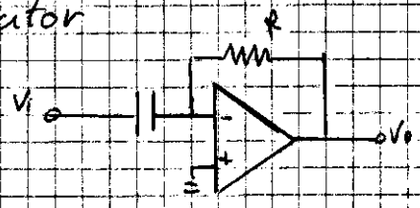
Let's talk now about the FILTERS: the analog signal MUST be filtered for two reasons: elimination of the noise at different frequencies and limited band to avoid ALIASING (it is called anti-aliasing filter).

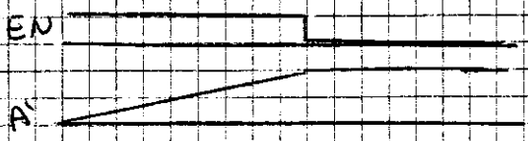


Another way is the RAMP CONVERTER; instead of the DAC there is an integrator, so while EN = 1 A' increases as a ramp.

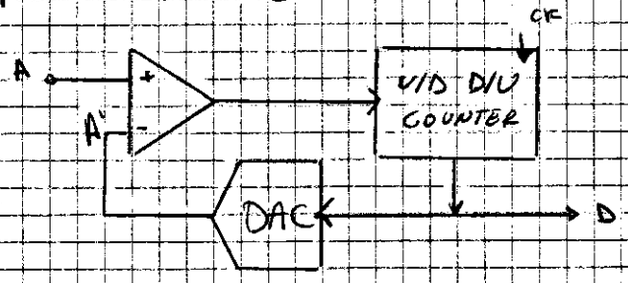


The integrator can be usually made as an operation amplifier with a resistence and a capacitor

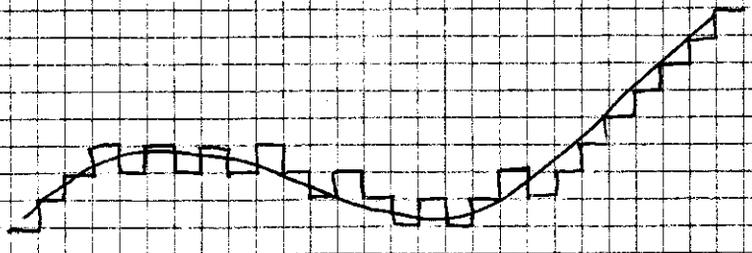




TRACKING ADC



The UP/DOWN - DOWN/UP counter allows to follow signal changing. Counter doesn't need to start from ZERO again. However in the worst case, that is at the beginning with the higher value, the time needed is 2^N cycle for N bits at most.



If the signal changes too quickly, A' cannot change quickly and there is an OVERLOAD.

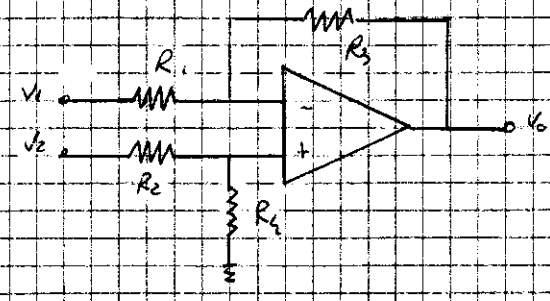
So all the FEEDBACK, the RAMP and the TRACKING AD converters have just one voltage comparator, so a complexity of 1, but a conversion time of 2^N for N bits. We will see a better converter that has just one converter

$$SNR_{TOT} = \frac{P_S}{P_N} = \frac{P_S}{P_a + P_q} = \frac{1}{\frac{P_a}{P_S} + \frac{P_q}{P_S}} = \frac{1}{\frac{1}{SNR_a} + \frac{1}{SNR_q}} =$$

$$\simeq \frac{1}{\frac{1}{4000} + \frac{1}{700}} = 595{,}7 = 55 \text{ dB}$$

$$ENOB = \frac{SNR_{TOT_{dB}} - 1{,}76}{6} = 8{,}87 \longrightarrow 9 \text{ bits are enough!}$$

We didn't consider the amplifier in the chain:



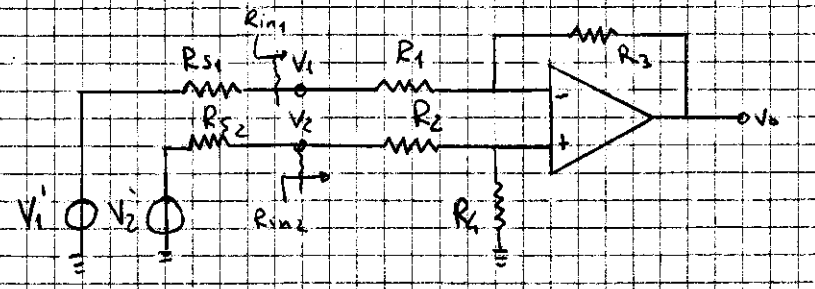
$$V_o = -\frac{R_3}{R_1} V_1 + \frac{R_4}{R_2 + R_4}\left(1 + \frac{R_3}{R_1}\right) V_2 \Rightarrow \frac{R_3}{R_1}(V_2 - V_1)$$

assuming that $R_1 = R_2$ and $R_3 = R_4$ that are the conditions to minimize the bias error because the equivalent resistances from the two inputs of the op.amps are equal.

$$V_o = A_d (V_1 - V_2) + A_{cm} \frac{V_1 + V_2}{2}$$

In our ideal circuit $A_d = -\frac{R_3}{R_1}$, $A_{cm} = 0$

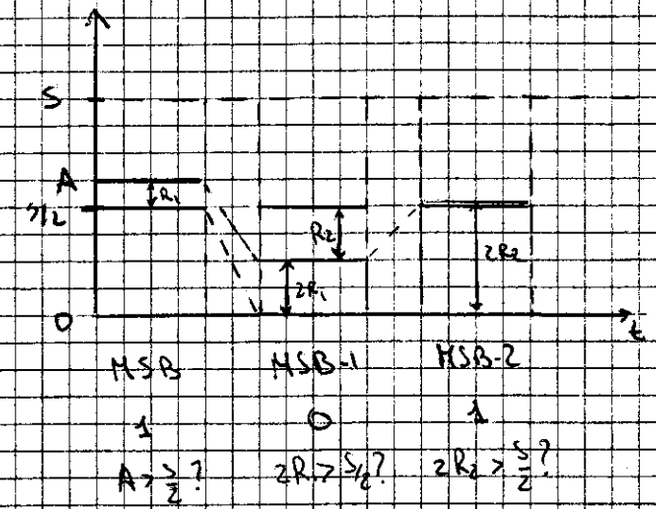
In the real circuit each voltage input has a series resistance:



$R_{in1} = R_1$

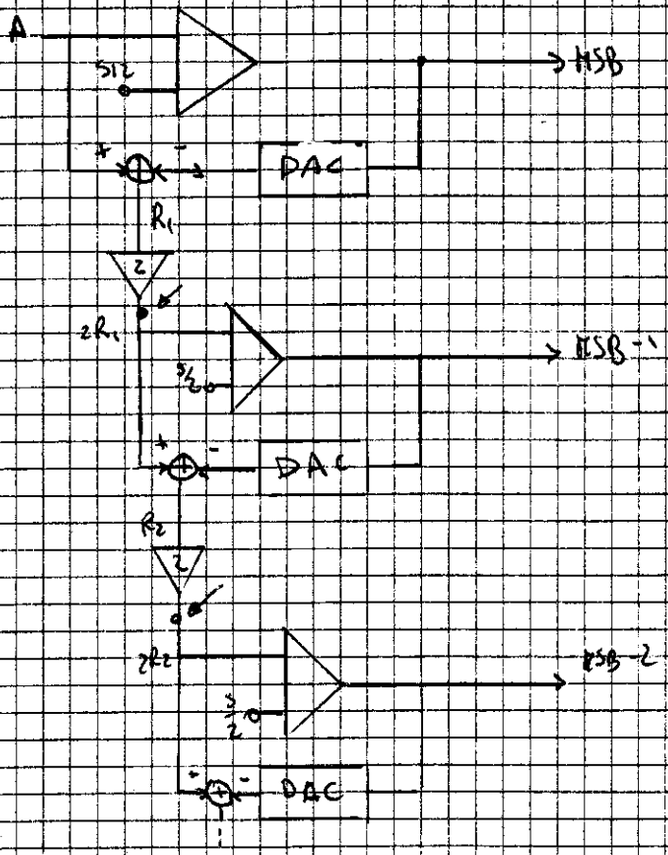
$R_{in2} = R_2 + R_4$

RESIDUE ADC



Each residue is multiplied by two and comparated with S/2 ; for each bit one comparator is needed, one 1-bit DAC to build the approximation, one analog adder to evaluate the residue, and one amplifier with gain 2.

