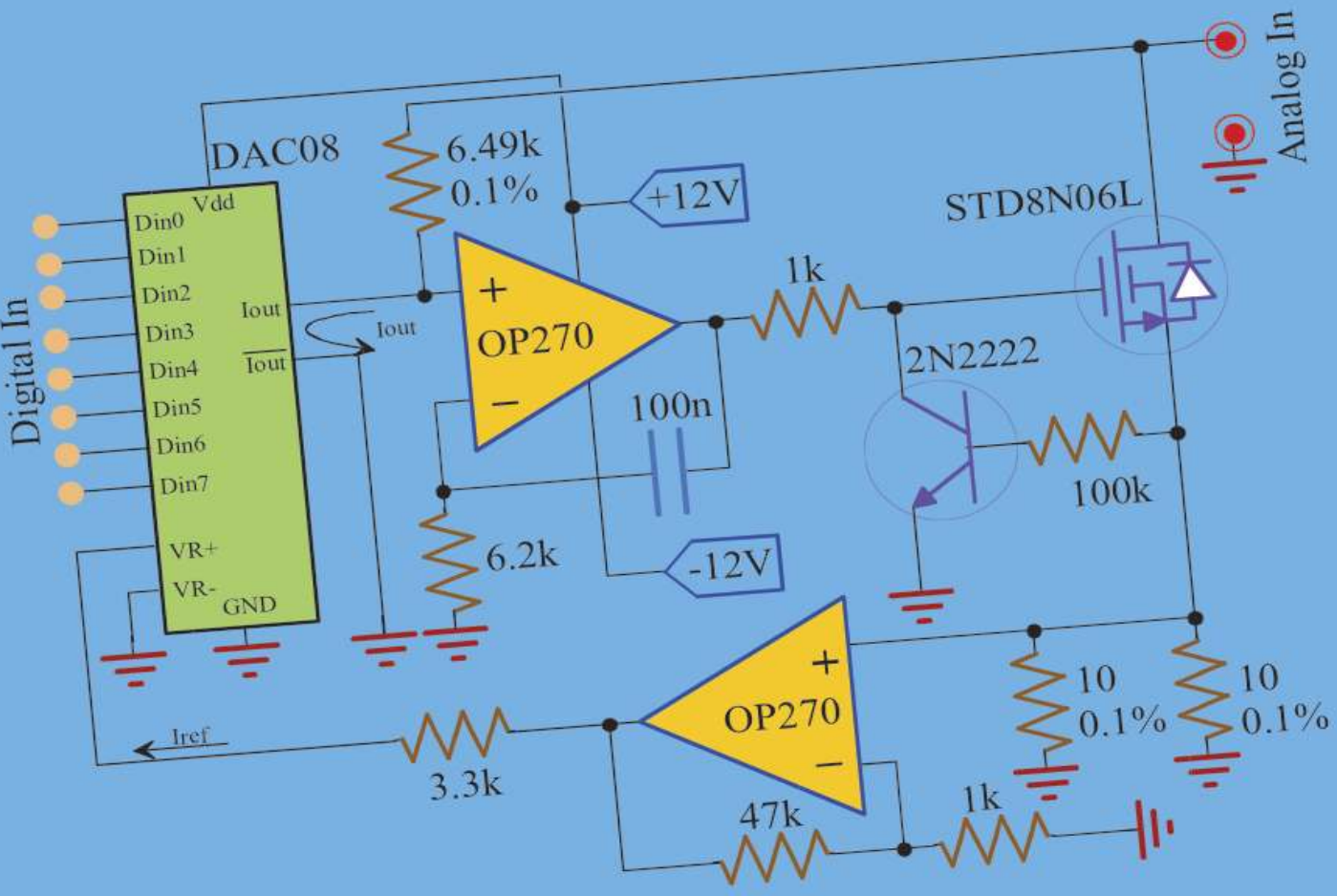


Electronic Systems

Operational amplifiers and circuits
INA, OTA, CFA, Norton, ISO advanced OpAmps
negative feedback, stability and frequency compensation
Sample&Hold circuits, DAC and ADC converters
oversampling and Sigma-Delta modulators, MicroControllers

Franco Zappa

SOCIETÀ EDITRICE
ESCULAPIO

Franco Zappa

Electronic Systems

Operational Amplifiers and circuits,
INA, OTA, CFA, Norton, ISO advanced OpAmps
negative feedback, stability and frequency compensation,
Sample&Hold circuits, DAC and ADC converters,
oversampling and Sigma-Delta modulators,
MicroControllers

SOCIETÀ EDITRICE
ESCULAPIO

Publishing Manager: Alessandro Parenti *Editorial Staff:* Gabriella Gatti, Giancarla Panigali, Carlotta Lenzi

40131 Bologna - Via U. Terracini 30 - Tel. 051-63.40.113 - Fax 051-63.41.136
www.editrice-esculapio.it

Edizione digitale: settembre 2014

ISBN: 9788874885510

Edizione digitale realizzata da **Simplicissimus Book Farm srl**

Summary

2 Circuits with OpAmps

7 DAC converters

8 ADC converters

Index

A

B

C

D

derivator

E

F

G

gain

H

I

M

multiplexer

N

O

P

S

T

Total Harmonic Distortion (THD);

U

V

virtual ground;

W

Z

zero crossing detector;

Acknowledgments

This book is the result of years of teaching in the “Electronic Systems” and “Electronics” courses at the Master in Electronic Engineering at Politecnico di Milano. Many grateful thanks to all PhD students who contributed to let this book become real.

This book deepens and complements concepts, components, and circuital stages introduced in the book “Electronics”, which provided the basics to analog and digital electronic component and electronic circuits design. Aim of this “Electronic Systems” book is to provide further advanced topics, integrated circuits and circuital configurations to students and engineers working in the many fields of design and development of electronic products and systems.

This is the first release in English of the book. Many grammar errors are still to be corrected and the quality of the English text will be further improved, for sure. I wish to apologize for such an inconvenience. However I preferred to deploy a text book as soon as possible to help students in the Master in Electronic Engineering.

My best appreciations to some of my PhD students, namely Simone Bellisai, Andrea Bahgat Shehata and Bojan Markovich, for the reviews, the translations, the lunches together and the nice atmosphere in the labs. My best gratitude to Hazar Yuksel, for a comprehensive revision of most chapters and an in-depth improvement of its fluency. Eventually my devoted gratitude to enthusiastic and charming Federica Villa, whose support did, does and will keep me teaching while smiling and the other way round, too.

Milano, October 2012.

Franco

To Rita,
Valentina,
Marco,
and Gabriele

1

Operational Amplifiers

“Everybody screamed when I kissed the teacher
and they must have thought they dreamed when I kissed the teacher
All my friends at school
they had never seen the teacher blush, he looked like a fool
Nearly petrified ‘cos he was taken by surprise
When I kissed the teacher
Couldn’t quite believe his eyes, when I kissed the teacher
My whole class went wild
As I held my breath, the world stood still, but then he just smiled
I was in the seventh heaven when I kissed the teacher.

One of these days
Gonna tell him I dream of him every night
One of these days
Gonna show him I care, gonna teach him a lesson alright

I was in a trance when I kissed the teacher
Suddenly I took the chance when I kissed the teacher
Leaning over me, he was trying to explain the laws of geometry
And I couldn’t help it, I just had to kiss the teacher”

“When I kissed the teacher”, ABBA

1.1. Amplifier Stages

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. Consider, for instance, Fig. 1.1, which schematically shows the connection between a signal generator, represented by its Thevenin equivalent, and a voltage amplifier. Let R_S be the series resistance of the source generator (Source), and R_{in} the input resistance of the amplifier. The actual signal at the input of the amplifier stage is equal to $V_{in}=V_S \cdot R_{in}/(R_S+R_{in})$ with an evident signal loss if R_{in} is smaller than R_s.

Of course, the amplifier operates only on V_{in}, which is a portion of the signal V_S; the signal V_{in} best approaches the signal V_S supplied by the generator when R_{in} is large compared to R_S. It follows that a good voltage amplifier must have very large input resistance compared to the resistance of the signal source; ideally, an amplifier should behave like an open circuit.

A voltage amplifier has to not only properly read the signal to be amplified, but also fully supply the amplified signal to the output load. Again, if the output stage of the amplifier is represented with its Thevenin equivalent, the voltage that arises across the load R_L is given by:

$$V_L = V_O \cdot \frac{R_L}{R_L + R_O}$$

To maximize the voltage transfer from the input to the load, $R_O << R_L$ must hold. Therefore, the output of a voltage amplifier must have impedance much lower than the load to which the amplifier will be connected.

Let us now consider the case of an input current signal that must be amplified and then provided to the load as an output current signal (Fig. 1.2). I_S is the current source with parallel source resistance equal to R_S; the real input current can be computed with the current divider formula:

$$I_{in} = I_s \cdot \frac{R_s}{R_s + R_{in}}$$

It is evident from this equation that, in order to maximize I_{in}, $R_{in}<<R_S$ must hold. Therefore, the input impedance of a current amplifier must be small compared to the impedance of the signal source. In other words, the input stage of a current amplifier must be an excellent “current reader”, ideally a short circuit. However, with an analogous argument, you see at the output that $R_O>>R_L$ must hold so as to maximize the fraction of the current actually supplied to the load R_L, i.e. the amplified signal has to be delivered to the load with high impedance.

You can also have *trans-resistance* amplifiers that read a current and deliver to the load a voltage proportional to the current (Fig. 1.3); these amplifiers must have low input and output impedance. On the contrary, a *trans-conductance* amplifier (Fig. 1.4) reads an input voltage and delivers a current proportional to the input voltage. In this case, input and output impedance should both be high.

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.

1.2. Differential Amplifiers

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.

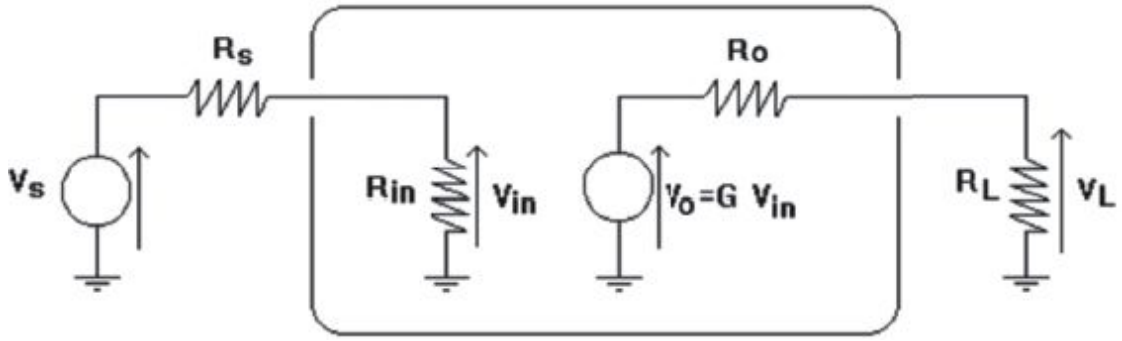


Fig. 1.1: Voltage amplifier.

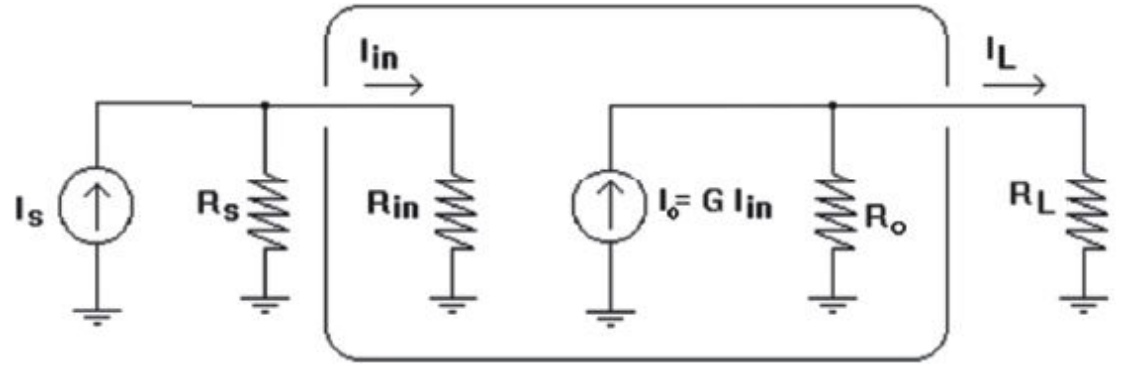


Fig. 1.2: Current amplifier.

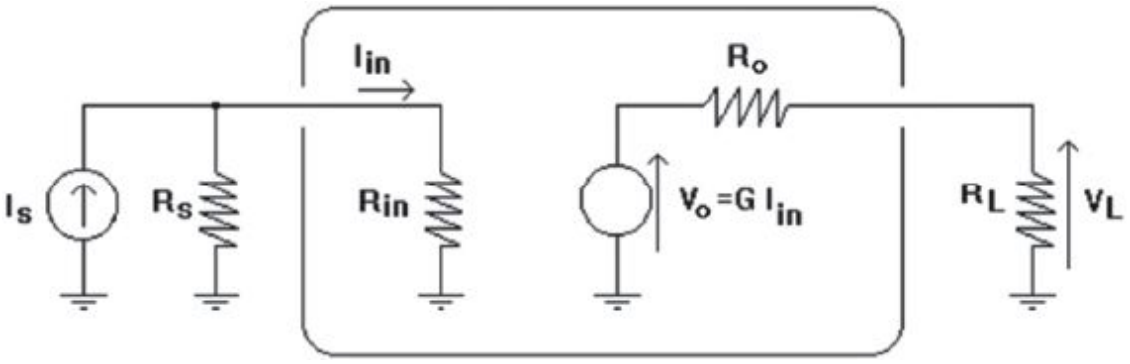


Fig. 1.3: Trans-resistance amplifier.

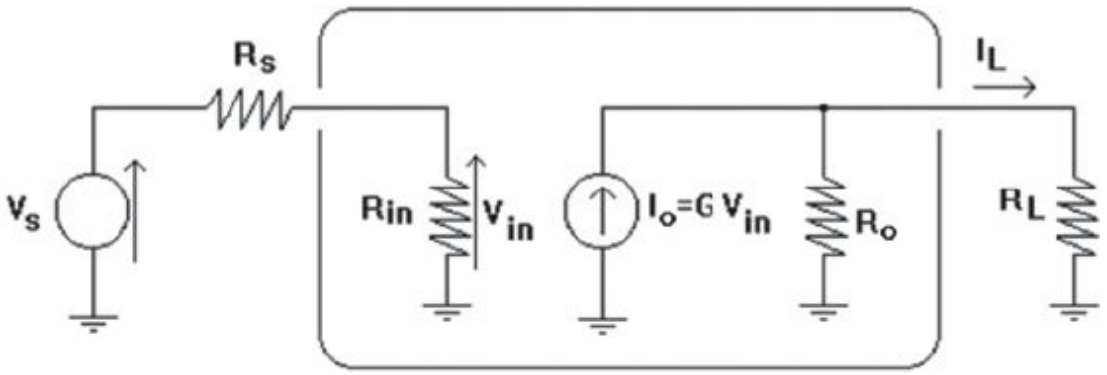


Fig. 1.4: Trans-conductance amplifier.

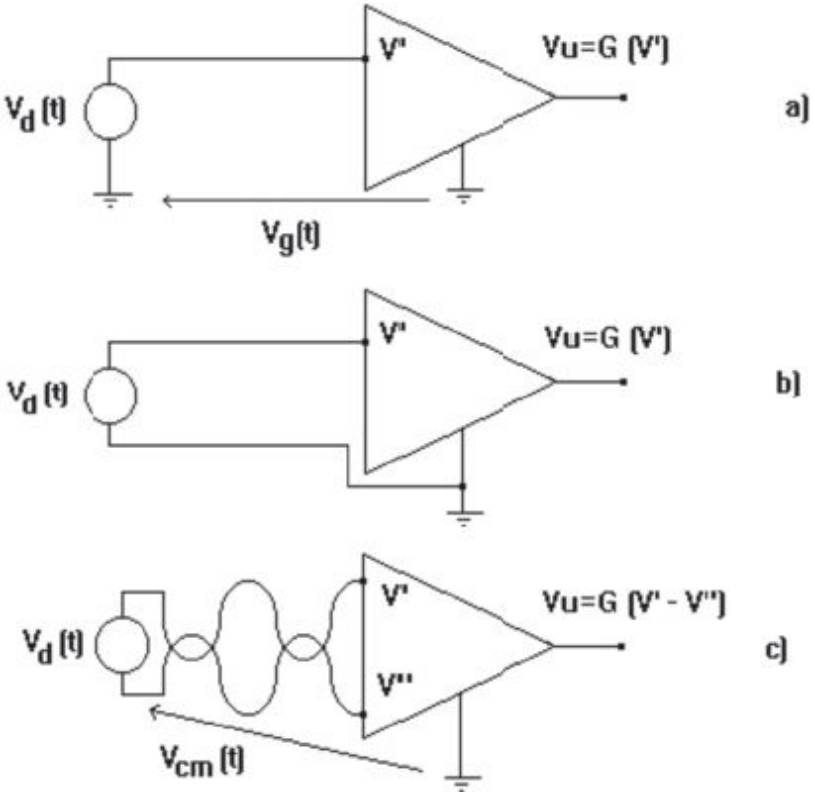


Fig. 1.5: Connection schemes of a sensor to an amplifier: a) single-ended through ground; b) single ended through two wires; b) differential sensing.

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=50μV 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/50μV)=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.).

1.3. FEEDBACK

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.