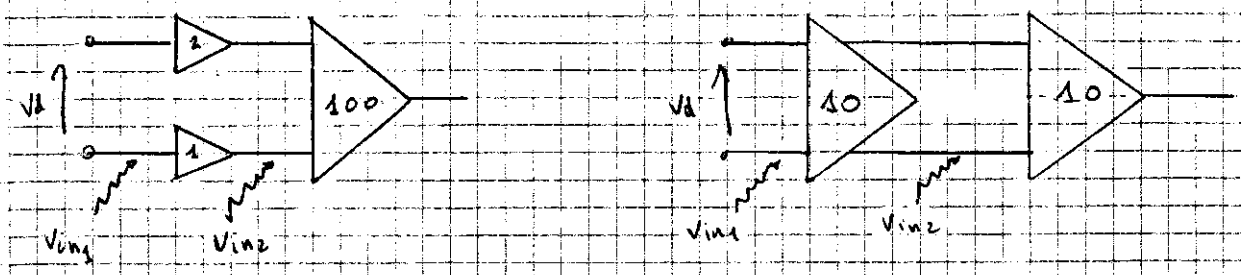
So it needs N comparators and has a conversion time of N ; no benefits in comparison with the successive approximation ADC.



This is the way how to build a RESIDUE ADC. Really it is useful only as starting structure for pipeline ADC. The idea of the pipeline is those that while a digital output is computed at step 2 (MSB-1), the previous circuit is free and you can send it the second analog value; so it operates on different samples at the same time

PIPELINE SUBRANGING ADC

The only difference with the residue ADC, from a structural point of view, is that some S/H are placed after the amplifiers (•). It uses N comparators and the con-



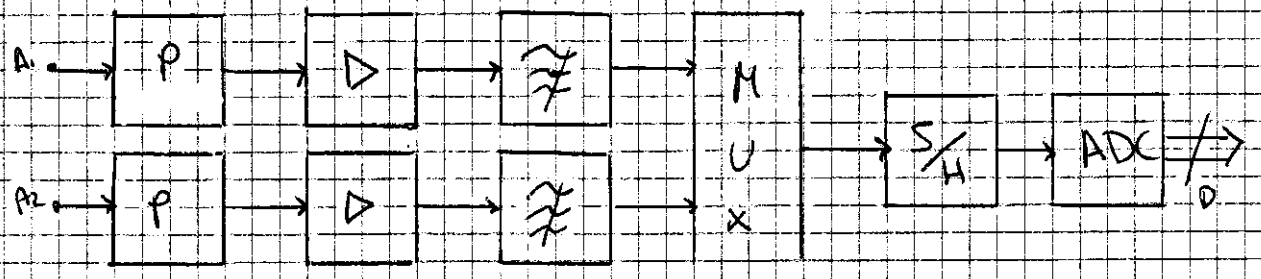
$V_o = A_d V_d$ gain $V_o = A_1 \cdot A_2 V_d = A_d V_d$

$V_{on_2} = V_{in_2} \cdot A_d$ offset error (V_{in_2}) $V_{on_2} = V_{in_2} A_2$

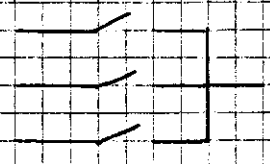
$V_{on_1} = V_{in_1} \cdot A_d$ noise error $V_{on_1} = V_{in_1} \cdot A_d$

As we can see, Offset error is less in the second choice because of the divided gain.

MULTIPLEXER



A multiplexer is useful to use S/H and ADC for several channels; it selects one of the N inputs and stops the others; so it is made of N switches; a switch is a MOS transistor that has an equivalent series resistance R_{on} and a leakage current I_{off}

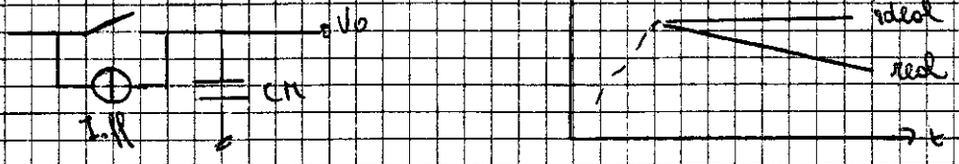


SAMPLE & HOLD

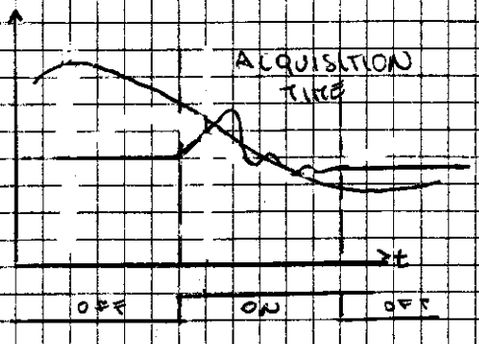
The sample & hold unit has two functions: it samples the analog signal at $t = kT_s$, $k = 1, 2 \ldots$, and keep the sample value at the output as long as required for A/D conversion.

In order to have a small pedestal error C_H should be high, but it would become slower. The smaller is C_H, the faster it is, the bigger is this error.

The switch has a leakage current I_{off} that causes an HOLD error:

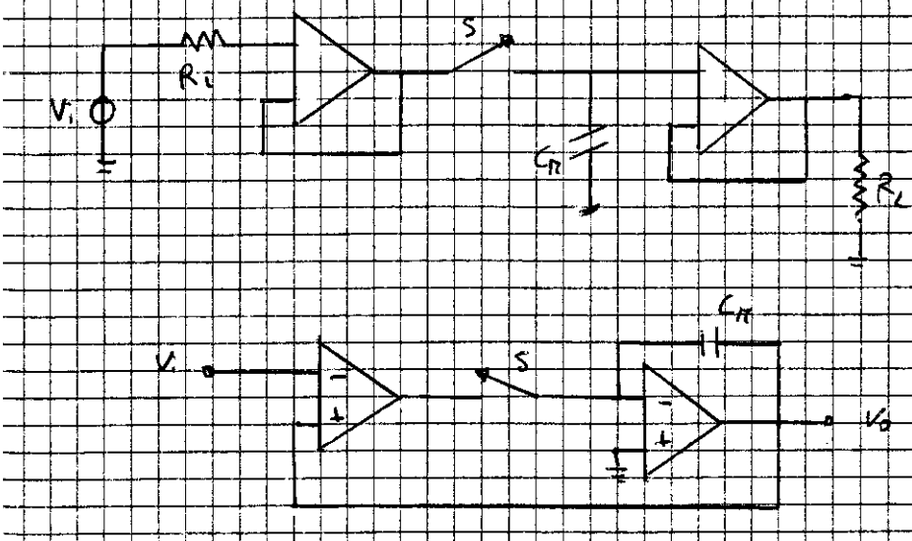


C_H should discharge only through R_L, but actually it discharge also through I_{off}



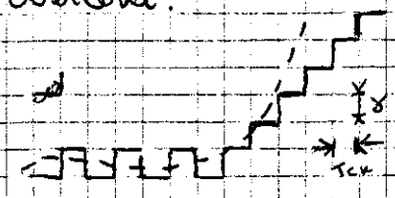
C_H needs some time to charge and thus cause the acquisition delay. In order to have a small acquisition time, C_H should be very small, but this would cause a worse hold error.

In order to solve these problems we can put two followers:



Coming back to converters, we will talk now about DIFFERENTIAL CONVERTERS: consider the tracking ADC; it is possible to obtain the output starting just from

Also, the variation of the signal cannot be too fast. otherwise the digital output cannot follow the input, and there is an overload.



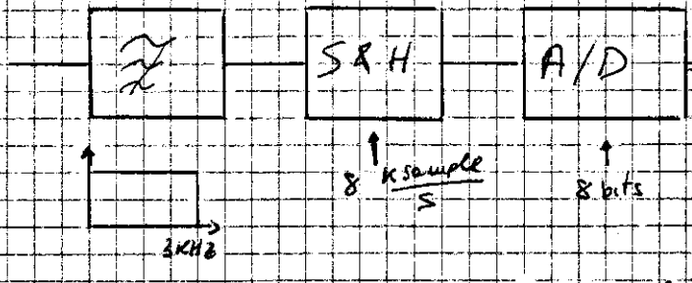
$$S_{R\,max} = \omega V < \frac{\delta}{T_{CK}}$$

$$\text{So}: \quad \frac{\delta}{2} < V < \frac{\delta}{\omega T_{CK}}$$

while for a standard converter is $\frac{E_q}{2} < V < \frac{S}{2}$; the delta converter does not require high precision devices, but has a limited range depending on the frequency.

EXAMPLE

In order to study the advantages of OVERSAMPLING consider this pratical case: we need a 8-bit A/D system for voice signal, limited to 3 kHz



The bit rate is $8\,bit \times 8\,\frac{k\,sample}{s} = 64\,\frac{k\,bit}{s}$; the SNR_q is

$SNR_q = 6 \cdot N = 6 \cdot 8 = 48\ dB$; the range of variation in a NYQUIST converter is $\frac{E_q}{2} < V < \frac{S}{2}$; the ratio is $\frac{S}{E_q}$;

$$E_q = \frac{S}{2^{N+1}} \rightarrow \frac{S}{E_q} = 2^{N+1} = 2^9 = 512.$$

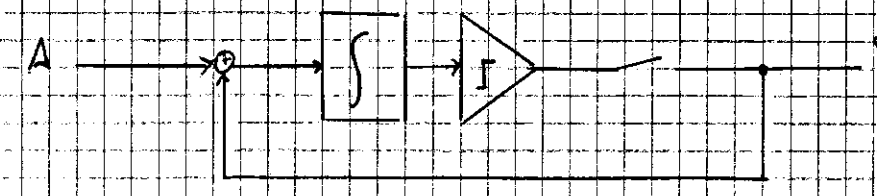
If we want the same variation in a Δ converter (also called OVERSAMPLING converter), we have

$$\frac{\frac{\delta}{\omega T_{CK}}}{\frac{\delta}{2}} = \frac{2}{\omega T_{CK}} = 2^9 \rightarrow T_{CK} = \frac{1}{\omega \cdot 2^8}$$

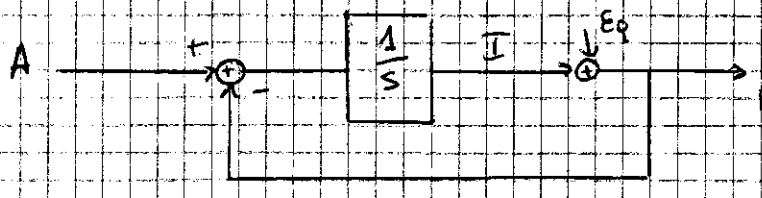
$$\omega = 2\pi \cdot f_s = 2\pi \cdot 3k \rightarrow f_{CK} = 2\pi \cdot 3k \cdot 2^8 = 4.8\,\frac{M\,bit}{s}$$

So the bit rate is much higher in a Δ converter

derivator. So this scheme starts from the STANDARD CHAIN and adds on integrator at the input (with the goal of decreasing amplitude when frequency goes up - in fact if $A = V \sin(\omega t)$, then $A_1 = \frac{V}{\omega} \cos(\omega t)$) and also a derivator at the output.
However this diagram can be improved:
the decoder is made of an integrator and a derivator that have an inverse function; so they can be substituted with a wire. At the input there is the sum between two values integrated; it is the same if we integrate the difference of the two signals. So:



Let's study it in the frequency domain, taking into account the quantization noise:

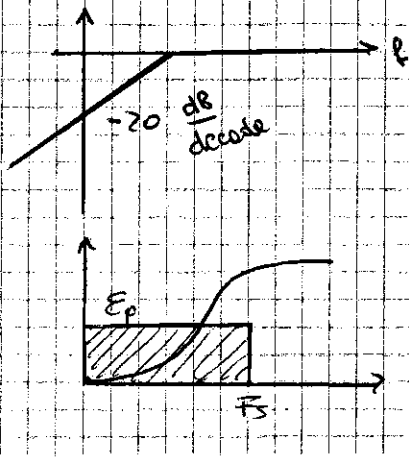


$A = 0 \rightarrow$

$$D = I + \varepsilon_p =$$

$$= -\frac{D}{s} + \varepsilon_p$$

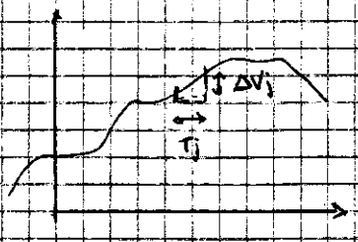
$$D\left(1 + \frac{1}{s}\right) = \varepsilon_p \rightarrow \frac{D}{\varepsilon_p} = \frac{s}{s+1}$$



The circuit behaves as a high-pass filter for the noise, that means that noise is present only at higher frequencies, so it is easier to be reduced with the reconstruction filter.

12 poles are needed to ensure $SNR_q = 72$ dB @ 100 KHz

Consider now a sine wave input at 50 KHz; we want to evaluate the SNR_j (jitter) due to the limited variation of the ADC; assume $T_j = 1{,}2$ ns



$$\Delta V_j = SR_{max} \cdot T_j$$

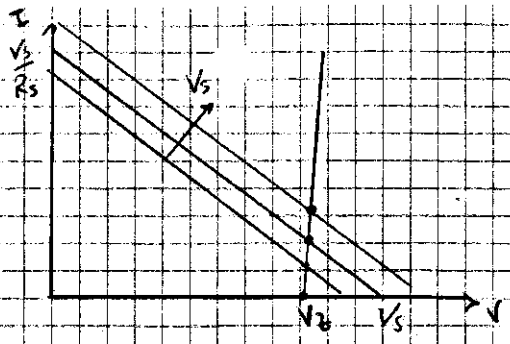
$$SR_{max} = \omega V = \frac{5}{2} \cdot 2\pi \cdot 50K$$

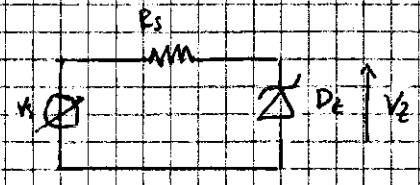
$$\Delta V_j = \frac{5}{2} \cdot 2\pi \cdot 50K \cdot 1{,}2 \cdot 10^{-9} = 5 \cdot 188 \cdot 10^{-6}$$

$$SNR_j = \frac{P_S}{P_{N_j}} = \frac{P_S}{\frac{\Delta V_j^2}{12}} = \frac{P_S}{\frac{5^2 \cdot 188^2 \cdot 10^{-12}}{12}} = \frac{\left(\frac{5}{2}\right)^2 \cdot \frac{1}{2}}{\frac{5^2 \cdot 188^2 \cdot 10^{-12}}{12}} =$$

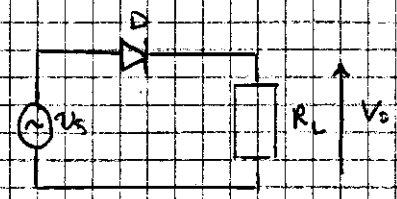
$$\simeq 76 \text{ dB}$$

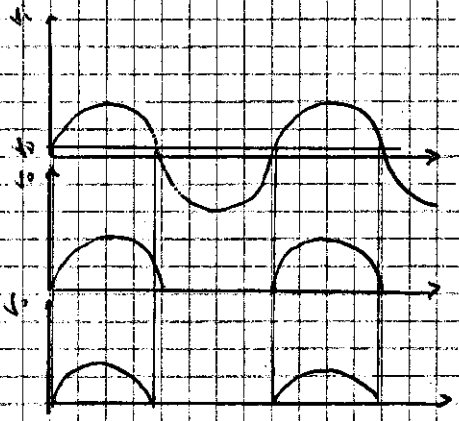
for current and voltage shown in the circuital symbol, breakdown voltage is positive; imagine to use it in a simple circuit with a voltage generator and a resistor; even if the voltage in the generator changes, the one of the zener diode remains always the same.



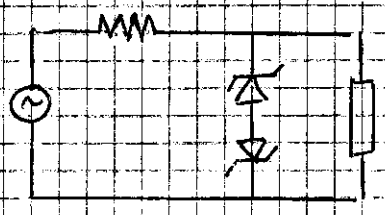


Diodes are used in rectifier circuits: the circuit on the right offers to the load only the positive half of the input sine voltage; the first output is shown using the ideal model of short for the diode, the second using the model with 0,7V and the resistance.
With 4 diodes it is possible to obtain a full wave rectifier.
We can also put a capacitor in parallel with the load to have less variations in output voltage.

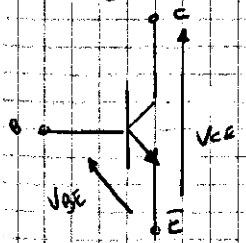




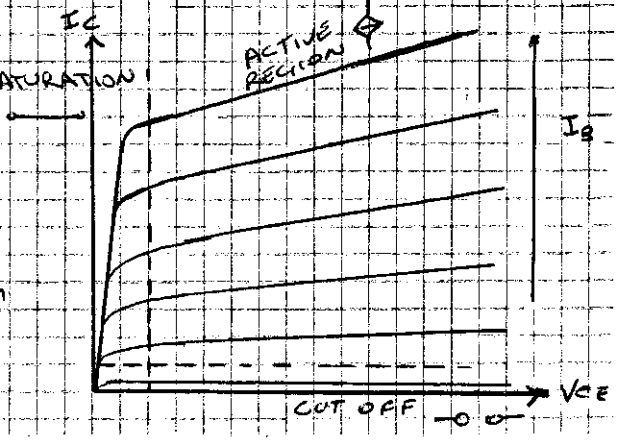
To limit voltage instead we can use a couple of zener diodes.



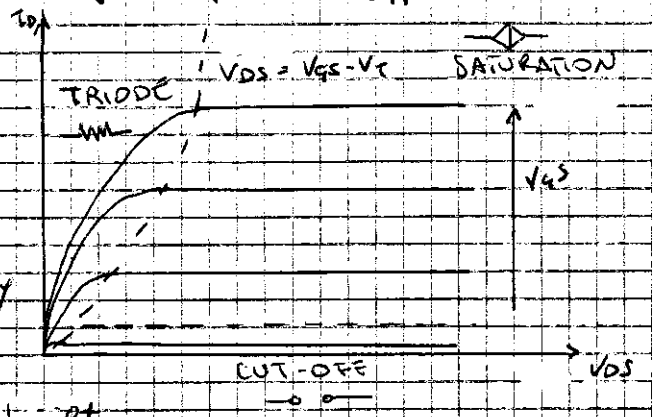
BJT



The BJT is used as an amplifier in the active region; instead is used as a switch in saturation (ON) and cut-off (OFF).