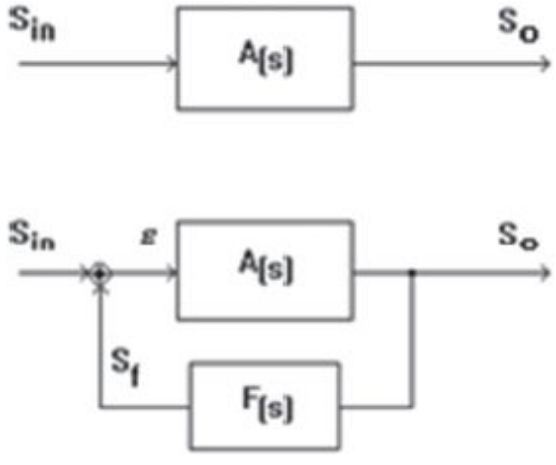


Fig. 1.6: From the simple amplifier stage (a) to the feedback amplifier (b).

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.

1.3.2 Properties of Ideal Feedback Circuits

Consider the system in Fig. 1.6a, where the amplifier has a transfer function A(s). In a circuit of this type, the changes in the components (due to temperature or aging) are directly reflected as changes in the performance of the amplifier, its gain, the position of the poles, etc. Furthermore, the non-linearity of the components of the amplifier results in the generation of amplified unwanted harmonics appearing in the output signal. To overcome these drawbacks, we can use the solution shown in Fig. 1.6b, obtained by adding the stage characterized by the transfer function F(s) that detects the value of the quantity S_o at the output and generates a signal S_f proportional to it. This signal, called the feedback signal, is added to the input quantity S_{in} by means of an adder node. In this way, the error signal ε is generated, which drives the block A(s). The resulting system is known as a feedback system. Its elements, i.e. the forward block A(s) and the branch of the reaction F(s), identify a loop called the **feedback loop**.

Obviously, the introduction of the feedback branch causes a change in the transfer function S_o/S_{in} of the original system. With reference to Fig. 1.6b, simple node balances lead to the following expressions:

$$\varepsilon = S_{in} + S_f = S_{in} + (S_o \cdot F(s)) \qquad S_o = \varepsilon \cdot A(s) = [S_{in} + (S_o \cdot F(s))] \cdot A(s)$$

which yield the transfer function of the feedback amplifier:

Eq. 1.1

$$G(s) = \frac{S_o}{S_{in}} = \frac{A(s)}{1 - A(s) \cdot F(s)} = \frac{A(s)}{1 - Gloop} = -\frac{1}{F(s)} \cdot \frac{1}{1 - 1/Gloop}$$

These expressions are valid when the ***loop gain*** $G_{loop}=A(s)\cdot F(s)$ is negative or at worst positive but less than one. Similarly, we obtain the expression of the signal that drives the forward amplifier:

Eq. 1.2

$$\varepsilon = \frac{S_{in}}{1 - A(s)\cdot F(s)} = \frac{S_{in}}{1 - Gloop(s)}$$

and that of the feedback signal:

$$S_f = \varepsilon \cdot A(s)\cdot F(s) = \frac{S_{in}\cdot Gloop(s)}{1 - Gloop(s)}$$

In the case of $|G_{loop}|>>1$, we can neglect 1 at the denominator of Eq. 1.1, and the expression of the transfer reduces to:

Eq. 1.3

$$G(s) = \frac{S_o}{S_{in}} \cong -\frac{1}{F(s)}$$

This equation highlights how, under ideal conditions, the transfer does not depend on the forward block anymore, but depends only on the characteristics of the reaction block. This is a very important result because the forward block is affected by high tolerance, thus it will be enough to ensure a sufficient margin of tolerance of the feedback block. Moreover, as seen in Eq. 1.2, if G_{loop} is negative and very large in magnitude ($G_{loop}<<-1$), ε (error signal) tends to zero from below with increasing G_{loop} (in the absolute sense).

1.3.3 Negative and Positive Feedbacks

It is useful to think about the meaning of Eq. 1.1, which depends on the sign of the loop gain G_{loop}. It has, in fact, a crucial role in defining the characteristics of the feedback circuit.

Suppose that the two transfer functions A(s) and F(s) have different signs, such as $A>0$ and $F<0$. Consequently, $G_{loop}<0$. If a positive step is applied as the input signal S_{in} to this circuit (Fig. 1.7a), the forward block gives a positive signal at the output. The change in the output generates a feedback

signal S_f that gets to the summing node with an opposite sign with respect to the forcing signal and is then subtracted from the input signal (obtaining $\varepsilon < S_{in}$). This feedback, in which the feedback signal tends to reduce the fraction of the input signal actually applied to the forward amplifying stage, is called **negative feedback**. The feedback is negative whenever the product $A \cdot F$ is negative, i.e. when the loop gain G_{loop} is negative; in the example, F is negative, and A is positive.

Suppose instead that the two blocks that make up the loop have gains with the same sign (Fig. 1.7b), for instance $A>0$ and $F>0$, which therefore leads to $G_{loop}>0$. In this case, the signal that comes back to the summing node has the same sign as the forcing signal and is added to the input signal ($\varepsilon > S_{in}$). A system with these characteristics is called a **positive feedback** system. The feedback is positive whenever the product $A \cdot F$ is positive, i.e. when the loop gain G_{loop} is positive. It is clear then that in a positive feedback system, even a small interference at the input could lead to a divergence of the signal circulating in the system itself. In fact, it would continue to increase, making the output inevitably diverge (instability). Probably, the components involved are not damaged because they end up going out from their range of linearity.

Both positive feedback systems and negative feedback systems are used in electronic systems. For instance, oscillators and clock generators belong to the first category; amplifiers, shapers, and filters belong to the second. After having introduced the topology of the feedback systems and their classification, it is now possible to quantitatively estimate the reproducibility of the transfer function of an amplifier designed according to the scheme of H. S. Black. Differentiating Eq. 1.1 with respect to A, we obtain:

$$\frac{dG}{G} = \frac{dA}{A} \frac{1}{1 - Gloop}$$

If the loop gain is negative, and its magnitude is greater than unity, the percentage change of the forward transfer A(s) slightly or negligibly affects the transfer G within the system.

Ultimately, in a negative feedback system, reducing the gain by a factor of $(1-G_{loop})$ with respect to the gain of the forward block allows making the

circuit less sensitive to variations in A precisely by the same factor ($1-G_{loop}$). For a system that has positive feedback and that is stable ($G_{loop}<1$), we instead have a transfer greater than the forward one, but it is also sensitive to variations in A through the usual factor ($1-G_{loop}$).

These considerations have an immediate practical application. Suppose that we have to design an electronic amplifier with high stability of the gain (variation of ±0.1%). You might not think of making such an amplifier open-loop, i.e. by adopting the scheme of Fig. 1.6a. In fact, all the characteristics of such an amplifier would heavily depend on the parameters of the transistors, the passive elements used, their variation with temperature, etc. The alternative to this solution is given by a negative feedback circuit, like the one shown in Fig. 1.6b. If the goal of the design is the realization of a voltage amplifier with an amplification of 100±0.1%, it is very easy and cheap to use a stage with an amplification of $5\cdot10^4$±50% and then make use of negative feedback with a loop gain $|G_{loop}|=500$, obtaining the amplifier with the desired requirements.

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.

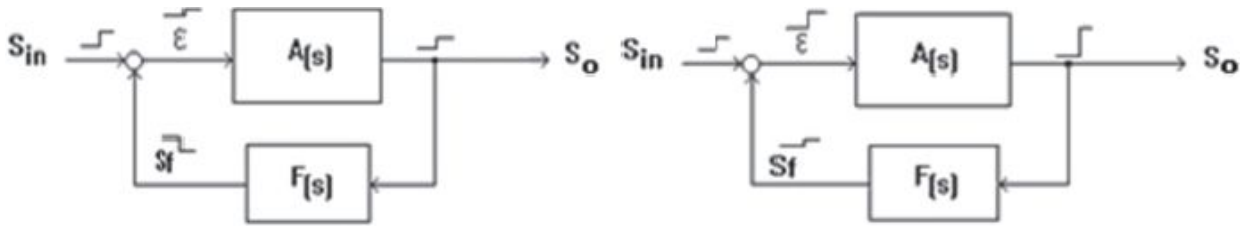


Fig. 1.7: Negative (left) and positive (right) feedback system.

1.3.4 Effect of the Feedback on the Bandwidth

Thus far, so as to introduce the properties of feedback systems, we have only considered the behavior of feedback amplifiers in a range of frequencies that were briefly indicated as center band or mid-frequency band, where the transfer functions A(s) and F(s) are independent of the frequency. Now, on the contrary, we extend the analysis to the more general case with the goal of determining the transfer function of a feedback circuit in the frequency domain and the influence that the feedback has on the transfer function itself.

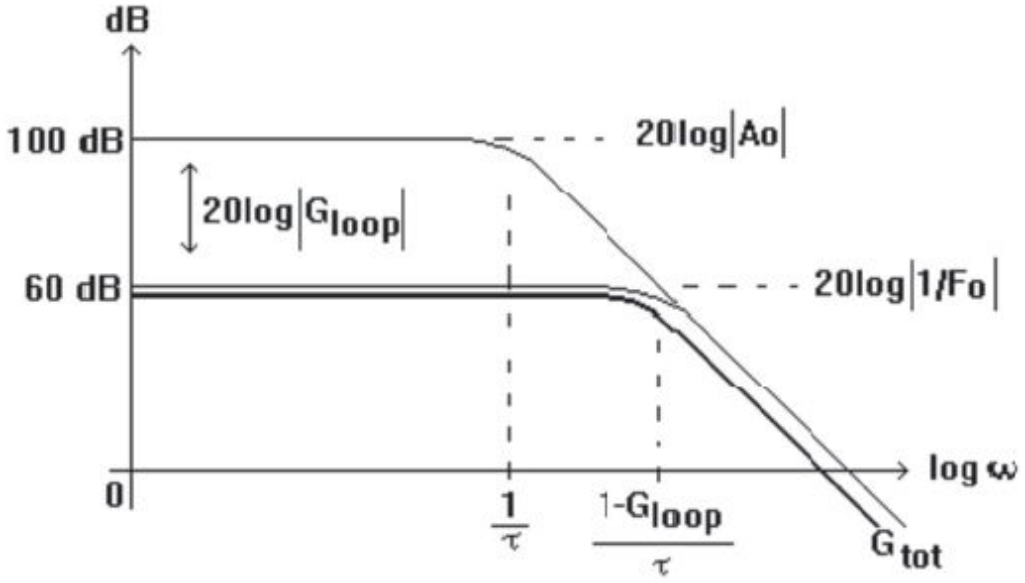


Fig. 1.8: Single pole system negatively feedback: Bode plot of the magnitude of the function A(jω), of the real transfer G(jω)and of the ideal transfer $1/F_0$.

An amplifier has been classified as a negative feedback amplifier if the sign of its loop gain at mid-frequencies is negative. This condition mathematically expresses the fact that the feedback signal S_f comes back to the summing node exactly out of phase with the signal S_{in} applied (Fig. 1.7a), giving a small control signal ε to the forward block A(s). Nonetheless, if in the forward stage and/or in the feedback one reactive elements are present, the functions A(s) and F(s) are not constant, but, as the frequency changes, they change in magnitude and phase. Therefore, in real amplifiers, there can be a frequency for which the additional phase introduced by the

reactive elements is ±180° with respect to the phase shift at the center band. For this frequency, the feedback signal comes back to the summing node in phase with the applied signal (Fig. 1.7b), and if the loop gain at that frequency is greater than 1, the signal is also amplified. Such a situation is obviously regenerative, and the system is unstable.

These intuitive considerations can be mathematically proved. In general, we can say that a negative feedback system is not unconditionally stable. The position of the poles of the transfer function (Eq. 1.1) depends on not only the reactive elements of the circuit, but also the value of the loop gain. The information about stability is therefore fully contained in the function $G_{loop}(j\omega)$.

Before addressing the complex problem of the frequency response of a feedback circuit (that will be widely discussed in chapter 3), it is important to analyze the particular case of a feedback circuit with a single pole in the forward block A(s) that has no zeros. The term G(s) can be analytically calculated according to Eq. 1.1:

$$G(s) = \frac{\frac{Ao}{1+s\tau}}{1-\frac{Ao \cdot Fo}{1+s\tau}} = \frac{Ao}{1-Ao \cdot Fo} \cdot \frac{1}{1+s\frac{\tau}{1-Ao \cdot Fo}}$$

The pole of such a transfer function is at the frequency $f_p=(1-G_{loop}(0))/(2\pi\tau)$. If the circuit is a positive feedback circuit, and $A_0 \cdot F_0>1$, the pole is real and positive, and the circuit is unstable. If, instead, $0<A_0 \cdot F_0<1$, there is an increase in the amplifier gain with respect to the value A_0, but also a reduction in the bandwidth. If, on the other hand, the feedback is negative, $A_0 \cdot F_0 <0$, we have a reduction in the amplification, but simultaneously an increase in the bandwidth by the usual amount $(1-A_0 \cdot F_0)= (1-G_{loop})$. It is easy to verify, for this single-pole system, that the **Gain BandWidth Product (**GBWP) remains constant whatever the value of G_{loop} is. Hence, in the design phase, we can modify the loop gain according to the specifications by conveniently reducing the gain to broaden the bandwidth. Strictly speaking, this simple conclusion is valid only in the case of a single-pole system. In general, it is much more complicated to determine the effect

of the feedback on the singularity of the circuit. However, in the spirit of the result just obtained, we can generally say that *the negative feedback, together with other benefits in terms of stabilization of the transfer and of impedential matching, also has a beneficial effect on the frequency response of the circuit since it tends to widen the bandwidth*.

Fig. 1.8 shows the trend of the magnitude of the forward gain A(jω), of the real transfer G(jω), and of the ideal transfer $-(1/F_0)$ for the single-pole negative feedback system. The distance between the curve of A(jω) and the line $|1/F_0|$ is just:

$$20\log |A_0/(1+j\omega\tau)| - 20\log |1/ F_0| = 20 \log |G_{loop} (j\omega)|$$

and quantitatively represents the amount of the feedback. The frequency at which A(jω) crosses the line $1/F_0$ is the frequency at which $|G_{loop}(j\omega)|=1$.

At low frequencies, where $|G_{loop}(j\omega)|>>1$, the real transfer coincides with the ideal transfer, i.e. $G(j\omega)\cong-(1/F_0)$. With decreasing $|G_{loop}(j\omega)|$, G(jω) significantly deviates more and more from $-(1/F_0)$, and for $\omega\rightarrow\infty$, the curve of |G(jω)| and that of |A(jω)| overlap, as expected from Eq. 1.1 for $|G_{loop}(j\omega)| \rightarrow 0$. It follows that only at low frequencies, where $|G_{loop}(j\omega)|>>1$, the transfer of the system depends only on the elements that are involved in the feedback block and is thus not influenced by the margin of tolerance and by the active elements present in the forward block. With increasing frequency, on the contrary, the transfer is influenced again by the forward block.

1.4. Operational Amplifier (OpAmp)

The operational amplifier (OpAmp) is a differential voltage amplifier with a very high gain, which is widely used as an amplifying block in feedback circuits. For such circuits, it is very important to have an amplifying stage with a very high gain to obtain a high loop gain. In the design of analog circuits, the OpAmp is widely used due to its almost ideal characteristics and because it allows amplifying, controlling, and generating waveforms having frequencies ranging from DC to many MHz. The OpAmp can be also used to develop, in a simple way, the following algebraic functions: sum, difference, multiplication, division, integration, and differentiation.

The inside of the operational amplifier is composed of a series of transistor amplifiers while its circuit representation is shown in Fig. 1.9. The inputs, identified with the symbols + and –, are called the non-inverting input and the inverting input, respectively. The corresponding voltages measured with respect to the ground are denoted by V_p and V_n, respectively, while the output voltage is denoted by V_o.

Putting two different voltages to the input terminals V_p and V_n, we obtain an output V_o related to the differential voltage $V_d=V_p–V_n$ through the differential gain A_d. Applying the same voltage to these two terminals, the output is linked to the common mode voltage $V_{cm}=(V_p+V_n)/2$ through the common mode gain A_{cm}. It can be written:

$$V_O \triangleq A_d \cdot \left(V_p - V_n\right) + A_{cm} \cdot \left(\frac{V_p + V_n}{2}\right)$$

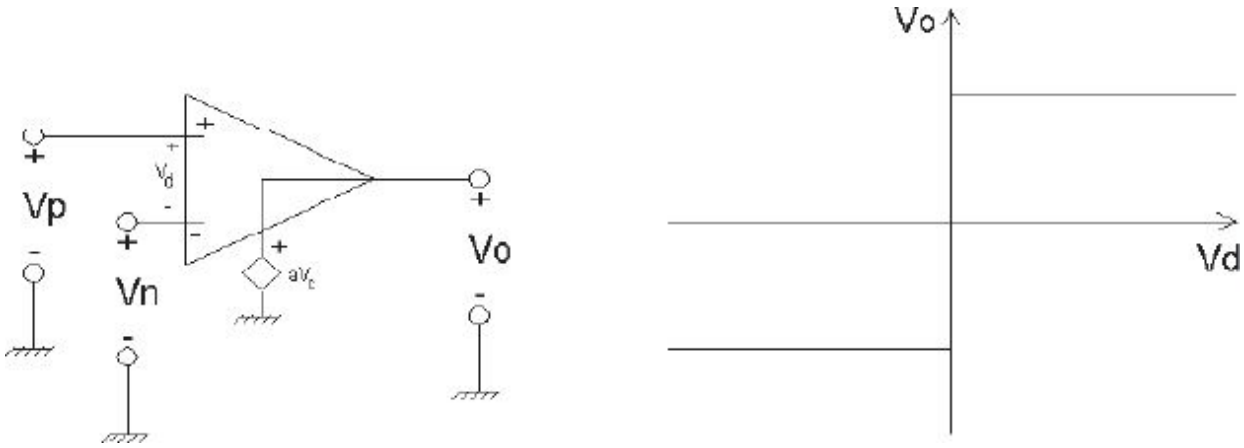


Fig. 1.9: The ideal OpAmp: circuit diagram (left) and input/output characteristic (right).