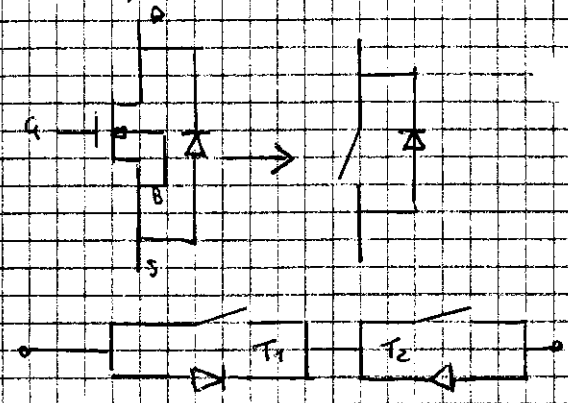
substain higher currents and voltages. However this kind of transistor can be built in single pieces, so it is used only in small numbers for power applications.



Triode and cut-off regions are used to implement switches.

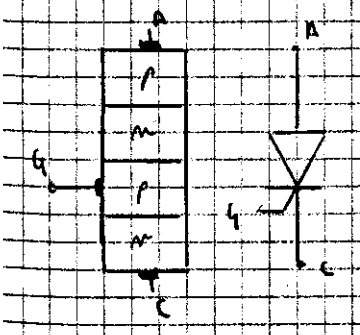
Mosfet use majority carriers, so they have a higher switching speed

In the way vertical MOSFET are built, there is a diode between bulk and drain; so the model, if we want to use the MOSFET as a switch is:

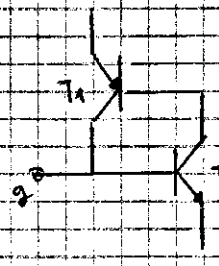


Current can flow from S to D, so this is a monodirectional switch; in order to have a complete switch we have to put two transistors back to back.

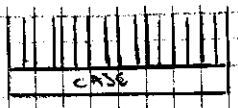
THYRISTOR



Gate is the control input of this device; actually, it is the name of connecting two BJT as in figure:



giving a small current to g, T_2 turns on and this I_{C_2} becomes base current for T_1, I_{b_1}; so both device conduct until current goes to 0



Usually we have a structure useful to dissipate heat better: it is called heatsink. Its thermal resistance θ_{rca} depends on air velocity (that's why we use fans in the pcs)

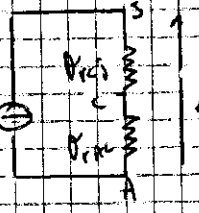
EXAMPLE

$P_d = 5\,W$

$T_a = 50°C$

$T_j = ?$

From the BST data sheet we have $\theta_{rjc} = 3{,}125\ \frac{°C}{W}$ while the heatsink has $\theta_{rCA} = 3\ °C/W$.



$$T_j = T_A + \Delta T =$$

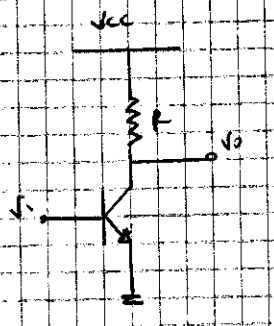
$$= T_A + (\theta_{rjc} + \theta_{rCA})\, P_d =$$

$$= 50\ °C + 6{,}125\ \frac{°C}{W} \cdot 5\,W \simeq$$

$$\simeq 80°C$$

Power Circuits

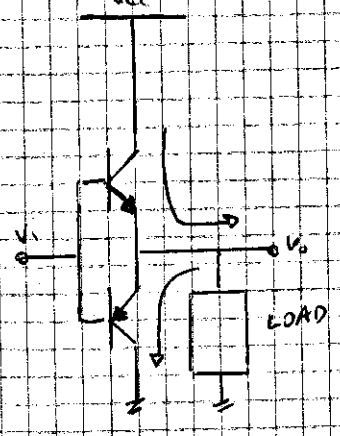
An amplifier produces "power amplification"; power amplifiers must handle large amounts of power (> 1 W)



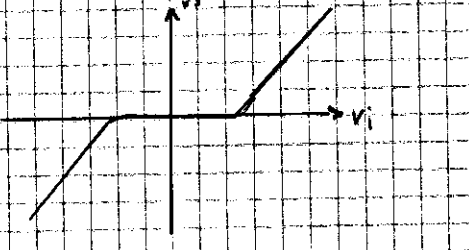
CLASS A

I_c circulates all the time

CLASS B



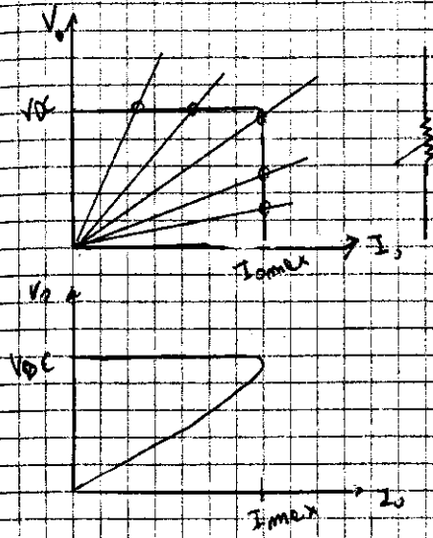
Each transistor conducts for half period; if $|V_i| < 0{,}7V$ then the circuit is off, no current to the load.



two following parameters:

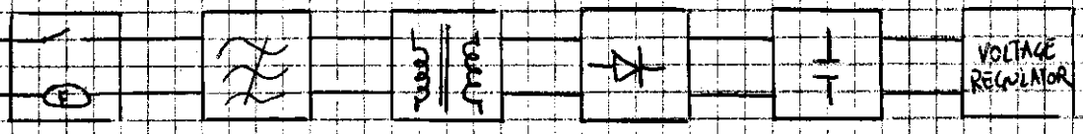
→ load regulation $R_o = \frac{\Delta V_o}{\Delta I_o}$ (output resistance)

→ line regulation $S_v = \frac{\Delta V_o}{\Delta V_i}$



Another parameter is the behaviour in load changings: if it has a LIMITER CIRCUIT, increasing R too much the voltage will decrease but the current will be the same. Or, it can have a FOLDBACK that switches off the circuit if I_o is too high.

A circuit is also characterised by Input (in rush current, that is the current needed when it is switched on; and also, the hold-up time, that is how long output voltage is sustained after a power failure. During this time the system is put in a safe condition, saving data in memories and moving mechanical parts to a safe position etc.



The first step is a switch with a FUSE: the fuse is a wire that blows when current is too high - it is a thermal behaviour on the energy of the input: a FAST FUSE doesn't accepts spikes even for a very short time, while a SLOW FUSE doesn't accept spikes only if they are present for a longer time.

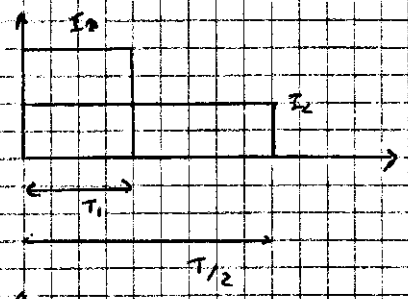
FUSE → FAST

→ SLOW

We also have to evaluate the peak current of the diode in order to choose the right components; we can evaluate it remembering that the charge that goes into C must be equal to the one that goes out.

It is difficult to evaluate it with the strange shape of I_D, so we can approximate, in 3 different ways:

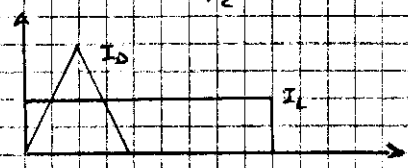
1. RECTANGULAR



$$\tilde{I}_D T_1 = I_L \frac{T}{2}$$

$$\tilde{I}_D = \frac{I_L \cdot T/2}{T_1}$$

2. TRIANGULAR



$$\frac{\tilde{I}_D \cdot T_1}{2} = I_L \cdot T/2$$

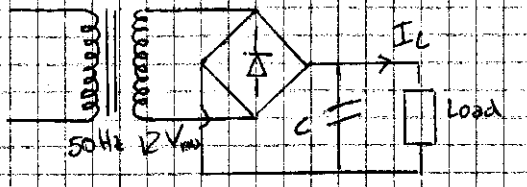
$$\tilde{I}_D = \frac{I_L \cdot T}{T_1}$$

3. PARABOLIC

$$\tilde{I}_D = 1{,}5 \frac{I_L \cdot T/2}{T_1}$$

In order to be safe I choose a diode that can support the higher current (worst condition) → triangular approximation.

EXAMPLE

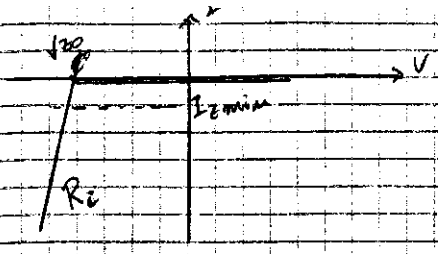


We have 12 V_{rms} at the output of the transformer; we want to evaluate the output DC V knowing that $I_L = 200$ mA, $C = 4000$ μF and $V_D = 1{,}5$ V

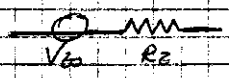
$$V_R = \frac{I_L \cdot T/2}{C} = 0{,}5\ V$$

$$V_{DC} = V_P - 2V_D - V_R/2 = \sqrt{2}\, V_{in} - 2V_D - \frac{V_R}{2} =$$

$$= (12 \cdot \sqrt{2} - 3 - 0{,}5)\ V = 13{,}7\ V$$

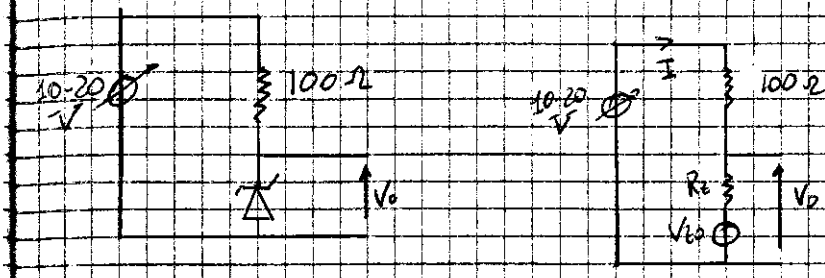


Voltage reference V_{z0} in series with a resistance R_z; remember that current must be higher than $I_{z\,min}$ to have a proper working.



EXAMPLE

$V_{z0} = 5V \quad R_z = 10\,\Omega \quad I_{z\,min} = 5\,mA$



$$I_{min} = \frac{10\,V - 5\,V}{110\,\Omega} = 45\,mA$$

$$I_{max} = \frac{20\,V - 5\,V}{110\,\Omega} = 130\,mA$$

$V_{o\,DC\,min} = I_{min} \cdot R_z + V_{z0} = 5{,}45\,V$

$V_{o\,DC\,max} = I_{max} \cdot R_z + V_{z0} = 6{,}3\,V$

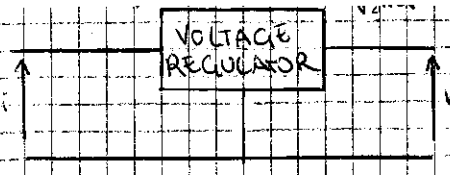
$\Delta V_o = 0{,}85\,V$

$$P_{max_{Dz}} = V_{z\,max} \cdot I_{max} = 6{,}3\,V \cdot 130\,mA = 0{,}819\,W$$

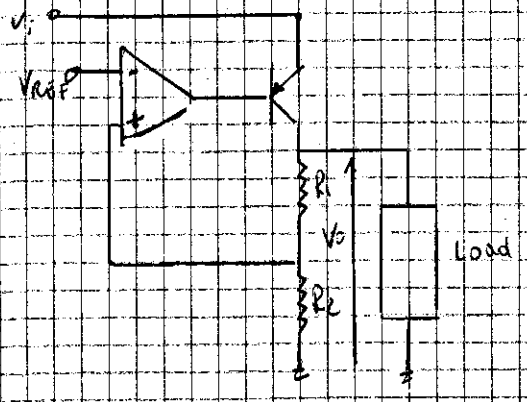
$$S_i = \frac{\Delta V_o}{\Delta V_i} = \frac{0{,}85}{10} = 0{,}085$$

If $\Delta V_i = 2V \rightarrow \Delta V_o = \Delta V_i \cdot S_i = 2 \cdot 0{,}085\,V = 0{,}17\,V$

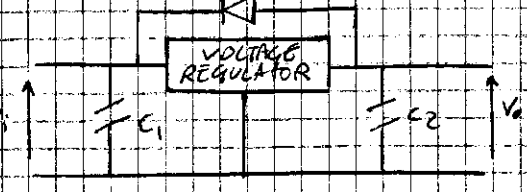
Series regulator instead needs a controlled variable resistor, and a feedback loop with a voltage reference and an Op Amp. These regulators have a voltage drop so $V_o < V_i$. They are however better than shunt regulators, so are widely used. One of the fields in which SHUNT regulator is needed is the credit card circuit (credit card is the load) because in series regulator the current is the same on card and device, and reading this current it is possible to obtain the secret code of the card; instead in shunt regulators the current divider load / diode doesn't allow this reading



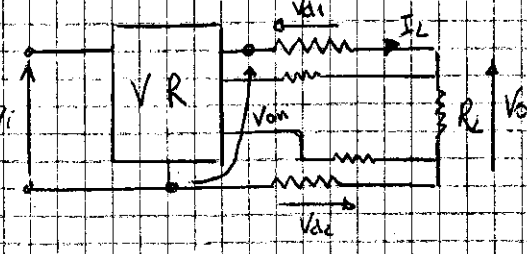
exist and they are called



this way the equivalent resistance and inductance are kept at the desired values.



and a diode to allow current flows and $V_o = 0$ V when $V_i = 0$ V, because otherwise $V_o > V_i$ could damage the voltage regulator.



some regulators have two more pins for voltage; no current flows in those wires so they can be also thinner; these wires are called remote sense wires.

A quiescent current flows out from the pin down in the picture (next page); if this current is very low (LM317) we can use this device to build a variable voltage regulator.

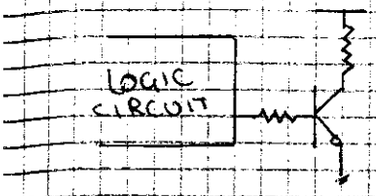
A voltage regulator has a DROPOUT in it; this reduces the available V_o; some regulators with low V_{drop} LOW DROP REGULATOR

Using a PNP transistor in a CE configuration, it is an amplifier (no more a voltage follower); pins of the Op Amp must be inverted to keep negative the loop.

Manufacturers specify both type and value of output capacitor; in

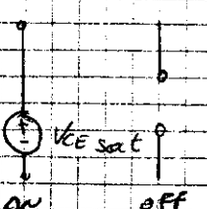
Good design rules for LDO regulators are the use of a capacitance C_1 to compensate wire inductance

Wires or lines in PCB have a resistance that cause voltage drops, so that V_{om} (coming out from the VR) is different from V_{or} in the load. To solve this problem



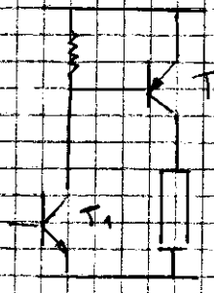
This circuit can work for the OFF state, but not for the ON state because the current is too low.

The equivalent circuit for the ON state is a voltage generator of $V_{CE\,sat}$, instead for the OFF state it is an open circuit; because of the leakage current I_{CBO} from Collector to Base it is necessary that the base is connected through a resistance to ground so that in the OFF state I_{CBO} can flow to ground and not through the BJT

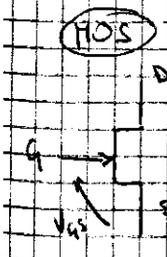


HIGH SIDE SWITCH

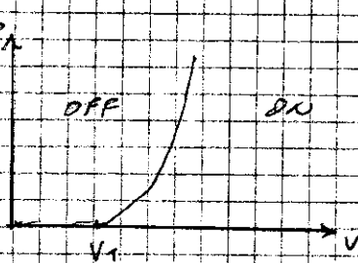
It is not possible to drive an npn BJT used as high side switch, because the ON state would need a voltage at the base higher than V_{cc}, so we can use a pnp as high side switch and it is possible to drive it with an npn, as shown in the picture. When T_1 is ON, base of T_S is at 0V so it is also on; when T_1 is OFF, base of T_S is near 12 V and so T_S is ON.



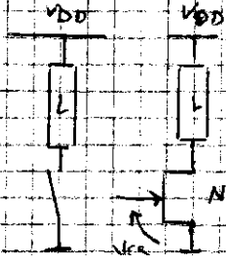
MOS



MOS transistor is OFF till $V_{GS} < V_T$; when $V_{GS} > V_T$ it is ON and we have $I_D = V_{DS} / R_d$



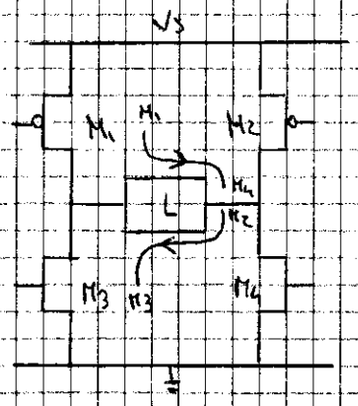
where R_d is the internal resistance that depends on temperature, V_{GS} and V_{DS}. I_D is always limited by R_d, because in any case $I_D \leq \frac{V_{DS}}{R_d}$



low side switch

An N-MOS is used as low side switch; if we used a P-MOS we wouldn't have known source voltage

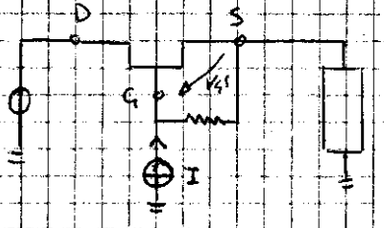
H-BRIDGE



This bridge allows to change the verse of the current in the load; if M_1 and M_4 are on, current flows in one direction, if M_2 and M_3 are on, it flows in the other direction; pay attention in order to don't have both the transistor of one side on at the same time: there would be a short between Vs and GND and everything would be destroyed.