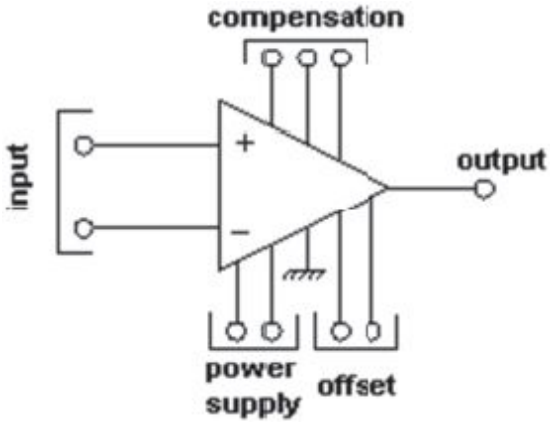


Fig. 1.10: Circuit diagram of a real OpAmp.

The ratio between A_d and A_{cm} expressed in dB is the CMRR, already defined as:

$$CMRR \triangleq \left(A_d / A_{cm}\right)_{dB} = 20 \cdot \log_{10}\left(\frac{A_d}{A_{cm}}\right)$$

Assuming A_{cm} is equal to zero, one can draw the equivalent circuit of the OpAmp as in Fig. 1.11 where, for the sake of completeness, the differential input resistance R_d and the output resistance R_o have been added. The approximation A_{cm}=0, when it can be considered valid, allows expressing the relationship between V_o and V_d as follows:

$$V_o \cong A_d \cdot V_d = A_d \cdot (V_p - V_n)$$

Rewriting the previous equation in a different way, we get:

$$V_d = \frac{V_o}{A_d}$$

Thus, we can obtain the differential input voltage V_d (and not the individual voltages V_p and V_n), knowing A_d and V_o. Bearing in mind that A_d takes values that far exceed a million and that the output voltage of an operational amplifier is usually less than 15V, it can be concluded that the operational

amplifier works in a linear region when the differential voltage V_d is some μV or even few nV.

1.4.1 Parameters of the Ideal OpAmp

Let us now define the characteristics of the ideal OpAmp shown in Fig. 1.9:

- 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;
- Output voltage equal to zero when the input voltages are equal to zero;
- No change of the parameters as the temperature changes;
- Infinite Slew Rate (SR); maximum speed at which the output can change, in V/μs;
- Response time equal to zero;
- Equivalent input current noise (i^2_n) and voltage noise (v^2_n) equal to zero;
- Input bias current (I_b) equal to zero;
- Infinite immunity to the change of power supply (PSRR).

Actually none of the above parameters is found in a real OpAmp, although some of them may approach their ideal values. For example, the low noise OpAmps have equivalent input voltage noise of about 2nV/√Hz, the OpAmps with FET input or in MOS technology have very high input impedance, input bias, and offset currents in the order of pA, there are operational amplifiers with very high slew rates, in the order of 100V/μs, and OpAmps with unity gain frequency of few GHz. To get an idea of what the market offers, just look through any catalog of manufacturers of linear electronic products.

1.4.2 Parameters of the Real OpAmp

The real operational amplifier is slightly different from the ideal OpAmp described until now. The real OpAmp, in fact, has always to be biased and, depending on the application and on the particular device chosen, may need

the input offset cancellation and/or the compensation in frequency. For this reason, the circuit symbol of the real operational amplifier is not what is shown in Fig. 1.11a, but is the one shown in Fig. 1.10, where, in addition to the inputs and the output, the terminals that allow the biasing, the offset cancellation, and the frequency compensation are highlighted. In general, the number of pins available depends on the type of the OpAmp chosen and on its characteristics.

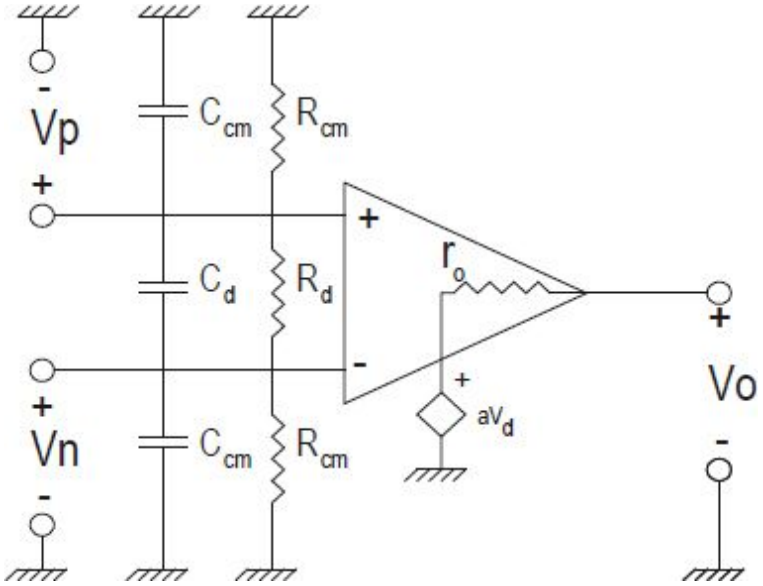


Fig. 1.11: Equivalent impedances of an OpAmp.

In the datasheet of every OpAmp, there is a section that lists the values of the parameters that are important for that specific operational amplifier. Here are a few:

Differential voltage gain A_{dO}: the differential voltage gain is defined as the ratio between the voltage measured at the output and the corresponding continuous differential voltage between the input terminals.

Common mode voltage gain A_{cmO}: the common mode voltage gain is defined as the ratio between the voltage detected at the output and the continuous common mode voltage applied to the input terminals. Usually, A_{d0} can vary from 1,000 to 1,000,000, depending on the type of the OpAmp considered while A_{cm0} has to assume values of hundreds of times lower than A_{d0}.

Open-loop unity-gain frequency: it is the frequency in correspondence of which the voltage gain crosses the 0dB axis, i.e. the voltage gain becomes equal to one. Such a frequency can, depending on the operational amplifier, take values ranging from a few kHz to some GHz. In general, the unity-gain frequency is denoted by f_t while, in the particular case where the frequency response of the OpAmp is that of a single-pole circuit, this frequency is called the "gain-bandwidth product".

Unity-gain rise time: it is the time t_r that the OpAmp, in voltage buffer configuration (i.e. with a voltage gain equal to 1), takes to go from 10% to 90% of the final value of the voltage. For this configuration, the bandwidth BW, close to f_t, is defined as follows: $BW=0.35/t_r$.

Full power bandwidth (FPB): it is the maximum frequency for which, given a sine wave with an appropriate amplitude at the input, an undistorted sine wave with the maximum amplitude obtainable can be seen at the output terminal. Such bandwidth is typically 10 or 100 times smaller than f_t.

Slew Rate: it gives an account of the ability of the OpAmp to quickly make large changes in voltage at its output. In fact, when we have a high voltage step at the input of a buffer, the response we get is not the typical "small signal" of a single-pole circuit with a rise time t_r, but is limited to a ramp with a maximum slope equal to the SR of the OpAmp, which can vary between 1 and 1'000V/μs.

Input impedance: it identifies the parasitic resistance and the capacitance related to the inputs of the operational amplifier. If such measurement is made between the input terminals V_p and V_n, we talk about differential impedance while if we use the ground and one of the two terminals, we get the common-mode impedance. As shown in Fig. 1.11, common-mode impedance is represented by a resistance R_{cm} on the order of hundreds of MΩ and by a capacitance C_{cm}, whose typical value is 3pF, while the differential-mode impedance includes a capacitance C_d on the order of 3pF and a resistance R_d whose value depends on the kind of the input stage used in the OpAmp. If the input stage is made with FET transistors, Rd has an order of magnitude of 10^{12} ohm while, in the case of OpAmp with an input stage made with BJTs, R_d takes values in the order of MΩ.

In the case of an OpAmp mounted on a printed circuit, the input impedance also contains two inductive parasitic terms, whose typical value is of a few nH, due to the parasitic inductance of the metal track.

Output impedance: it is outlined with an equivalent resistance measurable at the output terminal. This resistance can vary from tens of Ω in the case of BJT OpAmps to tens of kΩ in the case of MOSFET OpAmps.

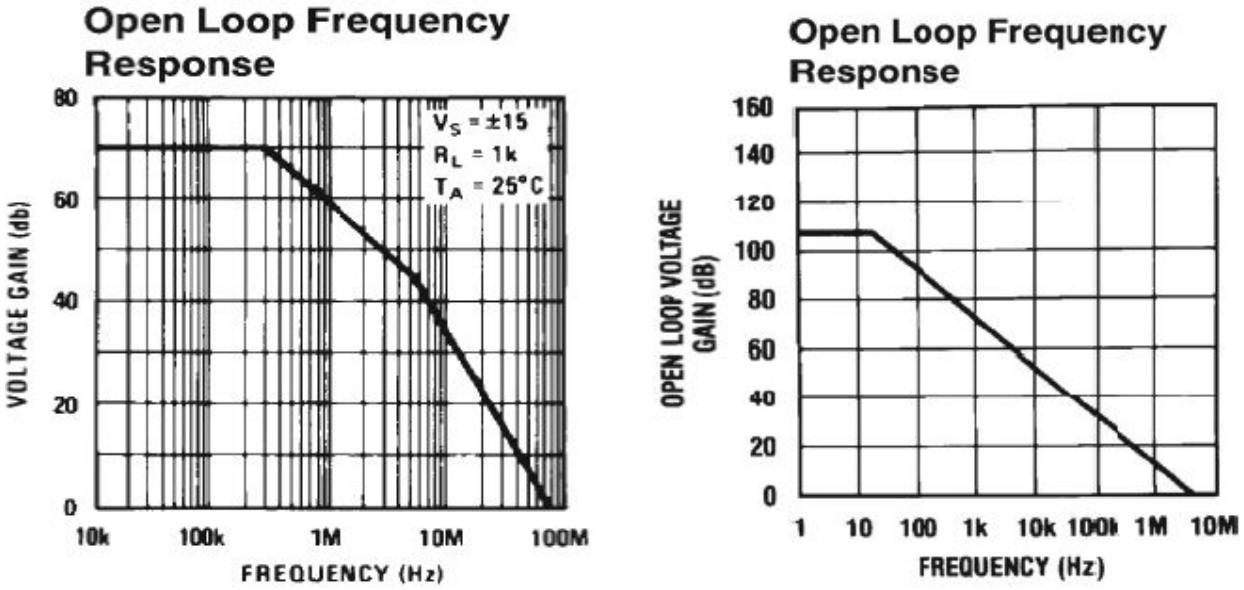


Fig. 1.12: Open-loop response of a non-compensated OpAmp (left) and of a compensated one (right).

1.4.3 Open Loop Frequency Characteristics of an OpAmp

When the OpAmp is used to amplify AC signals, we have to refer to the magnitude and phase Bode plots of the considered operational amplifier. A typical magnitude Bode plot of a *compensated OpAmp* is shown in Fig. 1.12, on the right: at low frequencies, the gain is constant and equal to A_{d0} while the –3dB frequency (open-loop bandwidth), compared to the value A_{d0}, identifies the position of the first pole of the OpAmp.

If there are other singularities before the intersection of the curve and the 0dB axis, the operational amplifier is said to be *non-compensated*, and the gain that corresponds to the frequency of the second pole is called A_{min}. This is an important parameter since it is the minimum closed-loop gain that can

be obtained from a feedback circuit in which the amplifying block is the selected OpAmp.

In the case of a compensated OpAmp, the intersection of the open-loop curve and the 0dB axis, indicated with the frequency f_t, coincides with the GBWP. In fact, as has been mentioned above, for each point where the plot has a slope of -20dB/dec, we have:

$$GBWP = f \cdot A(f)$$

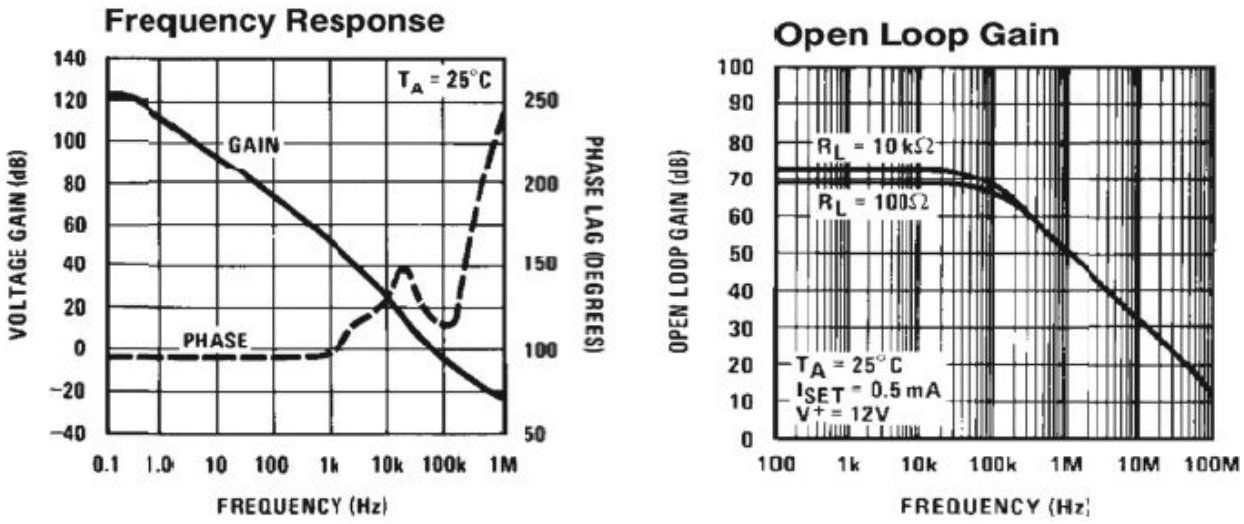


Fig. 1.13: Open-loop Bode plots of two OpAmps.

For the compensated OpAmps, the GBWP is fixed while for the non-compensated operational amplifiers, that value is either not defined or considered coincident with f_t and varies according to the type of the compensation employed. The GBWP is an important parameter for compensated operational amplifiers since it indicates the maximum bandwidth achievable with a voltage buffer containing the considered OpAmp. The more the distance is between the second pole and the GBWP, the larger the **phase margin**, a parameter that gives account of the stability of the circuit, will be.

OpAmp manufacturers provide Bode plots for frequencies lower and higher than the crossing frequency of the 0dB axis, as highlighted in Fig. 1.12. Nevertheless, remember that, regardless of the plots provided, to design a reliable feedback circuit containing OpAmps, we have to know the

trend of its Bode plots up to a decade after the 0dB axis and the phase margin of the amplifying block.

1.4.4 Internal Architecture of the OpAmp

Knowing the internal architecture of an OpAmp is important to understand the compensation methods for the input offset voltage and the frequency compensation methods of the non-compensated OpAmps. Given the diversity and the complexity of the circuit diagrams that are the basis of the OpAmps provided by the manufacturers, here we show just a generic block diagram, such as that shown in Fig. 1.14.

Input differential amplifier: this block determines the stability of the gain, the common mode rejection, the bias drifts, the input impedance, the slew rate, the bandwidth, and the noise of the OpAmp. Since the following stages have a negligible effect on these parameters, it is clear how important the design of good quality input stages is. Usually, the input stage can be schematized as in Fig. 1.15, although the actual circuit is much more complex. The current generator is designed such that the input stage has a high CMRR. The output is differential.

Second differential amplifier: this block is very similar to the previous one. Often, the current source is replaced by a resistor. The seven critical parameters determined in the input stage become up to a hundred times less restrictive in this second stage since the signal has already been amplified by the first stage. In this block, unlike the first, the output is single-ended.

Level translator amplifier: often, the level translator amplifier is introduced to have a translation of the DC voltage and to increase the gain of the total circuit. This translation is necessary because the quiescent output voltage of an OpAmp must be zero if the differential voltage is zero, and the outputs of the two stages analyzed above do not meet such a condition. To implement this stage, a single transistor in common source configuration is used, as shown in Fig. 1.15. The load impedance consists of two diodes and a resistor: this creates a voltage divider compensated to drive the output power stage. This stage has also the task of transforming the high output impedance of the second amplifier to low impedance, thus allowing the driving of the last stage.

Output power amplifier: it provides high current gains, large bandwidth, and low output impedance. Since this stage must be able to drive devices

outside of the OpAmp package, it must be capable of delivering high output currents. To protect the IC against an overload current, protections are put on the output stage transistors to avoid damage. Usually, the compensation network that we previously referred to is placed on the second or on the third stage of the OpAmp, depending on the type of the compensation implemented.