FLOATING SWITCH

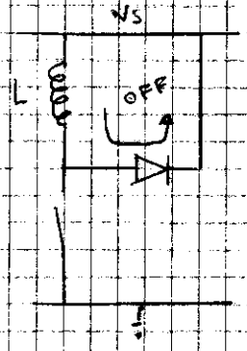
An example of floating switch is the switch of a S&H unit; this circuit allows to drive a floating switch



$I = 0 \rightarrow V_{GS} = 0 \rightarrow$ OFF

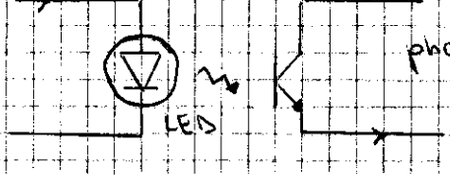
$I = I_C \rightarrow V_{GS} = R I_C \rightarrow$ ON

OVERVOLTAGE PROTECTION

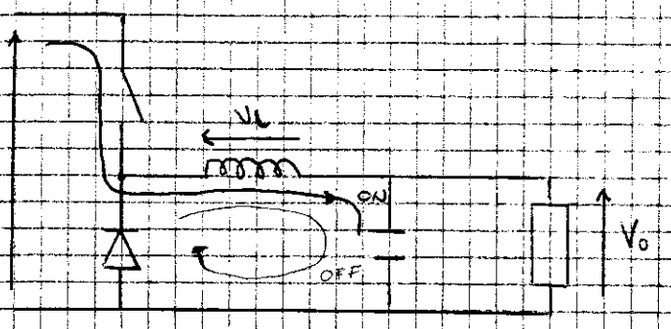


Consider an inductive load; when switch goes OFF the current in the inductance cannot go to 0 immediately. A path for current to flow out is needed. So a diode is used in this way.

Finally, we didn't discuss yet the need of a galvanic isolation between power section and control section. It can be provided with a transformer or an optocoupler



BUCK



ON: $V_L = V_i - V_o$

$$\Delta I_L = \frac{V_L \cdot T_{on}}{L} = \frac{(V_i - V_o)\, T_{on}}{L}$$

OFF: $V_L = -V_o$

$$\Delta I_L = -\frac{V_o\, T_{off}}{L}$$

$$\Delta I_{on} = -\Delta I_{off}$$

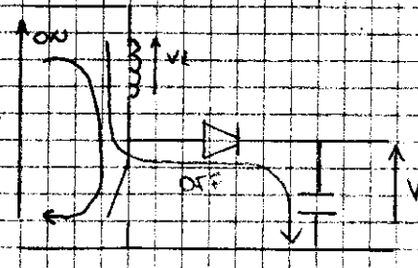
$$V_i T_{on} - V_o T_{on} = V_o T_{off}$$

$$V_i T_{on} = V_o T_s$$

$$\frac{V_o}{V_i} = \frac{T_{on}}{T_s} = D \qquad \frac{V_o}{V_i} = D$$

We can even replace the diode with a switch driven in the reverse way.

BOOST



ON: $V_L = V_i$

$$\Delta I_L = \frac{V_i\, T_{on}}{L}$$

OFF: $V_L = V_i - V_o$

$$\Delta I_L = \frac{(V_i - V_o)\, T_{off}}{L}$$

$$\Delta I_{on} = -\Delta I_{off}$$

$$V_i T_{on} = -V_i T_{off} + V_o T_{off}$$

$$V_i T_s = V_o T_{off}$$

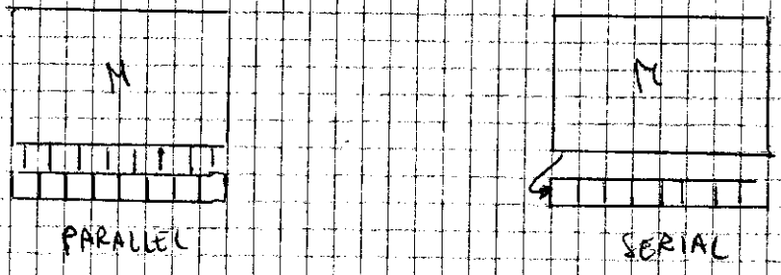
$$\frac{V_o}{V_i} = \frac{T_s}{T_{off}} = \frac{1}{1-D} \qquad \frac{V_o}{V_i} = \frac{1}{1-D}$$

Semiconductor Memories

In all fields where a programmable processor is needed (embedded systems for example), a memory from which processor takes the instructions is needed.
Let's first introduce taxonomy by consider these parameters:

- size (total number of bits)
- organization (size of a word)
- external access (serial or parallel)
- Read/write capability
- Volatility

From the point of view of external access, a parallel access is faster, but requires more wires than a serial one.

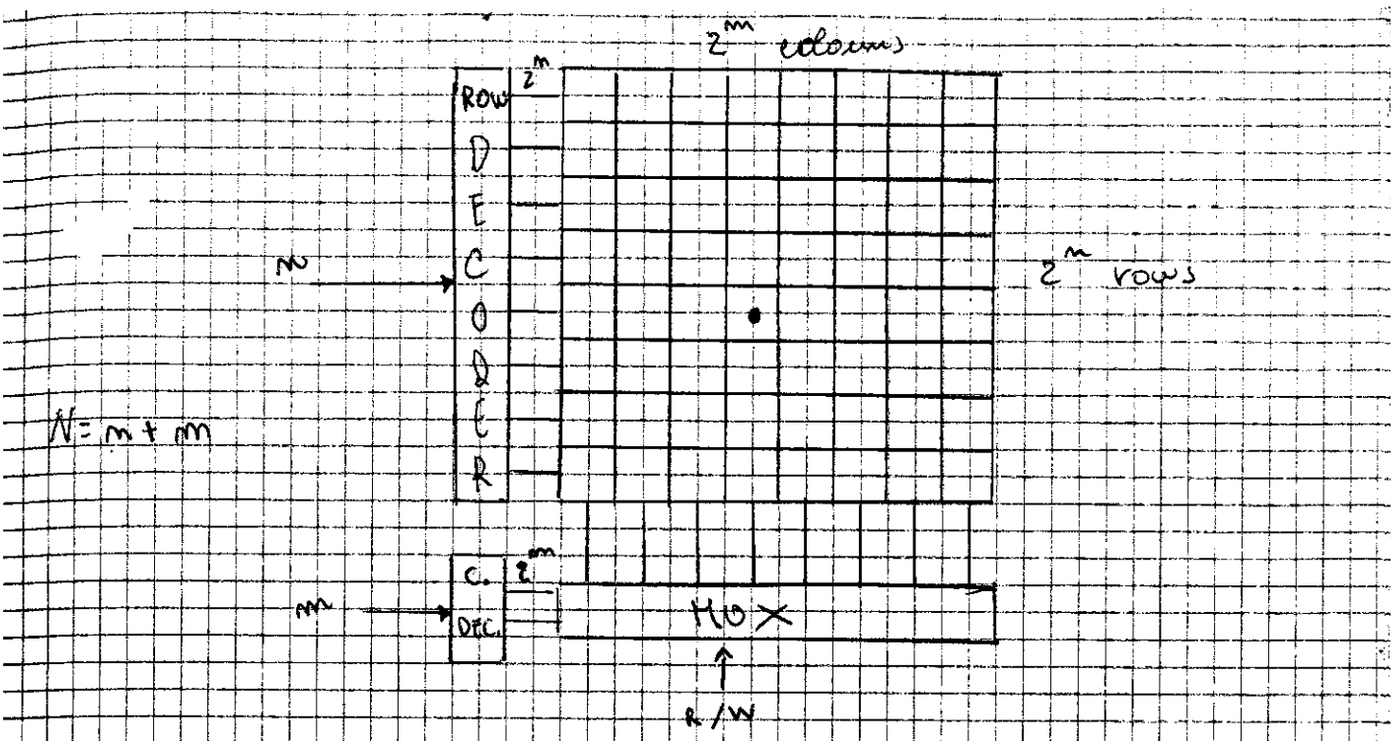


From the point of view of read/write capabilities instead we can distinguish:

- read only memories ROM
- read and write at the same speed READ WRITE M.
- read faster than writing PROGRAMMABLE M.

And, from the point of view of volatility:

- datas stored forever
- datas stored until V_DD
- datas stored for a limited time.



ROM

Read Only Memory - Nonvolatile ; it is programmed before being inserted in the embedded system ; they are used to store programs for general purpose processors, to store constant data needed by system or implement COMBINATIONAL CIRCUIT (just by putting in the right row - referred to a certain input - the output values).

Costs overview: the cost to produce one ROM (or any other device) can be divided in RECURRENT and NON-RECURRENT the first one is the real cost of fabbrication, and it is constant; the second one instead depends on the number of pieces producted: non-recurrent costs are the cost of the mask and the cost of the farm i.e.; so the total cost of one piece is

$$\text{TOTAL COST} = \text{RECURRENT COST} + \frac{\text{NON-RECURRENT COSTS}}{\text{N° PIECES}}$$

The produces must keep this cost as low as possible, to earn more money. In this context non recurrent costs have to be kept as low as possible.

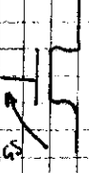
Instead of a fuse, an ANTIFUSE can be used; it performs the opposite operation of a fuse: it is a dielectric that can become a short circuit.