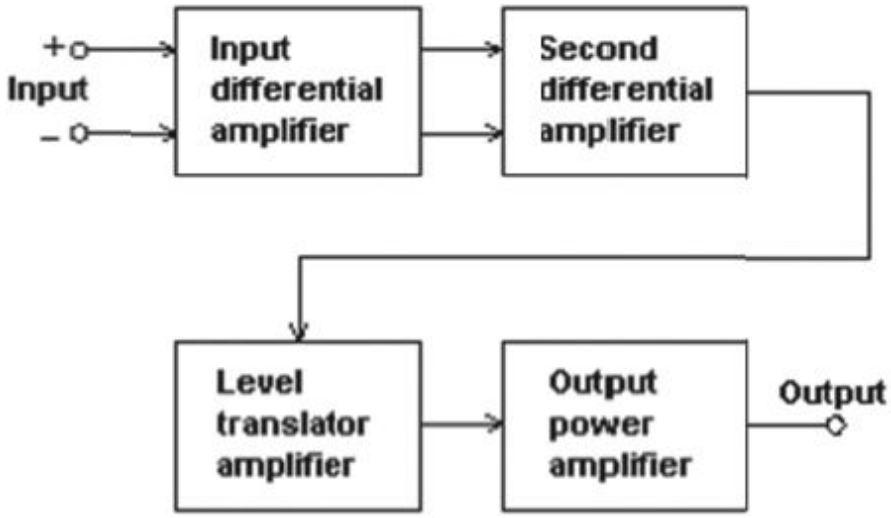


Fig. 1.14: Block scheme of the architecture of the OpAmp.

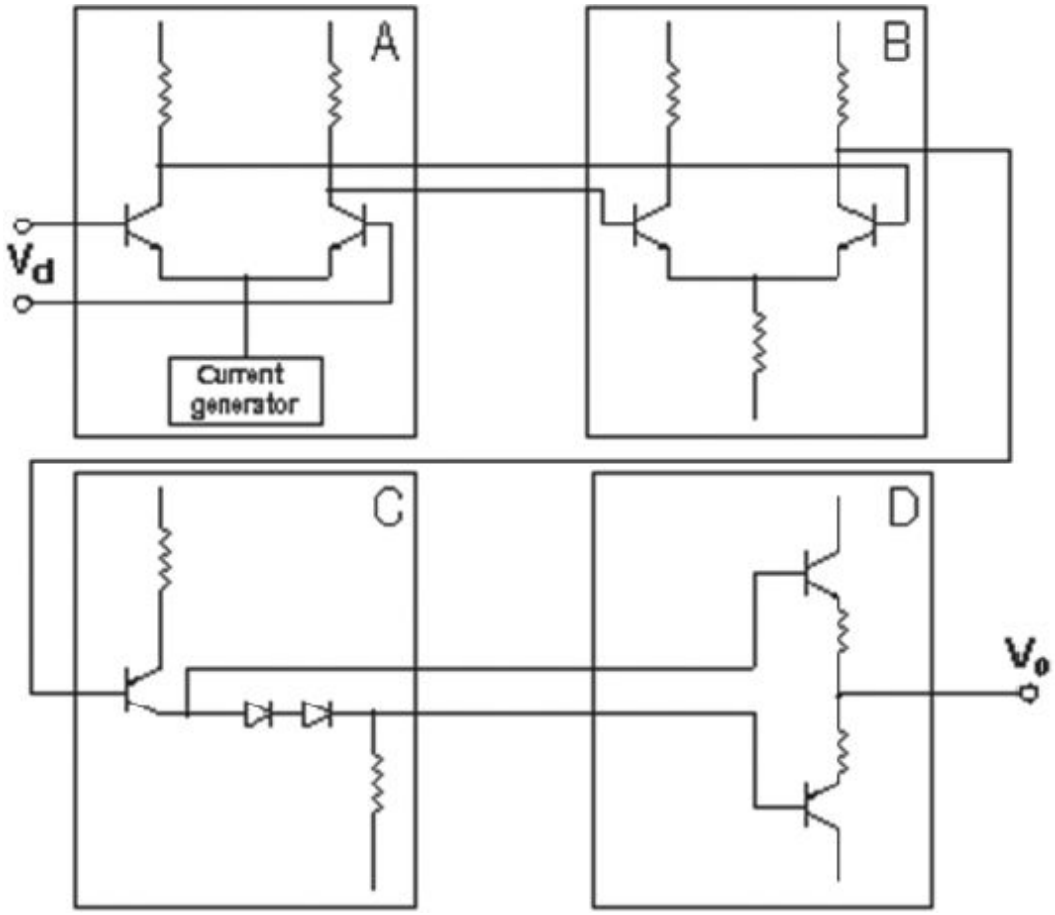


Fig. 1.15: Main stages of an OpAmp: (A) Input differential amplifier. B) Second differential amplifier. (C) Level translator amplifier. (D) Output power amplifier.

1.4.5 Compensated and Non-Compensated OpAmps

Having described the general characteristics of the OpAmps, we now consider operational amplifiers available on the market: compensated, non-compensated, and decompensated, focusing on their characteristics and the differences between them.

Compensated operational amplifiers

The typical open-loop frequency response of an internally compensated OpAmp is shown in Fig. 1.12: its characteristic is to have a constant gain-bandwidth product. The idea behind the internal compensation is to free the designer of the problem of stability, at least as regards circuits with resistive or capacitive feedback, as in the case of amplifiers and integrators, respectively. The manufacturer achieves this goal by ensuring that the phase rotation introduced by the amplifier never reaches *–180°* (one of the necessary conditions for the stability of the circuit).

If resistive feedback is applied to such operational amplifiers, since resistance does not introduce a phase shift, the phase margin of the circuit depends only on the amplifying block. In the case of the OpAmp in Fig. 1.16, on the left, the poles are located at a frequency of 2Hz and 2MHz while the GBWP is equal to 1MHz; the phase shift introduced by the device is therefore equal to –arctg(1MHz/2Hz)–arctg(1MHz/2MHz)≈–117°. By using resistive feedback with this device, assuming that $0<F(s)\leq 1$ (see Fig. 1.7), the minimum phase margin guaranteed will be $\phi_m = 63°$.

Internally compensated operational amplifiers, having a constant GBWP, are also called unconditionally stable. For the 741 OpAmp family, the constancy of the GBWP is accomplished with a 30 pF built-in capacitance placed between the input and the output of the second stage. Having a device with unconditional stability means to pay the price for very low open-loop bandwidth: in the case of the OpAmp, we are talking about bandwidth of only 2Hz.

There are applications where the designer wants to have very large open-loop bandwidth. Therefore, it is possible to find compensated OpAmps with large bandwidth, which are, however, extremely expensive or to use internally non-compensated OpAmps.

Non-compensated operational amplifiers

The typical response of a non-compensated OpAmp is shown in Fig. 1.16, on the right. What distinguishes it from that of an internally compensated device is the presence of much wider open-loop bandwidth (in figure, it is of 300kHz instead of 2Hz) and a larger phase shift. While, in the case of

internally compensated OpAmps, the slope of the magnitude Bode plot is related to the value of the integrated capacitance, for non-compensated devices, the whole response can be given by the sum of the frequency responses of the individual stages.

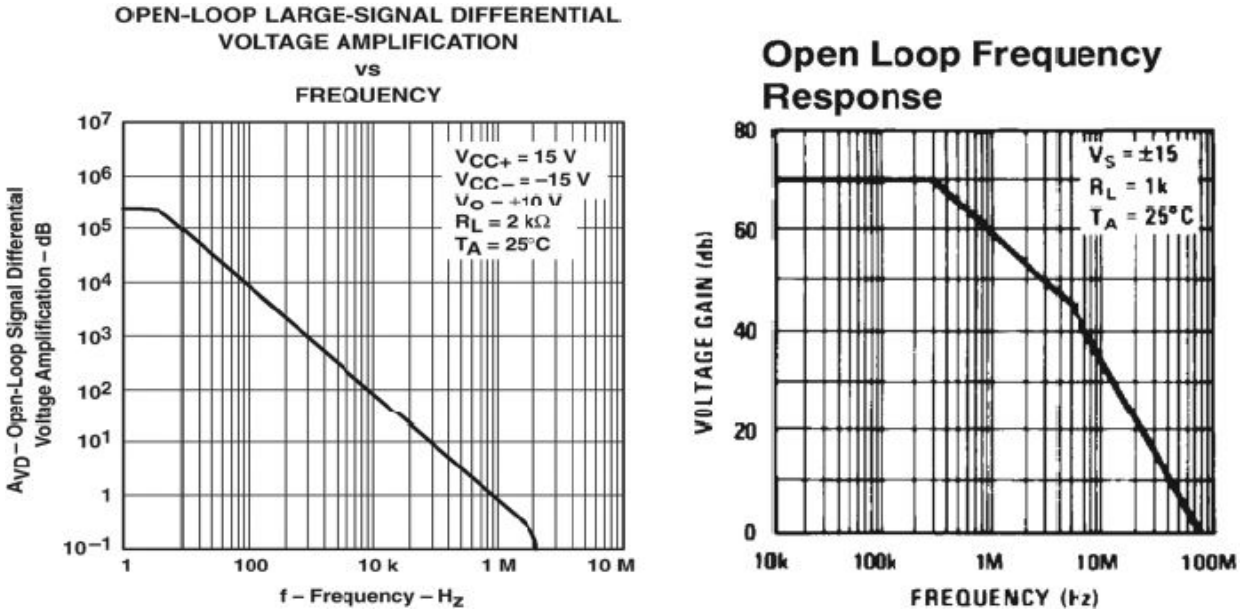


Fig. 1.16: Frequency response of a compensated OpAmp (left) and of a non-compensated one (right)).

For practical use, we can assume that the open-loop response of a non-compensated operational amplifier has two poles. The higher order singularities are deliberately ignored all the more as they go in the range of frequencies where $|A_d|$ is well below 0dB. If, despite everything, we want to consider higher order singularities, we must remember that there are practical limitations that make their determination uncertain. These factors are the uncertainties in determining the dc gain A_0 and the frequencies of the poles f_1, f_2 given by the thermal drift, aging, changes in the technological process, etc.

From the point of view of the stability, the most obvious difference between the plots of a compensated OpAmp and a non-compensated one is that the phase shift introduced by the latter can exceed –180° This is not surprising since each pole introduces a slope of –20 dB/dec and a phase shift of –90°

In the case of the frequency response of Fig. 1.16, we have A_0=70 dB, f_1=300kHz, and f_2=6MHz. Since the 0dB frequency is 80Mhz, assuming that the next pole is far from the unity-gain frequency, we can easily find that the phase margin of a buffer made with this OpAmp is equal to ϕ_m=180°-arctg(80MHz/300kHz)-arctg(80MHz/6MHz)≈ 4.5°, which is far below the acceptable minimum of 45°. Using the formula $f_1 \cdot A_0 = f_2 \cdot A_{min}$, we can obtain the value of A_{min}≈44dB. This value is the minimum amplification that can be achieved using the OpAmp under examination in a feedback circuit without introducing an additional compensation network into the circuit.

Decompensated operational amplifiers

In addition to the compensated and non-compensated operational amplifiers, there are also decompensated operational amplifiers. These devices can be seen as compensated OpAmps with bridge capacitance of small values with respect to the classical compensation. They are a trade-off between the benefits of the devices mentioned above. Indeed, internally compensating an operational amplifier with bridge capacitance of small value, we get a transfer function with two poles before f_t and with A_{min}, which can vary between 2 and 25. These operational amplifiers may be used without further compensation for ideal gain values greater than A_{min}.

1.4.6 Stability of Non-Compensated and Decompensated OpAmps

Now consider the stability of a non-compensated or decompensated OpAmp for different values of resistive feedback. For this study, we refer to the scheme on the left of Fig. 1.17, where one of the resistors is variable. Thus, we can freely change the gain of the feedback block $\beta=R_1/(R_1+R_2)$. To test the stability of the circuit, we have to plot the trend of $1/\beta$ on the graph of the open-loop gain A(s) of the OpAmp. At this point, we can identify the crossing frequency f_c between the two plots and calculate the phase margin of the transfer function. *We have to choose a minimum phase margin of 45° to ensure stability even when ambient parameters, such as drift and aging of the device, change.*

As an example, we report the results of tests performed with the SPICE simulation program using the LM6165 operational amplifier by National Semiconductor that is not compensated. Depending on the gain set by the ratio $1+R_2/R_1=1/\beta$, the circuit may be more or less stable. Basically, if we set a gain equal to or greater than A_{min}, the circuit will always be stable. If, on the contrary, $1/\beta<A_{min}$ holds, the circuit will be conditionally stable or unstable, depending on whether the intersection is between slopes of –40dB/dec or –60dB/dec, respectively (see Fig. 1.18, right).

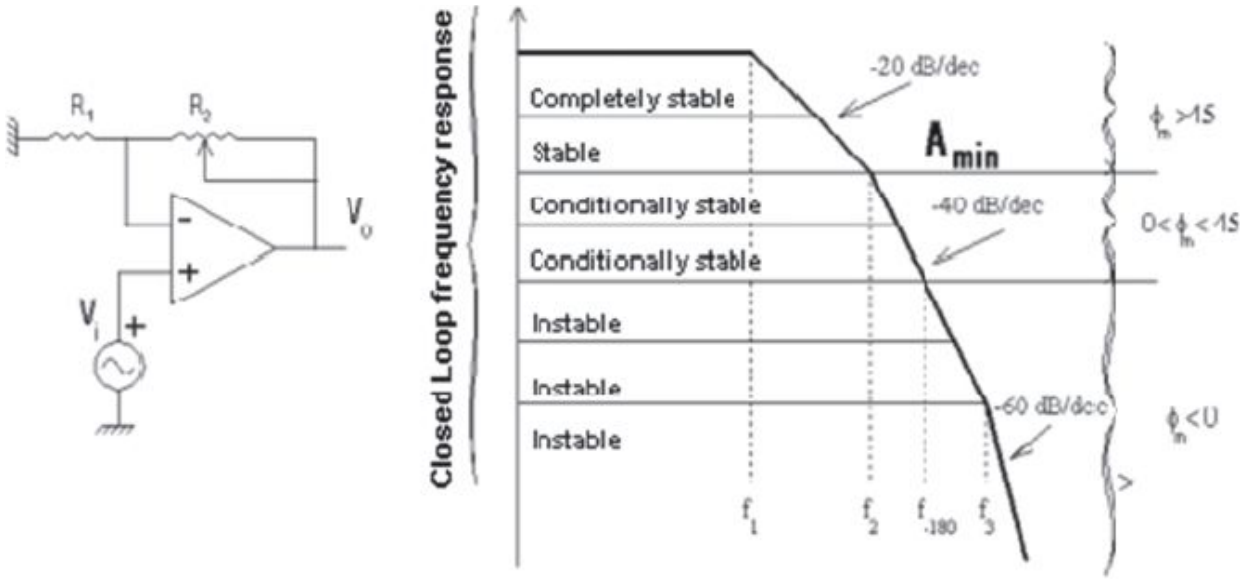


Fig. 1.17: Circuit for the stability analysis (left) and effect on the stability of a stage with an OpAmp with three poles before the 0dB axis (right).

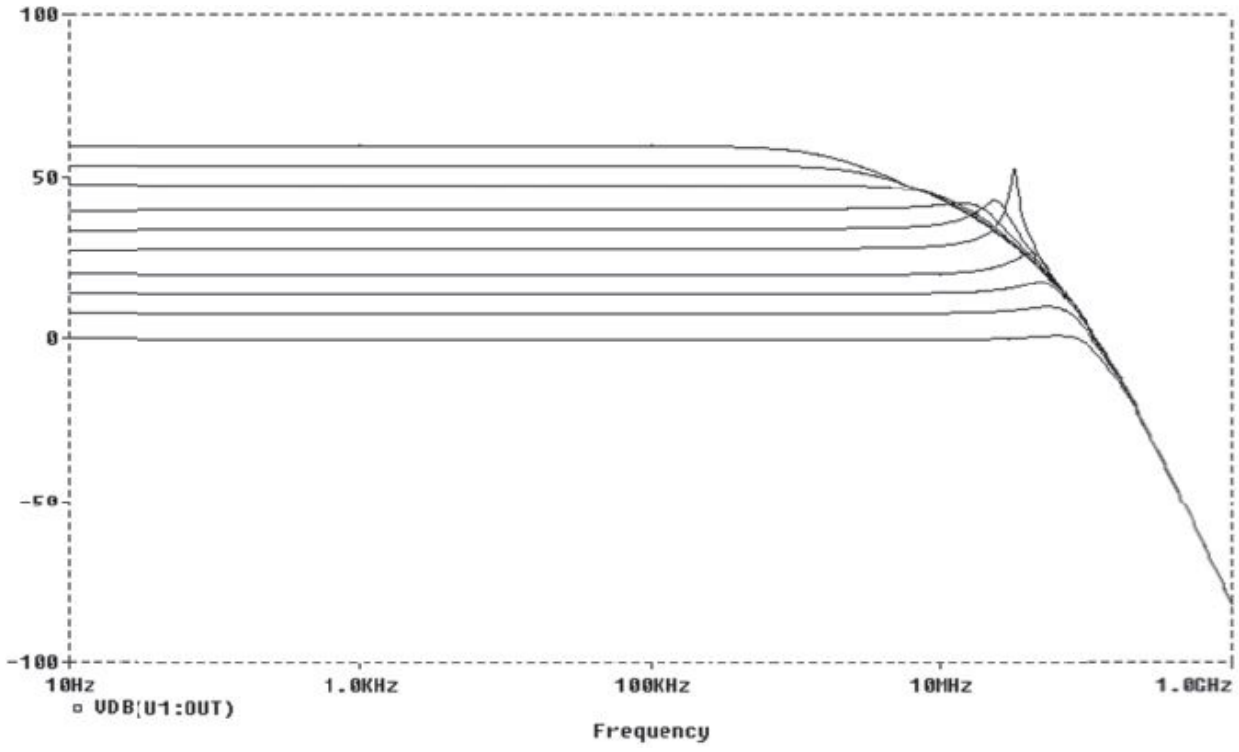


Fig. 1.18: Result of the Spice simulation with the LM6165 as the system gain $1+R_2/R_1$ changes.

The circuit used for the simulation is shown in Fig. 1.17a. The gain (in this case coincident with 1/β) varied from 1000 to 1, and the results are

plotted in Fig. 1.18. For low values of β, the function 1/β intersects the transfer function of the OpAmp in the area at the top, i.e. where the slope of $|A_d|$ is equal to –20dB/dec. In this case, the phase shift is about –90° therefore the phase margin is around 90° and the circuit is certainly stable. As 1/β decreases (i.e. as β increases) until it reaches A_{min}, we can note an increase of the peak of the closed-loop transfer function around the frequency f_C. For $1/\beta=A_{min}$, since the poles of the OpAmp are far apart more than a decade, the phase margin is only 45°. The stability worsens more and more with decreasing 1/β.

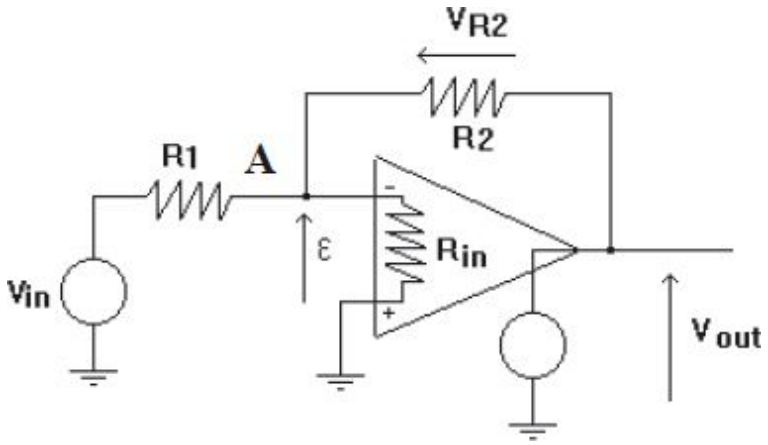


Fig. 1.19: Inverting configuration.

1.5. Consequences of Feedb on OpAmps

1.5.1 OpAmp in Inverting Configuration

Let us analyze the inverting configuration shown in Fig. 1.19. The output voltage is expressed by the following relation:

$$V_{out}=G_0\cdot(V^+-V^-)=-G_0\cdot V_A$$

(V_A: potential at node A). The current flowing through the resistor R_1 is:

$$I_1 = (V_{in} - V_A) / R_1$$

while the current flowing through the input resistance R_{in} is:

$$I_3 = (V_A / R_{in}).$$

Applying Kirchhoff's current law at node A, we can calculate the current flowing through the resistor R_2 as the difference between the currents I_1 and I_3:

$$I_2 = I_1 - I_3 = \frac{V_{in} - V_a}{R_1} - \frac{V_a}{R_{in}}$$

Applying Kirchhoff's voltage law, we obtain:

$$V_{out} = V_a - V_{r2} = V_a - R_2 \cdot I_2 = V_a - R_2 \cdot \frac{V_{in} - V_a}{R_1} + \frac{R_2 \cdot V_a}{R_{in}} = V_a \left(1 + \frac{R_2}{R_1} + \frac{R_2}{R_{in}}\right) - \frac{R_2}{R_1} \cdot V_{in}$$

Since $V_a = \frac{-V_{out}}{G_o}$, we derive the complete expression of V_{out}, which is equal to:

Eq. 1.4

$$V_{out} = \frac{-V_{out}}{G_o}\left[1 + R_2\left(\frac{1}{R_1} + \frac{1}{R_{in}}\right)\right] - \frac{R_2}{R_1} \cdot V_{in}$$

$$V_{out}\left[1 + \frac{1}{G_o}\left(1 + R_2\left(\frac{1}{R_1} + \frac{1}{R_{in}}\right)\right)\right] = -\frac{R_2}{R_1} \cdot V_{in}$$

IDEAL TRANSFER

If G_0 has a very high value (infinite in the case of the ideal OpAmp), we have:

$$V_{out} = -\frac{R_2}{R_1} \cdot V_{in}$$

The two resistors R_1 and R_2 represent the feedback block β; $-(R_2/R_1)$ is the ideal transfer of the inverting configuration of the operational amplifier. Since $V_A=-(V_{out}/G_0)$, replacing V_{out}, we obtain:

$$V_a = -\frac{R_2}{R_1} \cdot \frac{V_{in}}{G_o}$$

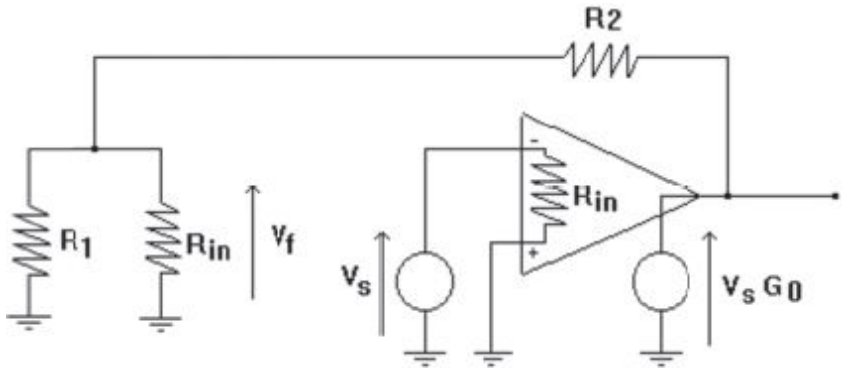


Fig. 1.20: Calculation of the loop gain.

Therefore, if G_0 were infinite, V_A would be equal to zero. This voltage V_A is precisely the error signal that is made infinitesimal in the ideal case: *it is as if the OpAmp changed its output to make the point A tend to the same potential of the non-inverting terminal (in this case, connected to the ground). For this reason, the node A is called the* **virtual ground**, *because it goes to the same potential of the non-inverting terminal, although the two are not electrically connected to one another. For this reason, since* $V_{R1}=V_A=0V$, *the current flowing in* R_1 *is* $I_1=V_{in}/R_1$ *and is the same as that*

flows in R_2. Since $V_A=0$, the output voltage will be equal to the voltage across R_2, i.e. $V_{out}=(V_A-I_1 \cdot R_2)=-(R_2/R_1) \cdot V_{in}$.

Applying a step at the input increases V_A while V_{out} decreases since the OpAmp would see at its input a negative error signal ε. Thanks to the feedback network, part of the signal V_{out} comes back to the node A, thus making the differential signal equal to zero. It is the feedback, therefore, that makes the node A fixed in voltage and equipotential to the voltage applied to the non-inverting terminal.

REAL TRANSFER

Let us analyze the circuit shown in Fig. 1.19 in real conditions. The Eq. 1.4 becomes:

$$\frac{V_{out}}{V_{in}} = \frac{-\frac{R_2}{R_1}}{1+\frac{R_1 \cdot R_{in} + R_2 \cdot R_{in} + R_2 \cdot R_1}{G_o \cdot R_1 \cdot R_{in}}} = \frac{-\frac{R_2}{R_1}}{1+\frac{1}{\frac{G_o \cdot R_{in} \cdot R_1}{R_{in} \cdot R_1 + R_2 \cdot (R_1 + R_{in})}}} = \frac{-\frac{R_2}{R_1}}{1+\frac{1}{\frac{G_o \cdot (R_1 // R_{in})}{(R_1 // R_{in}) + R_2}}} = \frac{\frac{1}{\beta}}{1+\frac{1}{A\beta}}$$

where: $Gid = \frac{1}{\beta} = -\frac{R2}{R1}$ and $Gloop = -A\beta = -\frac{Go \cdot (R1 // Rin)}{(R1 // Rin) + R2}$.

The real transfer is hence expressed by the following relationship:

Eq. 1.5

$$G_r = \frac{G_{id}}{1 - \frac{1}{G_{loop}}}$$

where G_{id} is the ideal gain, and G_{loop} is the loop gain. Since, in general, it is difficult to find A directly, it is recommended to calculate G_v through G_{id} and G_{loop}.

LOOP GAIN (G_{loop})

To calculate G_{loop}, we must do the following:

1- turn off forcing generators (by shorting voltage sources and opening current ones);
2- ideally break the feedback loop at an arbitrary point;
3- reconstruct the impedance seen by the circuit before and after the cut (Fig. 1.20) so that the behavior of the circuit is the same as before;
4- apply a probe signal (voltage V_S or current I_S) to the point where we made the “cut” and assess the voltage V_f or current I_f on the other side of the loop.

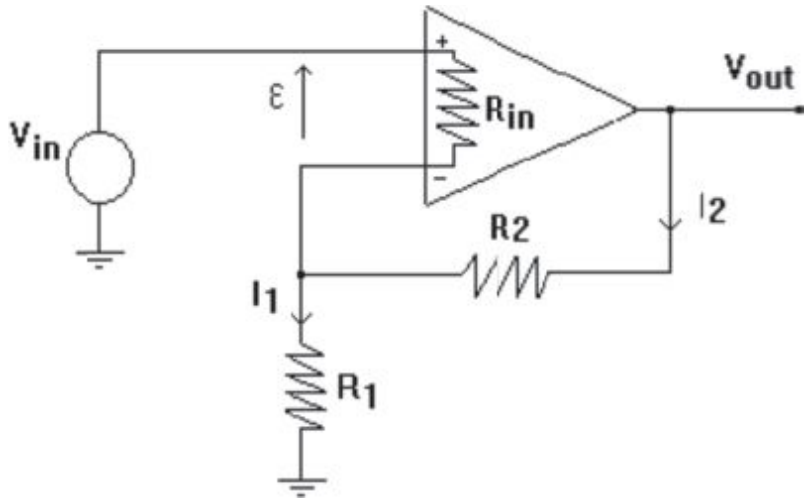


Fig. 1.21: Non-inverting configuration.

It follows: $G_{loop}=V_f/V_s$ or $G_{loop}=I_f/I_s$

In the following example, we suppose to inject a voltage signal V_f (Fig. 1.20).

$$V_f = -G_o \cdot V_s \frac{(R_1 // R_{in})}{R_2 + (R_1 // R_{in})}$$

$$G_{loop} = \frac{V_f}{V_s} = -G_o \frac{(R_1 // R_{in})}{R_2 + (R_1 // R_{in})}$$

Therefore, the real transfer is:

$$G_r = \frac{G_{id}}{1 - \frac{1}{G_{loop}}} = \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{A \cdot \beta}} = \frac{A}{1 + A \cdot \beta}$$

And then:

$$A = -\frac{R2}{R1} \cdot \frac{Go(R1 // Rin)}{R2 + (R1 // Rin)}.$$

NOTE: it is possible to make the cut at any point of the circuit provided that the impedance seen from that point is “reconstructed” exactly. For this reason, it is advisable, however, to operate in the way just shown; that is,

cutting close to R_{in} since R_{in} is usually high (theoretically infinite) and can thus be considered an open circuit without the need to “reconstruct” it.

1.5.2 OpAmp in Non-Inverting Configuration

Fig. 1.21 shows the schematic of an operational amplifier in the non-inverting configuration.

IDEAL TRANSFER

If G_{loop} tends to infinite, then the error signal ε tends to zero, we have $I_1=I_2=(V_{in}/R_1)$, and:

$$V_{out} = V_{R1} + V_{R2} = V_{in} + \left(\frac{V_{in}}{R_1}\cdot R_2\right) = V_{in}\cdot\left(1+\frac{R_2}{R_1}\right)$$

Hence: $G_{id}=(V_{out}/V_{in})=1+(R_2/R_1)$

Again, the inverting terminal acts as a **virtual ground**, meaning that the OpAmp varies its output to make the inverting terminal equipotential to the positive one (eliminating the voltage error ε).

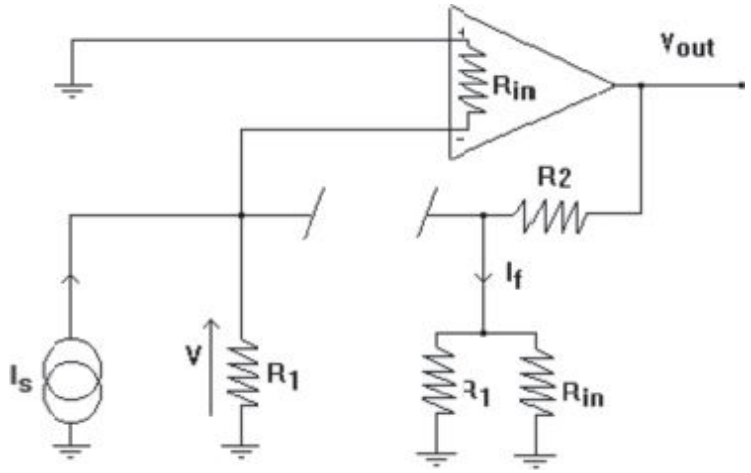


Fig. 1.22: Calculation of the loop gain.

LOOP GAIN

Let us insert a current probe generator I_s and reconstruct the impedance seen by the circuit before cutting. The circuit is shown in Fig. 1.22. We have $V = I_s \cdot (R_1 - R_{in})$

then: $V_{out} = -G_o \cdot I_s \cdot (R_1 // R_{in})$ $\quad I_f = -\dfrac{G_o \cdot I_s \cdot (R_1 // R_{in})}{R_2 + (R_1 // R_{in})}.$

Therefore: $\quad G_{loop} = \dfrac{I_f}{I_s} = -\dfrac{G_o \cdot (R_1 // R_{in})}{R_2 + (R_1 // R_{in})}.$

REAL TRANSFER

$$G_r = \frac{G_{id}}{1 - \frac{1}{G_{loop}}} = \frac{1 + \frac{R_2}{R_1}}{1 + \frac{R_2 + (R_1 // R_{in})}{G_0 \cdot (R_1 // R_{in})}}$$

1.5.3 Variation of Resistances of the Stage

The feedback of the OpAmp changes the values of the impedance seen from the outside (i.e. looking toward the inputs of the stage and toward the output). Let us see how.

INPUT RESISTANCE

Consider the circuit in Fig. 1.23, left. Let us calculate the equivalent resistance seen from the point A by injecting a voltage signal V_s. Knowing that $V_{out}=G_0 \cdot R_{in} \cdot I_s$, we have:

$$(I+I_s)\cdot R_1 + R_2 \cdot I = G_0 \cdot R_{in} \cdot I_s \qquad I(R_1+R_2) = (G_0 \cdot R_{in} - R_1)\cdot I_s \qquad I = \frac{(G_0 \cdot Rin - R1)\cdot Is}{R1+R2}$$

$$V_s - R_{in}\cdot I_s - (I+I_s)\cdot R_1 = 0 \qquad V_s = R_{in}\cdot I_s + \frac{R_1\cdot(G_o \cdot R_{in} - R_1)}{R_1+R_2}\cdot I_s + R_1 \cdot I_s$$

Since the equivalent resistance is given by the ratio V_s/I_s, we have:

$$R_{eq} = \frac{V_s}{I_s} = R_{in} + \frac{R_1}{R_1+R_2}(G_0 \cdot R_{in} + R_2) = R_{in} + \frac{R_1 \cdot R_2}{R_1+R_2} + \frac{G_0 \cdot R_1 \cdot R_{in}}{R_1+R_2} =$$

$$= R_{in} + (R_1 // R_2) + G_o \cdot (R_1 // R_2)\cdot \frac{R_{in}}{R_2}$$

If the system were ideal, G_{loop} would be ∞ (because $G_0=\infty$). In this case, $I_s=0$ would imply $R_{in}=\infty$. Conversely, for $G_0=0$, i.e. system without feedback, it would be:

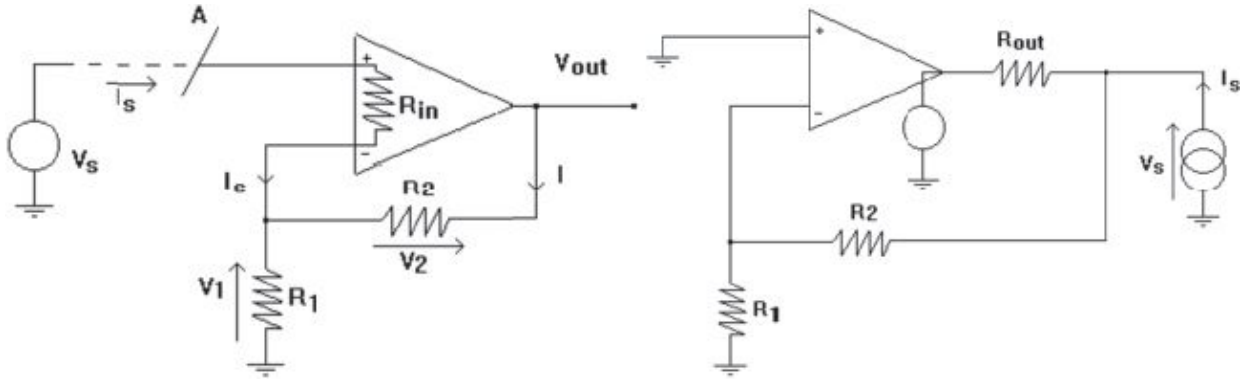


Fig. 1.23: Circuits for calculation of input (left) and output (right) resistances.

$$R_{eq} = \frac{V_s}{I_s} = R_{in} + (R_1 // R_2) = R_0$$

where R_0 is the resistance in the absence of feedback. Let us rewrite R_{eq}, highlighting G_{loop}:

Eq. 1.6

$$R_{eq} = \left[R_{in} + (R_1 // R_2)\right] \cdot \left[1 + \frac{G_0 \cdot (R_1 // R_2)}{R_{in} \cdot (R_1 // R_2)} \cdot \frac{R_{in}}{R_2}\right] = R_0 \cdot \left[1 - G_{loop}\right]$$

Therefore, with changing loop gain, the input equivalent resistance varies from Ro (the value with no feedback, i.e. $G_{loop}=0$) to $+\infty$ (for $G_{loop}=\infty$, the ideal case).