NRC are much lower in this case; OTP-ROM must be small: they cannot be tested at production time, because to be tested they must be programmed, so it is very important that they work all (nearly all) properly, that means a very high yield; since yield decreases with dimension, area must be small.

- EPROM

Erasable Programmable ROM.

First of all an introduction to some different devices.

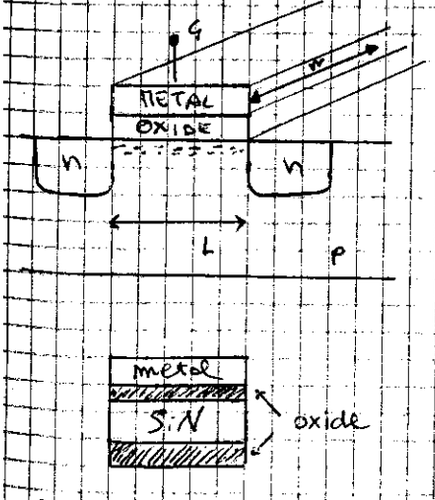


For our purpose an nMOS can be seen just as a switch:

$V_{GS} > V_T$ → ON → short

$V_{GS} < V_T$ → OFF → open

The idea to SET the state of the transistor (to store the bit) is to modify V_T, a very low V_T makes the MOS always on, while a very high V_T makes it always off.



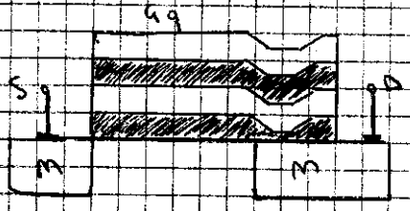
V_T can be controlled by putting positive or negative charges in the oxide: if I put negative charges, the channel is created with more difficulties, so V_T increases; if I put positive charges, the channel is easily created so V_T decreases. This method is used for SONOS transistors (Silicon - Oxide - Nitride - Oxide - Silicon);

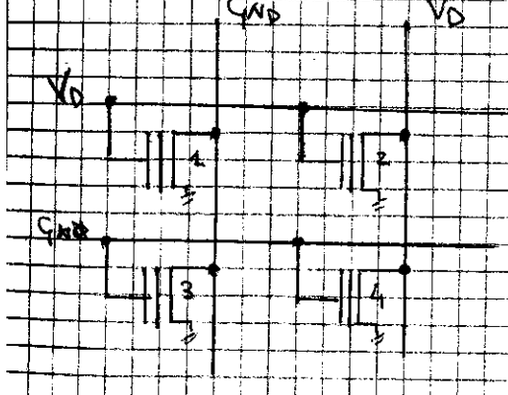
however this device is much bigger than a normal MOS, so they are used in environment with ionizing radiation only (military, space, nuclear plants)

rays; because $E = hf$, the this energy (high because f is high) is transmitted to the electron that can return to channel. So an EPROM package has a quartz window through which UV light can pass.

2nd - FOWLER NORDHEIM EFFECT

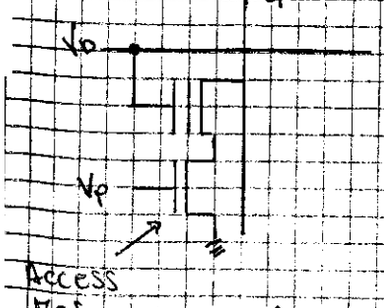
Fowler Nordheim effect requires a very thin oxide; this allow electrons to pass through the barrier by tunneling; the oxide in the thin zone is 8-10 nm wide; one problem is that this process is not self regulating, the cell is programmed giving a high voltage at the gate and putting source and drain to GND; to erase it instead I have to put gate to ground and give a high voltage at source.





But, imagine I want to programme transistor 1; I give high voltage at gate and put source to ground; in order not to programme also MOS 2, I have to give VD also at 2nd bit line; then, not to erase MOS 3, I'll put WL2 to ground, but so MOS 4 is erased.

Every time I programme one cell, I erase another one!!!



The solution is to put another transistor that controls the drain of our MOS; only if the programming line is on (V_P) then the MOS is programmed or erased; this cause a doubled space, or, with the same area, half memory, but we can have electrically erasing instead of lighting one.

- EEPROM (E^2 PROM)

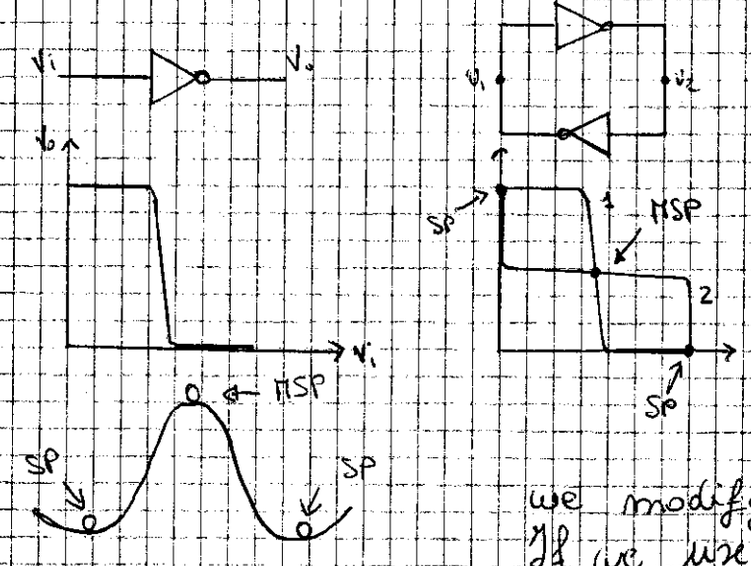
They use the described FOWLER NORDHEIM EFFECT: their advantage

we have $t = RC \ln \frac{2.3}{3} = 0,09 \cdot R_C$, i.e. it is 7 times faster.
Moreover, because of noise, a differential signal is always used, taking a line from V_{DD} and comparing it with the output using a sense amplifier.

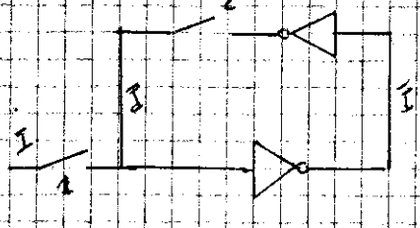
RAM

Rams are volatile memories that can be written and read at the same speed. Information can be stored in a capacitor or in an inverter. In the first case we have dynamic RAM, while in the second we have Static RAM

• SRAM - Static RAM

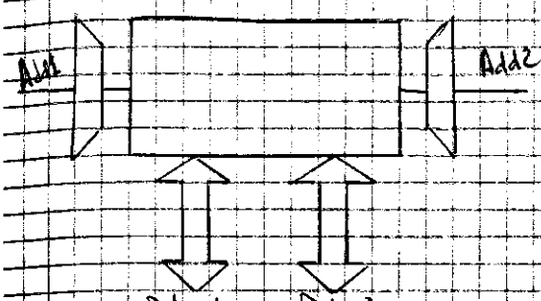


A couple of inverter can store the data (either 1 or 0) in the stable points SP, while MSP is a metastable point, like in a mechanical equivalent; the point is how can we modify the state.
If we use a NAND instead of an inverter we have another input that we can use to set the value; if I put **1** as input the NAND behaves as an inverter, instead if I put **0** I reset the cell.

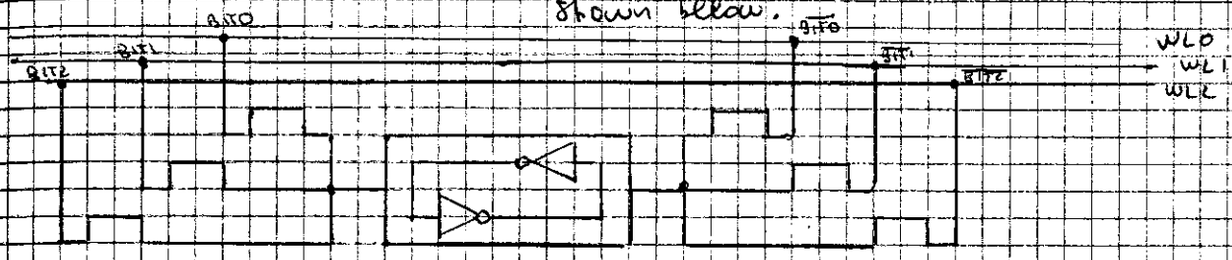


Another method is to use a couple of inverter with two switches that works in reverse (when one is open the other is closed); we close "1" to set the state I, then we close "2" to keep the fixed value. So an enable line is needed to close

• MULTIPORT RAMS

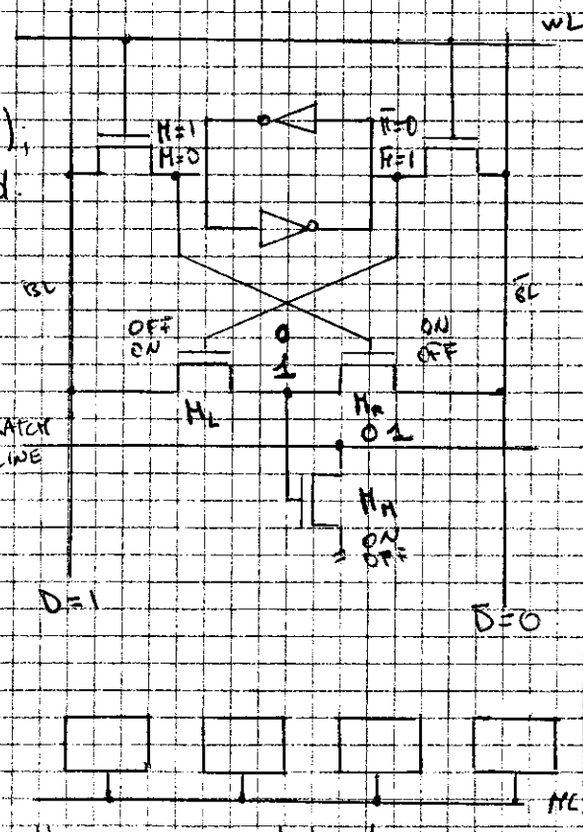


In some applications, and usually with multi-core processors, it is useful to have multiple access to memory; of course there are some restrictions: for example it is forbidden to write the same cell at the same time from two different processors; the implementation is shown below.