OUTPUT RESISTANCE

Assuming $R_{in}=\infty$, insert now a current probe generator I_s at the output (Fig. 1.23, right) and evaluate the voltage drop V_s across it: $R_{eq(out)}=V_s/I_s$. Ideally, we have $G_{loop}=\infty$, and $R_{eq(out)}=0$. Nonetheless, if $G_{loop}=0$ (i.e. $G_o=0$), we have $R_{eq(out)}=R_{0(out)}=R_{out}//(R_1+R_2)$ (resistance in the absence of feedback).

The full expression of the equivalent output resistance is the following:

Eq. 1.7

$$R_{eq(out)}=\frac{R_{o(out)}//(R_1+R_2)}{1-G_{loop}}$$

As can be seen, the resistance is reduced by the loop gain.

1.5.4 Variation in the Frequency Response of the Stage

Let us consider an operational amplifier whose gain Bode plot (of the forward block) is shown in Fig. 1.24, on the left. This is a compensated OpAmp with a pole at the frequency of 10Hz. The transfer function of the forward block of the operational amplifier is therefore of the type:

Eq. 1.8

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

where τ is the time constant of the pole, and G_0 is the dc gain. By closing the operational amplifier in a negative feedback with a block β, the new transfer function is obtained to be:

Eq. 1.9

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

The new dc gain will be: $\frac{G_0}{1+\beta\cdot G_0}$

And the new pole will no longer be at the angular frequency $1/\tau$ but at $\omega_p = (1/\tau)\cdot(1+\beta G_0)$.

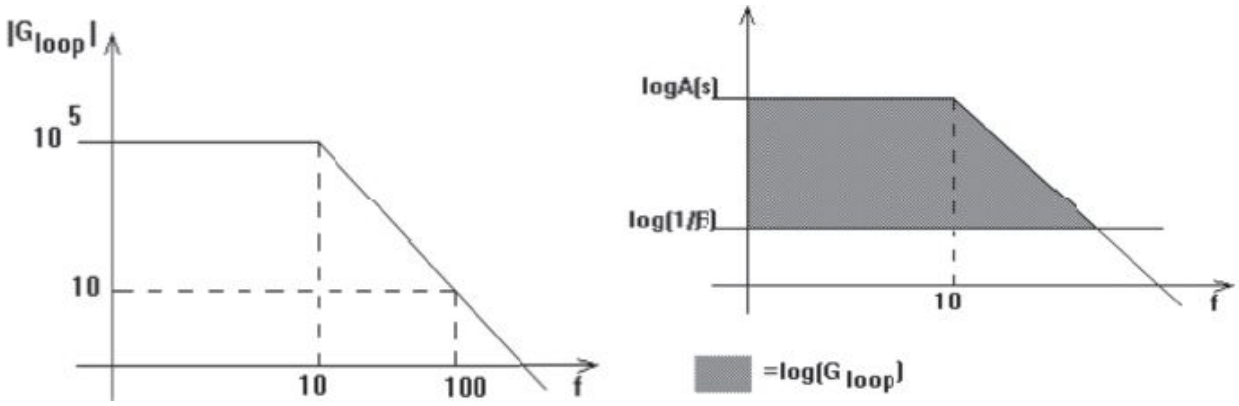


Fig. 1.24: Trend of the gain of the OpAmp (left). By superimposing to this the inverse of the gain of the feedback block, i.e. 1/β, we derive, from the distance between these two, the trend of G_{loop} at different frequencies (shaded area on the right).

By closing the negative feedback, we can note that the bandwidth goes from a value of 10Hz to a value of $10Hz(1+\beta\cdot G_0)$ and that the gain decreases by the same amount $(1+\beta\cdot G_0)$.

Considering the system in inverting configuration of Fig. 1.19, size the resistors so that the gain is 10 and verify its stability.

$$G_{id} = -(R2/R1) = -(10k\Omega/1k\Omega) = -10 = 1/\beta$$

The ideal gain does not faithfully represent the frequency response of the operational amplifier since the gain of the system varies by increasing the frequency.

In general:

$$G_r = \frac{A}{1+A\beta} \neq G_{id} = \frac{1}{\beta}$$

And:

$$G_r = \frac{G_{id}}{1-\frac{1}{G_{loop}}}$$

To know the real transfer function, we have to calculate the value of G_{loop}:

$$G_{loop} = \frac{V_f}{V_s} = -\frac{G_0}{1+s\tau}\cdot\frac{R_1}{R_1+R_2}$$

The Bode plot of the open-loop transfer function is shown in Fig. 1.24. Knowing that $G_o = 10^6$, we get $G_0 \cdot \frac{R_1}{R_1 + R_2} \cong 10^5$

Substituting the equation of the loop gain in the expression of the real gain, we have:

$$G_r = \frac{G_{id}}{1 - \frac{1}{G_{loop}}} = \frac{\frac{-R_2}{R_1}}{1 + \frac{1}{\frac{G_0}{1+s\tau} \cdot \frac{R_1}{R_1 + R_2}}}$$

To evaluate the stability of a system, the Bode plot of its loop gain (Fig. 1.24) must be analyzed. We can say that the system is stable if the crossing of the 0dB axis occurs with a slope of -20dB/dec. An index of stability is the phase margin, defined as: $M\phi = 180° - |\phi c|$, where ϕc is the critical phase, i.e. the phase of the loop gain at the angular frequency where the 0dB axis is crossed. To attain a degree of acceptable stability, the phase margin must be greater than or equal to 45°. In our case, the system Bode plot crosses the 0dB axis with a slope of -20dB/dec and will certainly be stable because: $\phi c = -90°$. The term –90° is introduced by the pole; then, the phase margin will be: $M\phi = 180° - |-90°| = 90^0$. The system, therefore, has a good value of phase margin and stability.

The analysis of the phase margin to access the stability of a system can be based on the simplified Bode criterion. A necessary and sufficient condition for the system to be stable is that there is no frequency such that when $|G_{loop}| = 1$, the phase is more negative then –180° because this situation would certainly cause the divergence of the output signal.

In fact, recalling $G_v(s) = A(s)/(1 - G_{loop})$, it is easy to see that each of the above conditions would make the denominator equal to zero, thereby resulting in the divergence of G(s). By introducing a second pole at a frequency below the crossing, the system will be unstable since it can get a phase shift of -180° with a gain greater than 1. Boundary condition for the stability is when the second pole is introduced at the crossing frequency because in that case, we have $M\phi = 45^0$. We have:

$$G_r = \frac{G_{id}}{1-\frac{1}{G_{loop}}} = \frac{\frac{-R_2}{R_1}}{1+\frac{1}{\frac{G_0}{1+s\tau}\cdot\frac{R_1}{R_1+R_2}}} = \frac{\frac{-G_0\cdot R_2}{R_1+R_2}}{\frac{G_0\cdot R_1}{R_1+R_2}+(1+s\tau)} = \frac{\frac{-G_0\cdot R_2}{R_1+R_2}}{\left(1+\frac{G_0\cdot R_1}{R_1+R_2}\right)\cdot\left(1+\frac{s\tau}{1+\frac{G_0\cdot R_1}{R_1+R_2}}\right)}$$

The pole has a time constant equal to: $\tau^* = \frac{\tau}{\left(1-G_{loop}(0)\right)} = \frac{\tau}{1+\frac{G_0\cdot R_1}{R_1+R_2}}$ where $G_{loop}(0)$ is the dc loop gain.

Generally, the frequency is given by $f=1/2\pi\tau$, then the frequency of the pole with time constant τ^* is $f^*=f\cdot(1- G_{loop}(0))$. We can note that with increasing $G_{loop}(0)$, the bandwidth of the system increases and that the real transfer decreases. In fact:

Eq. 1.10

$$G_r = \frac{A(s)}{\left(1-G_{loop}(s)\right)}$$

It is possible to obtain the forward gain using the real gain:

$$G_r = \frac{G_{id}}{1-\frac{1}{G_{loop}}} = \frac{\frac{-R_2}{R_1}}{1+\frac{(R_1+R_2)\cdot(1+s\tau)}{G_0R_1}} = \frac{\frac{-R_2}{R_1}\cdot\frac{G_0\cdot R_1}{(R_1+R_2)\cdot(1+s\tau)}}{\frac{G_0\cdot R_1}{(R_1+R_2)\cdot(1+s\tau)}\cdot\left(1+\frac{(R_1+R_2)\cdot(1+s\tau)}{G_0\cdot R_1}\right)} =$$

$$= \frac{\frac{-G_0\cdot R_2}{(R_1+R_2)\cdot(1+s\tau)}}{1+\frac{G_0\cdot R_1}{(R_1+R_2)\cdot(1+s\tau)}}$$

So we have: $A = \frac{-G_0\cdot R_2}{(R_1+R_2)\cdot(1+s\tau)}$. Rewriting the real gain in the formula $G_r = \frac{A(s)}{1+\beta\cdot A(s)}$, we observe that if $\beta A(s)>>1$, $G_r(s)$ tends to $1/\beta$. And if $\beta A(s) <<1$, $G_r(s)$ tends to $A(s)$. In Fig. 1.24 are shown the Bode plot of the forward gain (A) and the Bode plot of the ideal gain (Gid). Then the following equation holds: $\log(A(s))-\log(1/\beta)=\log(A\cdot\beta)=\log(G_{loop})$. Therefore, the

shaded area in Fig. 1.24, on the right, represents the Bode plot of the loop gain. Recalling Eq. 1.10, we observe that the Bode plot of the real gain will follow the Bode plot of the ideal gain until the loop gain $|G_{loop}|=A\cdot\beta$ will be much greater than 1. After that, it will follow the plot of A(s) for $|G_{loop}|<<1$.

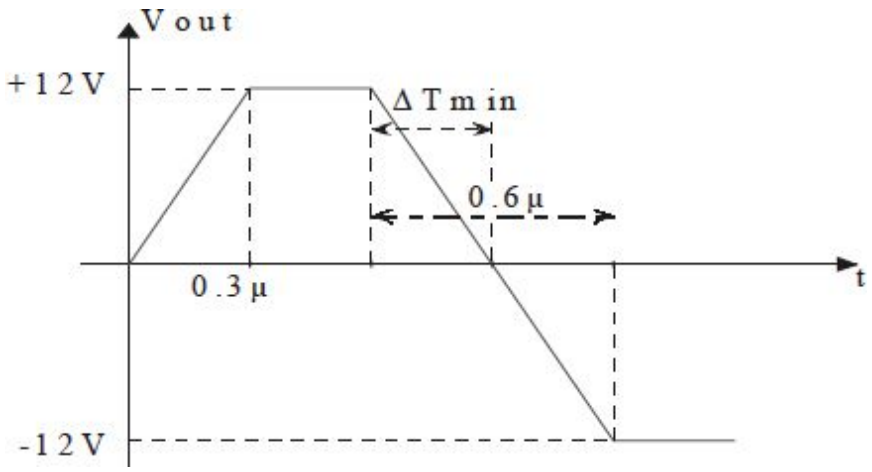


Fig. 1.25: Response of the circuit to a square wave.

1.6. Response to "Large Signal"

The switching speed of the output is limited by the internal Slew-Rate (SR) of the OpAmp, and this rate can cause high distortion. For example, Fig. 1.25 shows the triangular wave output in response to a square wave applied at the input, which is caused by an SR=40V/µs (then the output sweeps 24V in 0.6µs and not in an infinitesimal time).

The **Slew-Rate** of the OpAmp is defined as:

$$SR = \left(\frac{dVout}{dt}\right)_{max}$$

Typical values of the SR are 0.5÷50V/µs in voltage-mode OpAmps.

From this parameter, it is possible to obtain the maximum frequency of the square wave at the OpAmp input, which allows a full dynamic (even if distorted) response. This is obtained by:

$$f\max_{slew-rate} = \frac{SR}{2 \cdot \Delta Vo_{\max, peak-peak}}$$

An increase of the SR is achieved at the expense of higher consumption of the OpAmp.

Another important parameter related to the SR is the **Full Power Bandwidth**. This represents the maximum frequency of a sinusoidal input that causes a sinusoidal output with the maximum amplitude. It is clear that this is a more restrictive condition than the previous one. There are two different expressions to derive FPBW: in the case of resistive load:

$$FPBW = \frac{SR}{2\pi \cdot Vout_{\max}}$$

In other cases, it may also be the maximum output current of the OpAmp that limits the maximum slope, insomuch as it has to charge a capacitive load. In this case, the maximum slope is obtained using the usual equation of charging capacitor $I_{out\ max}=C_{out} \cdot dV_{out}/dt$.

1.6.1 Origin of the Slew-Rate

In Voltage Mode OpAmps, the differential input stage converts a voltage difference (V_{diff}) into a current signal. In the case of maximum imbalance, the maximum current available is the current of the "tail" of the stage shown in Fig. 1.26 with I_1. This represents the maximum current that charges the compensating capacitance Cc placed across the second stage (Miller compensation). The internal Slew Rate arises from this situation. We, therefore, have $SR_{theoretic}=I_1/C_C$.

Fig. 1.26: Buffer with OpAmp (left) and plots of the response of a buffer made with transistors (right).

Let us study the worst case, which, from the point of view of the SR, is the use of the OpAmp in buffer configuration. To this end, let us look at the following figures, in particular to Fig. 1.26, which shows the plots

(schematically) of experimental measurements, where we note a further deterioration with respect to the $SR_{theoretic}$.

Let us consider a pnp input stage (in which the "tail" stray capacitance Cs and the compensation capacitance Cc are also indicated) connected in voltage buffer configuration (Fig. 1.27). Let us then look at the rising edge of the input signal. The sudden rise of Vin brings Q2 to work as an emitter follower. The output voltage Vop(t) drives the follower. The current I_1-Is, flowing through Q2 and mirroring into Q4, gradually charges Cc. The output voltage Vop(t) is nothing but a replica of the voltage across Cc.

If Cs were not there, all the current I_1 would flow in Cc, and the charging of the capacitance would be linear with a voltage slope given just by the expression of the $SR_{theoretic}$:

$$\left.\frac{dVop(t)}{dt}\right|_{max,theoretic} = \frac{I_1}{Cc} = SR_{theoretic\,(withoutCs)}$$

Instead, it happens that the presence of Cs requires a new contribution of current such that:

$$\left.\frac{dVop(t)}{dt}\right|_{max} = \frac{I_1 - I_s}{C_c} \cong \frac{I_s}{C_s}$$

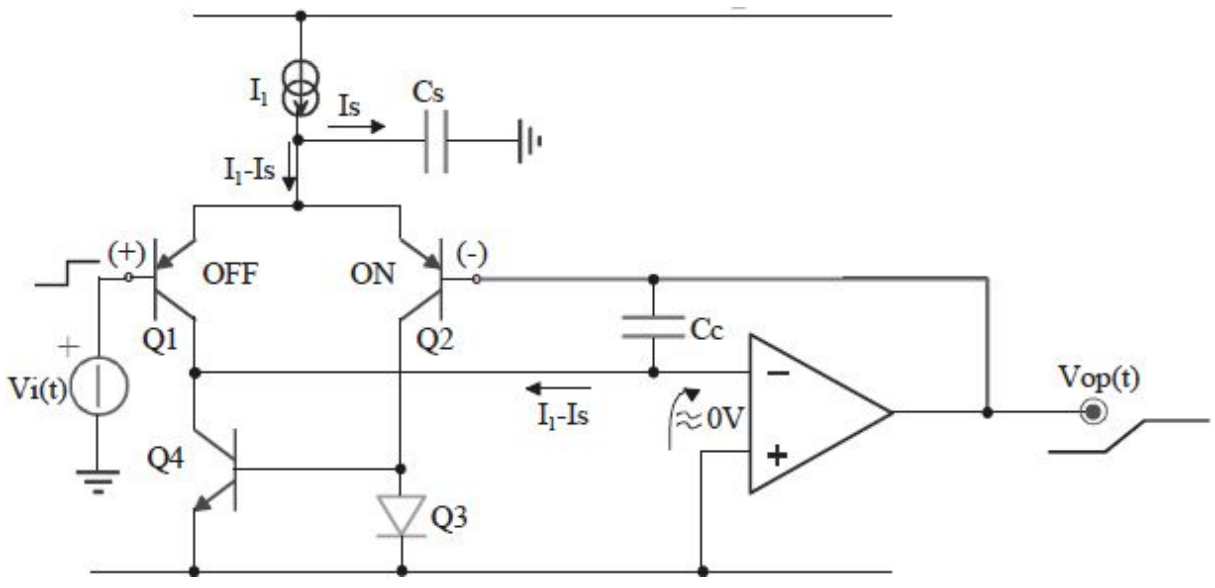


Fig. 1.27: Pnp input stage connected in voltage buffer configuration.

The last equality, in the expression above, is due to the role of the buffer that has Q2. Solving the equation, we obtain:

$$Is \cong \frac{I_1}{1+\frac{C_c}{C_s}}$$

which clearly shows that the increase in the output voltage follows a degraded slope.

$$SR_{degraded} \cong \frac{I_s}{C_s} = \frac{I_1}{C_s + C_c}$$

The reduction factor is equal to (1+Cs/Cc) and worsens as Cs becomes more comparable to Cc. In the case of an npn input stage, the behavior is reversed (Fig. 1.26).

From what just said, it is easy to understand that this effect is even worse if the OpAmp is designed for high-frequency applications (hence with small C_c). An example of this problem is found in practice. For example, the LM6118, which has C_c=5pF, I_1=500µA, and a $SR_{theoretic}$=100V/µs, due to the presence of a C_s=2pF, exhibits a $SR_{degraded}$=70V/µs which corresponds to a degradation of 30%. It is important to note that this degradation is only present in the case of non-inverting buffer, where the input generator directly drives an input of the OpAmp (on the left in Fig. 1.28 and Fig. 1.29). Instead, in the case of inverting buffer (A_v=-1), this effect is not present, as shown on the right of Fig. 1.28 and Fig. 1.29.