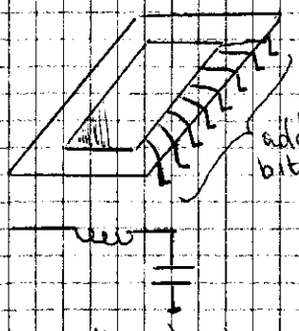


• ASSOCIATIVE RAMS

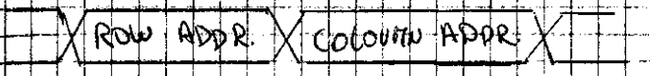
An associative RAM allows me to perform three operations: read, write and MATCH. Match operation finds the address of a word that I give to the memory (if it is already in memory), another line, called "match line" is needed. Let's see how it works for each bit: imagine first we have H=0 and we are looking for bit D=1; we give D to the bit line, and put the word line at low state. M_L (left) is on because H=1, while M_R is off. So gate of M_H (match) is at high value, so there M_H is on and match line is at 0. When instead H=1 we have M_R conducting so match line goes at 1. When I'm looking for a word all the cells are attached at the same match line, and only if all of them give 1 as result, the match is 1. Then an encoder is needed, that has the match lines as input and gives the address of the right line as output.



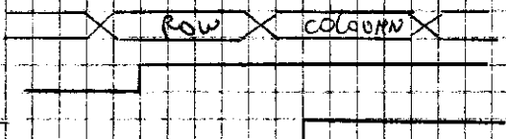
- Multiplexed addressing RAS and CAS



Consider the RAM in its package; it has pins to achieve the address to W/R; increasing the number of pins (bigger memory) the cost increases and also the parasitic effects of connections (connection wires have parasitic capacitances and inductances) increases. So it is better to have few pins as possible. In order to do this we can send row address and column address to the same pin.



The memory gets the row address and select the correct row; while all datas of the row are read and taken to the multiplexer, column address is get and used to select the right word among those of the selected row. This needs two other controls (pins):



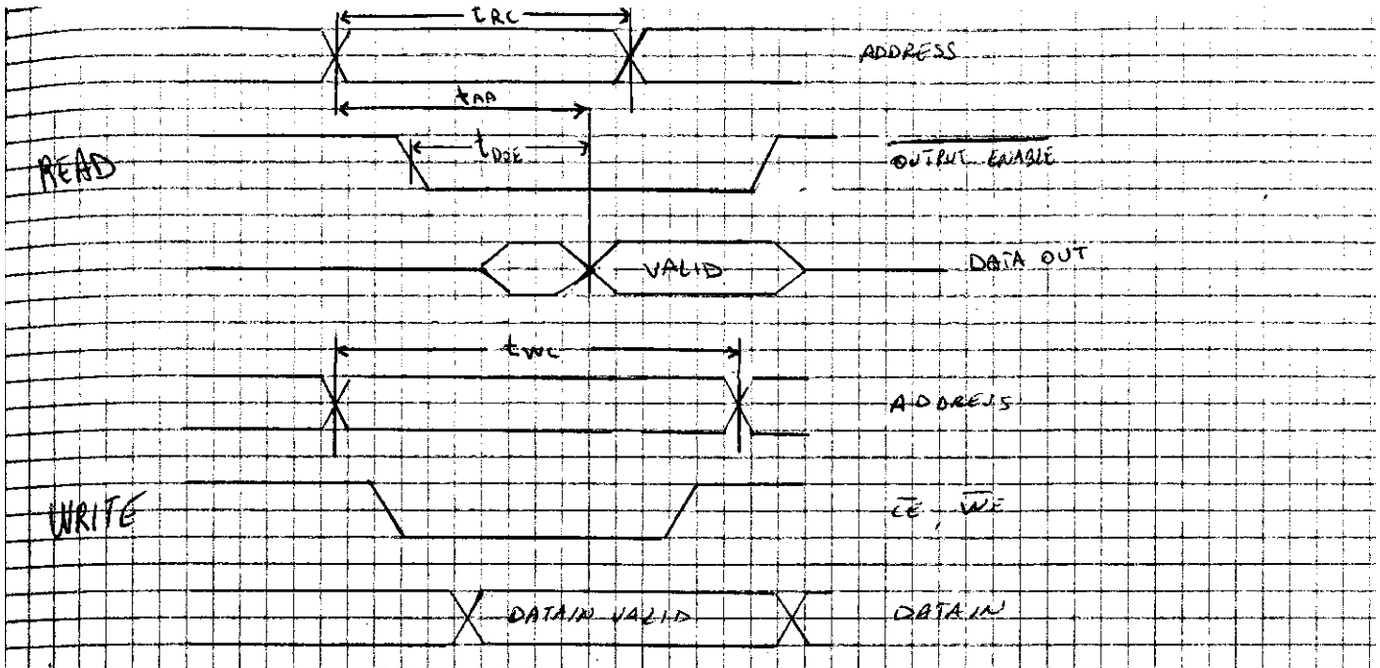
RAS ROW ADDRESS SELECT
CAS COLOUMN ADDRESS SELECT

If row addres dimension is the same of column address dimension then the number of pins is halved with this method.

There are some RAMs with variations: PSRAM (pseudo-static) have built-in memory refresh controller; NVRAM (non volatile) holds data after power removed by using a battery; SRAM with EEPROM or FLASH move the content to the non-volatile part before turning off.

Ferroelectric memories

Using ferroelectric materials, the logic state can be saved as direction of dipoles oriented due to polarization (with an electric field); so datas remain stored even without voltage supply; these memories cannot be as small as DRAM (or less memory with same dimensions - up to 256 MB), but they are more resistant to radiations than EPROM and FLASH.

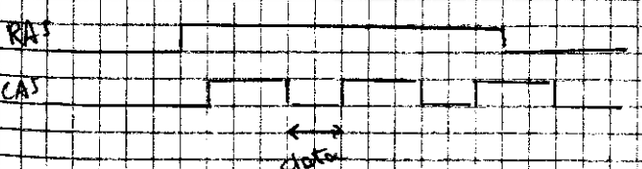


Because of these delays, access to memory is slow in comparison with processor work frequency. To reduce this gap two methods can be adopted:

1 - SPEED UP ADVANCED RAMS
2 - CACHE

① SPEED UP ADVANCED RAMS

Addresses that I usually needs are sequentially (both for datas and instructions); calling "PAGE" all the words of the same row, often selecting a row, and the needed column, all next columns of same page are read. Datas are given at output after a CAS signal and before next CAS signal; intro- ducing a latch before output buffer; time reduction is of 20, 30 %.



Instead of giving different CAS signals it is possible to synchronise the DRAM with the edge of the clock connected both to processor and memory. In this case a counter is needed; the elimination of time to detect RAS/CAS and RD/WR signals makes the RAM faster.

Direct mapping is simple and inexpensive but, on the other hand fixed locations for given blocks cause that if a program accesses 2 blocks that map to the same line repeatedly, cache misses are very high.
Fully associative instead has an expensive searching because all tags have to be compared with the one searched.

Replacement Algorithm: the question is how to choose the block to replace in a cache; this problem doesn't exist for direct mapped caches because each block can go only in one fixed position.
For fully associative and set associative caches we have four algorithms possible:

- Random → the cheapest one
- Least Recently Used (LRU) → a cycles counter is needed (it restarts counting when the block is accessed)
- First In First Out (FIFO) → FIFO structure needed
- Least frequently used → a counter is needed (how many accesses to that block) doubt: access in memory, then if another access in memory is needed, the most recent data is cancelled!

Write policy: obviously a cache block cannot be overwrite before memory updating; there are two different policies for memory updating:

- Write through → writes go to cache and memory; this cause a high traffic and slower writes, but memory is always consistent, so access from other CPU's or I/O devices to memory gets always the right datas.
- write back → updates are made in cache only initially, then memory is updated only if bits in cache for a block have changed. This cause non consistent memory, so I/O must access memory through cache; but, this reduces a lot traffic and jam-writes (successive updating of the same data).

Cache must have the right dimension; a too small cache cause a high miss rate, while a too big cache cause higher