Let us now look at the falling edge. We refer to Fig. 1.30. The fast edge is applied between the base and the emitter of Q1 and gives rise to a current equal to:

$$I_s(t) = C_s \cdot \frac{v_{ip}}{dt}$$

The current pulse causes a charge step over C_c and, therefore, at the output:

Instead, the tail current I_1 (constant) determines the linear charge of C_c according to the $SR_{theoretic} = I_1/Cc$. This explains the waveforms visible at the output.

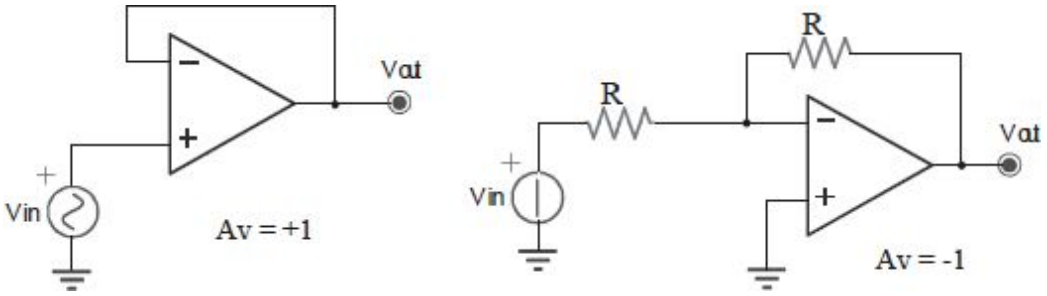


Fig. 1.28: Non-inverting buffer and inverting buffer.

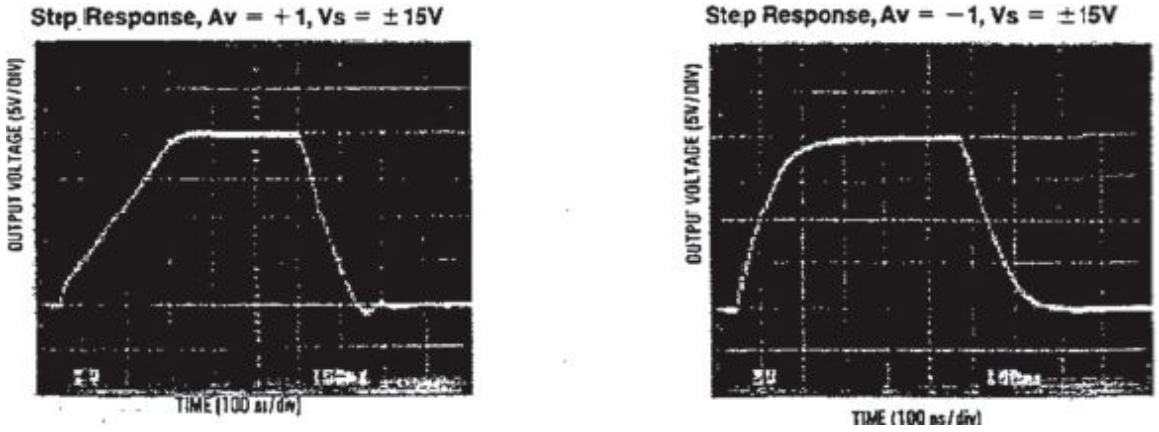


Fig. 1.29: Slew rate degraded and not degraded response for the LM6118.

Fig. 1.30: Circuit for the study of the response to the falling edge.

1.7. Response to "Small Signal"

1.7.1 Settling Time

This parameter represents the time that the output takes to go from 10% to 90% of the excursion in the case of a "small signal" (i.e. the signal is to keep the OpAmp internal components in their linear range). We observe, instead,

that the SR is a concept of a “large signal”. Fig. 1.31 shows the step response of a non-inverting buffer, which highlights the different patterns of response curves, depending on the amount of the input signal. For “small” intensity, there is a “linear” response with a single pole, i.e. with exponential charge with time constant τ (concept of “small signal”) related to the pole of the circuit. In this case, being buffer connected, the OpAmp has a closed-loop pole coincident with its GBWP. Instead, for “large” intensity of the signal, the response will be “non-linear”, which is limited by the SR. The threshold, the value ΔV below which a signal is “small” and above which is “large”, is obtained when the maximum slope of the exponential charge (equal to that at the origin with a value of dV/dtmax=Δ/τ) reaches the slew rate limit (i.e. dV/dtmax=SR), i.e. ΔV=SR·τ.

These values can be related to the parameters of the OpAmp, remembering that:

$$SR = \frac{I_1}{Cc} \qquad GBWP = \frac{gm_1}{2\pi \cdot Cc}$$

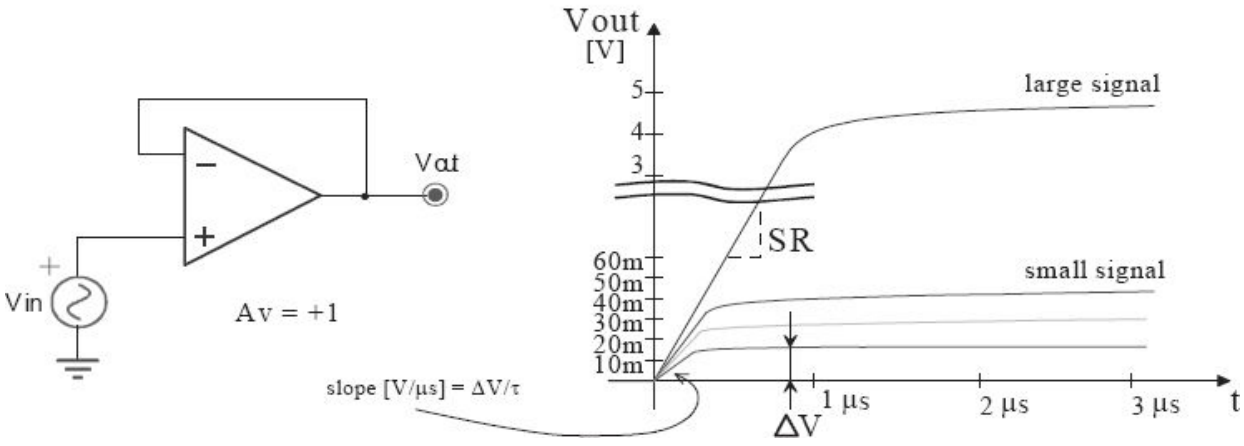


Fig. 1.31: Step response of an OpAmp in non-inverting buffer configuration.

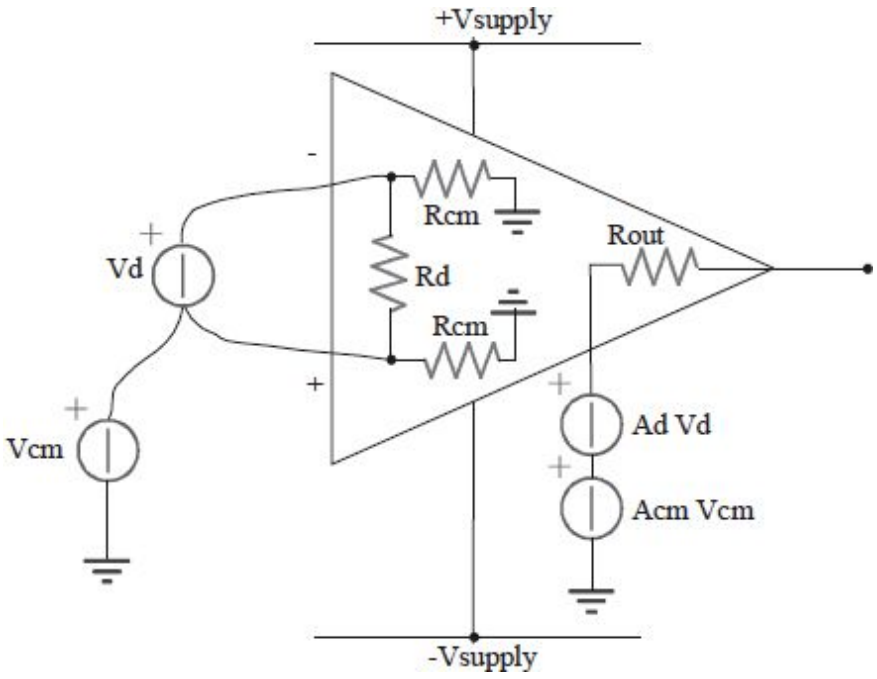


Fig. 1.32: Summary schematic of the electrical characteristics of the operational amplifier.

1.8. Electrical Characteristics

Operational Amplifiers (OpAmp) are integrated circuits that can achieve a very high voltage gain that is of the order of hundreds of dB. Precisely, because of this characteristic, OpAmps are well suited to the realization of electronic feedback circuits. In fact, if used as a forward block of a system, OpAmps allow you to achieve a high loop gain (by choosing a feedback block that does not attenuate too much).

Referring to the basic scheme shown in Fig. 1.32, which shows that you can apply two different types of signal to an OpAmp:

1- a “differential” signal, i.e. a voltage signal Vd applied between the positive terminal and the input terminal;

2- a “common mode” signal, i.e. a voltage signal Vcm simultaneously applied between the positive terminal and the ground and between the negative terminal and the ground.

It is obvious that the system has, at its input, different impedance, depending on the type of the signal applied. Typically, OpAmps have **common mode resistance** Rcm much higher than MW and **differential resistance** Rd that varies in a range of values between 100kW and 10MW

for devices made with standard technologies, but that achieves values of the order of 10^{12}W in more advanced technologies, such as BiFet and BiMos.

On the other hand, the **output impedance** Rout of an OpAmp is much smaller than its input impedance. Rout, in fact, takes values between 50W and 4kW. As will be seen later, the presence of high Rout can cause problems related to the stability of the feedback systems that we will design.

Let us look at how the two types of signal applied at the input are amplified.

Ideally, the **amplification of the differential signal** is infinite. Actually, the products on the market have gain values Ad=Vout/Vd ranging from 10,000 (80dB) to 200,000 (110dB). Because of this high gain, problems of accuracy arise since there is a large margin of tolerance for this parameter. It may thus happen that two copies from the same manufacturer have very different differential gain values. It is up to the ability of the circuit designer to make a project that works well whatever the value of Ad is in the range of tolerance of the device.

Conversely, the **amplification of the common mode signal** should ideally be equal to zero. Unfortunately, this is not possible, but it is possible to obtain values of Acm=Vout/Vcm still very low, between 1 and 10. The higher the difference between the values of Ad and Acm is, the better the OpAmp is. This figure of merit is expressed by the CMRR (**Common Mode Rejection Ratio**), normally expressed in dB, and defined as $20\text{Log}_{10}(\text{Ad/Acm})$. Typical values of this parameter are of the order of 90dB.

Another important parameter, which is often provided in datasheets, is the PSRR (**Power Supply Rejection Ratio**), defined as the ratio $\text{Vos}_{\text{supply}}/\text{DV}_{\text{supply}}$, where $\text{DV}_{\text{supply}}$ is the power fluctuation that causes a spurious variation in the output of the OpAmp $\text{Vout}_{\text{supply}}$, and $\text{Vos}_{\text{supply}}$ is the equivalent signal that should be provided to the input of the ideal OpAmp to cause the same fluctuation $\text{Vout}_{\text{supply}}$ of the output. Typical values are around 100dB.

In the documentation accompanying the product, there are also some indications of the output dynamics. The **Common Mode Input Voltage Range** expresses the voltage range within which we can change the inputs, ensuring the correct functioning of the OpAmp. If this is “rail-to-rail”, it will

mean that the inputs can be brought up to the highest extremes represented by the supply voltages of the OpAmp. The term *Output Voltage Swing* refers to the range within which the output can vary.

In Fig. 1.33 are shown the three **ranges of operating temperature** for electronic components and therefore also for the OpAmps.

Let us now examine the static errors of the OpAmp, i.e. non-idealities present in bias conditions (with no input signal). Fig. 1.34 helps us to

Fig. 1.33: Operating temperature ranges.

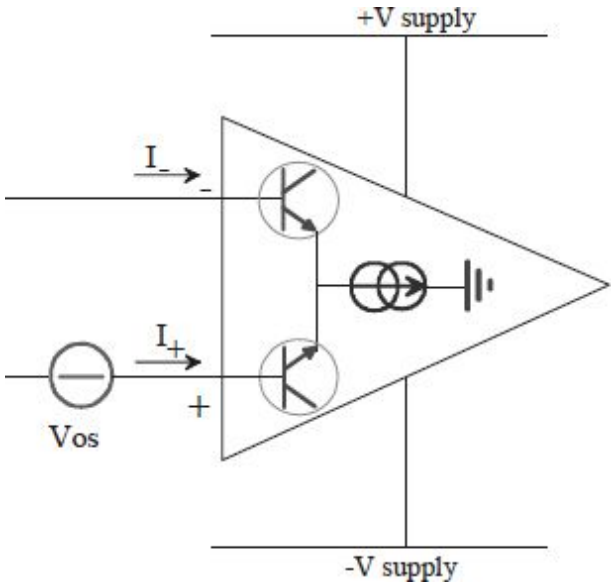


Fig. 1.34: Static non-idealities.

Fig. 1.35: Bias currents in the non-inverting stage with no resistance of the source generator (left) or high (greater than R1‖R2) (right).

see what happens in these conditions. Since the input impedance is not infinite, it happens that, once biased, the OpAmp absorbs (or supplies) current from the input terminals. Manufacturers quote on datasheets two parameters associated to this occurrence:

- **Bias current**: $I_B=(I_-+I_+)/2$ represents the average current at the two terminals. The order of the magnitude of this parameter depends on the technology used and ranges from few pA (MOS technology) to mA (bipolar technology);
- **Bias current offset**: $I_{OS}=I_--I_+$ takes values around 10% of I_B.

The OpAmp also presents an **offset voltage,** which is also shown in Fig. 1.34 through a voltage generator at the input. This expresses the fact that, even when no kind of an input signal is applied, the output is not exactly at the expected potential. It is as if at the input of the ideal OpAmp, a signal generator Vos were applied. Typical values for this offset are Vos≈±0.1mV÷±10mV. From this point of view, the OpAmps in bipolar technology are better than those in MOS technology.

The last important aspect of static non-idealities is represented by the **drifts**, both the bias current and the voltage offset. In fact, with the variation in the operating temperature, the electronic devices present in the operational amplifier change their characteristics (e.g., with the same current flow, the voltage drops across the pn junctions decrease with increasing temperature). This is reflected at the external terminals with a drift of the values for the parameters mentioned above. In particular, we observe a drift of I_B of the order of nA/°C and a drift of the offset voltage of the order of μV/°C, where the temperature of the silicon chip (and not that of the environment or of the package) is measured in Celsius degrees.

1.8.1 Compensation of the Bias Current

Let us now see how we can reduce the effects introduced by the bias currents through circuit solutions. We will refer to the non-inverting configuration of Fig. 1.35. In both cases, the introduction of the resistance R3, which is appropriately sized, cancels the effect of the bias currents. In the absence of R3, even without any applied signal, the output would have a voltage not equal to zero. In the circuit shown on the left of Fig. 1.35, the value would be equal to the product $I_{B-}\cdot R2$ and be positive or negative, depending on whether the bias current is absorbed or supplied by the OpAmp.

In the circuit on the left of Fig. 1.35, resistance R3 has been introduced at the non-inverting terminal, so the following relation is verified:

If the two bias currents are equal, we will choose the resistance to be equal to R3=(R1∥R2).

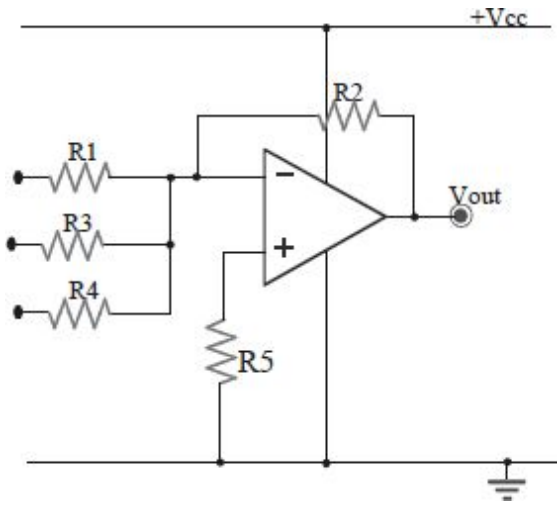


Fig. 1.36: Summing amplifier.

Fig. 1.37: Inverting stage properly AC decoupled (left) and non-inverting stage AC decoupled in the wrong way because it does not allow the flowing of I_B (right).

In the circuit on the right of Fig. 1.35, because of the high value of the series resistance of the signal generator, resistance R3 has been introduced in series with the inverting terminal of the OpAmp, so the following relation is satisfied:

If the bias currents are equal, to eradicate the error, we must choose R3=Rs-(R1∥R2). It is important to note that the resistance R3 was added to the circuitry so that the gain of the stage would not change.

To compensate the effects of the bias currents in the case of the summing stage of Fig. 1.36, we will have to size the resistance R5 as R5=R1//R2//R3//R4. Instead, we must not forget the presence of an offset of

the two bias currents, which remains unpredictable and cannot be compensated in advance.

Let us now make a very important practical observation related to the presence of these bias currents. Consider Fig. 1.37, where the two configurations are shown, inverting and non-inverting, with decoupling capacitors.

The solution on the right, which is apparently correct, is not viable just because of the presence of bias currents. In fact, the existence of a path to the ground to discharge the current coming from the non-inverting terminal is not guaranteed. The OpAmp, hence, cannot work and would ultimately saturate its output toward power supplies, V_{CC} or GND. Nevertheless, the circuit on the left of Fig. 1.37 can work properly because the bias current of the inverting terminal is guaranteed by the output of the OpAmp itself, through R2.

Fig. 1.38: Compensation of the Vos in the 741 (left) and in the LM301 (right).

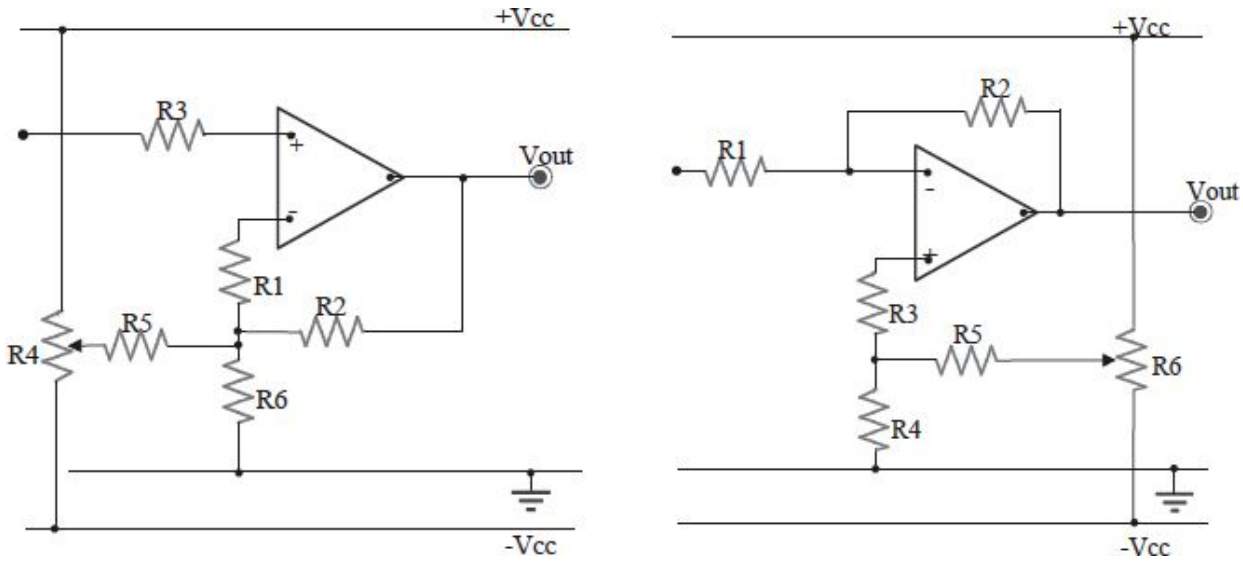


Fig. 1.39: Compensation of the Vos in an amplifier in non-inverting (left) and inverting configuration (right) with no offset-null pins.

1.8.2 Compensation of the Voltage Offset Vos

Now, let us study some methods for compensating the offset voltage. OpAmps are generally equipped with two pins of offset-null (Fig. 1.38) that can be used to connect an external resistor or a trimmer. If properly chosen, these pins allow compensating the initial offset voltage properly. This external resistance changes the OpAmp's internal voltage levels, thus positioning the output to the desired potential.

In the case in which the OpAmp is not equipped with the offset-null pins, we can proceed with the compensation by directly acting on the input terminals, through suitable resistive networks. In Fig. 1.39 are shown two examples of offset compensation. On the left, it can be seen, by means of a resistive partition and on the assumption that R4 (R5+R2∥R6) holds, that the voltage of the negative terminal is:

Note that we have neglected the effect of the bias current of the inverting terminal.

Fig. 1.40: Dual-polarity power supply.

Fig. 1.41: Single-polarity power supply (voltage values are typical of the LM324 OpAmp).

Fig. 1.42: Problem for bipolar signals in single-supply amplifiers.

The configuration on the right represents a situation similar to the previous one, but in the case of an inverting amplifier. Also in this case, appropriately sizing the resistive network, we are able to adjust the voltage of the positive terminal, thus bringing the output voltage to the desired value. We get when R6<R5+R4. Also in this case, the effect of the bias current of the non-inverting terminal has been neglected.

1.9. Power Supply

In this section, we will see how to power up the OpAmp and what the problems regarding the type of the selected power supply are. The OpAmp can be powered in two modes: Dual-Polarity and Single-Polarity. In Fig. 1.40, there is an example of ***dual-polarity*** power supply. In a dual-polarity OpAmp, the output can assume values of both positive and negative, but often it does not reach the values of the power supplies, neither the positive nor the negative one, as shown in Fig. 1.40 and Fig. 1.41. (because transistors inside the OpAmp always require the presence of a voltage at their ends for proper functioning in the active zone.) For example, using an OpAmp with a 5V ***single-polarity*** power supply (Fig. 1.41), its output voltage will take only positive values lower than 3.5V. In fact, above this value, the OpAmp comes out from its linear operation regime. Obviously, the voltage ranges depend on the internal circuitry of the OpAmp.

Now, let us see some typical problems related to the power supply of the circuit. Suppose that we use a single-supply OpAmp and that we want to amplify a bipolar signal, such as a sinusoidal signal with amplitude of 100mV. Please refer to the circuit in Fig. 1.42. The output in this configuration is biased at 0V; therefore, it does not allow amplifying the sinusoidal input signal because the output is forced to “cut off” the negative half-waves (below the lowest available power supply).

To get the amplified sine wave at the output, we will have to “raise” the output to the center of the dynamics and then “subtract” this dc value with a capacitive by-pass. Fig. 1.43 shows two examples of amplifiers with single-polarity power supplies, which can achieve the desired amplification. For the left amplifier:

Note that, in the circuit we are considering, the compensation of the bias currents is implemented since the impedance values in connection with both terminals of the OpAmp are the same (the inverting terminal sees 10kΩ, and the non-inverting terminal sees $20k\Omega \| 20k\Omega = 10k\Omega$). The capacitor C3 is used to filter the noise of the power supplies and of the two resistors while

C1 realizes the decoupling of the input, and C2 allows to obtain a bipolar signal at the output (positive and negative with respect to GND).

Fig. 1.43: Inverting (left) and non-inverting configurations (right) with single-polarity OpAmps.

Fig. 1.44: Single-Supply inverting amplifier with resistive load.

The amplifier on the right of Fig. 1.43, instead, has a gain equal to And this amplifier uses the capacitor C3 to prevent the output saturation (if C3 were not added, in the absence of a signal, the output would tend to 6V·11/1=66V). Unfortunately, this solution does not work for continuous signals (DC).

1.10. Output Currents

Now consider the amplifier of Fig. 1.44. We want to analyze what the maximum current that the OpAmp must be able to provide is in order to have proper operation of the circuit. There are two possible contributions of currents required from the OpAmp. The first contribution is due to the feedback branch: given a signal at the input of the stage, the system will respond by making a current flow through the feedback resistance that, in this case, is equal to the current forced by the signal generator. Let us call this current i_f. The second contribution is due to the current that must be supplied to the load. To determine what the maximum current required for the operational amplifier is, we have to consider the case of maximum output voltage excursion $Vout_{max}$. In this condition, at regime and at mid-frequencies, the required current is equal to:

To make the circuit properly operate at full dynamics, the OpAmp must be capable of delivering at least 6.6mA. Now, consider the OpAmp in inverting configuration and dual-polarity power supplied shown in Fig. 1.45. In

sinusoidal regime at frequency f and with an output signal with the maximum amplitude (+12V peak), the maximum current required to the OpAmp is equal to:

In this case, the load that the OpAmp has to drive is capacitive. On the assumption that the OpAmp LM324 is used, which has I_{outmax}=±20mA, the maximum rate in output change is determined by the following relation:

Fig. 1.45: Double-Supply inverting amplifier with capacitive load.

Fig. 1.46: Response of the circuit to a square wave.

The switching speed of the output is limited to this value provided that the internal SR of the OpAmp does not come into play before.

In practice, with an input square wave, to achieve the maximum dynamic output (say ±12V), we must wait for 0.3ms. The system has a bandwidth of:

This results, however, in a high distortion of the signal at the maximum frequency (at the output, a triangular wave appears in response to a square wave, as in Fig. 1.25).

1.11. EXERCISES

Ex. 1

The OpAmp has I_B=*50nA*, I_{OS}=*10nA*, V_{OS}=*5mV*, A_0=*100V/mV*, and *PSRR=60dB*; the weight sensor supplies v_S=*450µV/g*.

a) Introduce a resistor of *2.2kΩ*, or *47kΩ*, or *100kΩ* to reduce the effect of *IB* and calculate the errors (in "*grams*") due to I_{OS} and V_{OS} at the output.

b) Calculate the fluctuation of the power supply, which causes an error of *±1g*.

Ex. 2

The signal *vS*, sinusoidal with a frequency of *20Hz*, has a *100mV* peak. The OpAmp has *IB=50nA, VOS=5mV,* and *A0=100,000.*

a) If interference at *100Hz* with a peak of *0.6V* is superimposed on the 12 V, choose the PSRR to achieve a precision of *±1LSB.*

b) Calculate C*bypass* to have an error of *±1mV* due to the same interference in the power supply.

Ex. 3

Consider the following circuit.

a) Calculate the frequency response V_{out}/V_{in}, assuming that the OpAmp is ideal.

b) Draw the Bode plot (magnitude and phase) of the frequency response V_{out}/V_{in}. What function does this circuit perform? What is its bandwidth?

c) Assuming $V_{in}=A\cdot\sin(\omega_{in}t)$ with A=1V and ω_{in}=100krad/s, calculate the output $V_{out}(t)$.

d) If the operational amplifier has an SR of 0.5V/µs, is the voltage $V_{out}(t)$ still the one calculated at point c)? Why or why not?

e) Calculate the *maximum* output voltage due to the offset voltage and the bias currents of the OpAmp (V_{os}=5mV, $I_{b+}=I_{b-}$=100nA).

f) Consider the differential input resistance (R_{id}=500kΩ) of the operational amplifier. Calculate the resistance R^+ seen from the non-inverting terminal at dc (assume that the gain of the OpAmp is A_0=100dB).

Modify now the circuit assuming C1=∞.
Assume for the OpAmp $A(s)=A_0/(1+s/\omega_0)$ with A_0=100dB and $\omega 0$=10rad/s.

g) Calculate the loop gain of the new circuit and draw the Bode plots (magnitude and phase). Determine the phase margin approximately. Is the circuit stable enough? Give reasoning for your answer.

h) Calculate the bandwidth of the circuit.

Ex. 4

Consider the following circuit.

a) Calculate the ideal transfer function $v_{out}(s)/v_{in}(s)$, assuming that the OpAmp is ideal.
b) Draw the Bode plots (magnitude and phase, with values written in the abscissa and the ordinate) of the transfer function calculated in point a), in the frequency domain.
c) Given a sine wave with a peak of 100mV and a frequency of 100Hz as an input, calculate the amplitude, frequency, and phase shift of the output signal.
d) Determine the dc output voltage due to the offset voltage of the OpAmp.
e) Assume that the OpAmp is real. Calculate the loop gain of the circuit and draw the Bode plots (magnitude and phase). Approximately determine the phase margin and find out whether the circuit is sufficiently stable.
f) Connecting a load R_L=10kΩ to the output, determine the peak current that the OpAmp must deliver at full dynamics ($V_{out,\ peak}$=±5V) at 100kHz (at this frequency, consider C already “shorted”).

Ex. 5

In the circuit on the right, the *gate* control comes from a HCMOS gate.
a) Explain the operation with a *2Vpp* sinusoidal input, calculating the gain v_{out}/v_{in} and the input impedance, for both levels of the *gate*.
b) Determine the effect of a leakage current of the MOS of 10nA.

Ex. 6

Consider the following amplifier, whose input is applied a voltage step of 1V.

a) Calculate the step response of $V_{out}(t)$, assuming that the SR of the OpAmp is infinite and that GBWP=10 MHz. Draw the behavior of $V_{out}(t)$.

b) If the SR is equal to 10V/μs, is the last answer still true? Justify your answer and draw $V_{out}(t)$.

c) Given that R=1kΩ how much current does the OpAmp provide when the transient has ended?

d) If the OpAmp has I_{outmax}=3mA, what is the limit to the load resistance that can be connected to the output, given the input signal? Is it an upper or a lower limit?

a) The circuit is a non-inverting configuration with the ideal gain $G_{id}=1+9R/R=10$. The bandwidth of the non-inverting configuration is therefore $BW=GBWP/G_{id}=1MHz$ and can also be obtained from the Bode plot of V_{out}/V_{in}.

$G_{op}=A(s)$ is the open-loop gain while the real gain G_{real} (dashed) represents the gain V_{out}/V_{in}. From the Bode plot, we can see that G_{real} has a pole at f_P=1MHz that limits the bandwidth. The corresponding time constant is $\tau=1/(2\pi f_P)$=159ns. Since the OpAmp's SR is infinite, there is no restriction on the maximum slope of V_{out}. The step response of the non-inverting configuration is then an exponential with the time constant τ.

b) To assess whether an SR of 10V/µs modifies the step response obtained in a), we must calculate the maximum slope of the exponential drawn previously. The maximum slope is at the origin of the axis and is 10V/159ns=62.89V/µs. Since now the OpAmp limits the maximum slope of the output V_{out} to 10V/µs, the trend of V_{out} is different from that calculated in a). As long as the slope of the exponential does not go below 10V/µs, V_{out} is SR limited.

The moment when the SR ceases to limit the slope of V_{out} can be obtained by calculating the point where the slope of the exponential is exactly 10V/µs. From that point on, the slope of the exponential is lower than 10V/µs and, therefore, is no longer restricted by the SR:

$$V_{out}\ t) = 10V\left(1 - e^{-t/\tau}\right)$$

$$\frac{dV_{out}}{dt} = \frac{10V}{\tau} e^{-t^*/\tau} = 10\frac{V}{\mu s}$$

leading to t*=292ns, to which corresponds V_{out}*=10V/µs·t*=2.92V.

c) When the transient is over, V_{out}=10V, and then the current delivered by the OpAmp is the current flowing through the series of the resistors 9R+R=10kΩ. The supplied current is $I_{out}=V_{out}/(9R+R)$=1mA.

d) Since through the resistors (9R, R) flows a current of 1mA, through the additional load resistor can flow at most the remaining current deliverable by the OpAmp, i.e. 3mA-1mA=2mA. Therefore, the load we can connect to the output must have resistance greater than R_{LOAD}=10V/2mA=5kΩ.

Ex. 7

Given the circuit on the right:
a) find the relation *Vout(Va, Vb)*;
b) calculate the impedance seen by *Va*;
c) calculate the impedance seen by *Vb*;
b) calculate the impedance seen by *Vout*;

a) $V_{out} = 2 \cdot V_a - V_b$

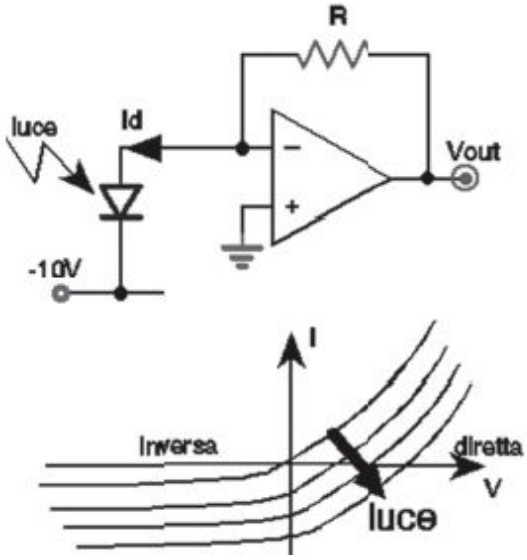
b) the impedance is infinite since there is never current absorption.

c) the impedance is just *100k*Ω that is in series with *Va*.

d) the impedance is zero because even if the OpAmp were not ideal, there would be the negative feedback to reduce the impedance seen.

Ex. 8

Given the circuit on the right, the photodiode provides a current *Id=0.2μA* every *1μW* of incident light.
a) Plot the I-V characteristics of the diode where it is working.
b) Size *R* such that, for a light of *500μW*, the output signal is *Vout=5V*.



a) the load line is at *-10V*, as shown.
b) for *500μW*, the photocurrent is *0.2·500=100μA*.
Therefore, we have to choose a resistance value equal to R*=5V/100μA=50k*Ω to have *5V* at the output.

Ex. 9

Consider the amplifier shown on the right.

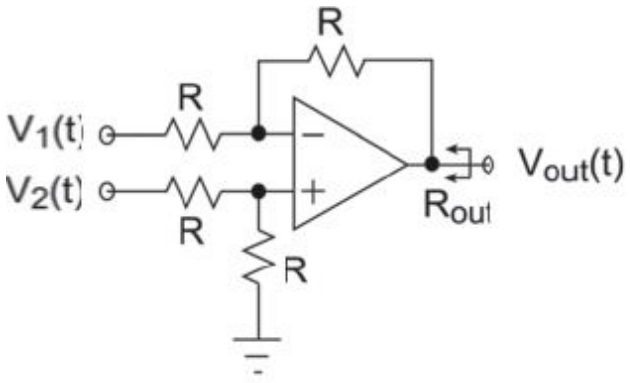
a) Assuming that the OpAmp is ideal, derive the expression of $\mathbf{v_{out}(t)}$ and draw it.

b) Assuming that the OpAmp has a GBWP of 1MHz, find the bandwidth of the amplifier for each input $\mathbf{v_1(t)}$ and $\mathbf{v_2(t)}$.

c) What should the value of the minimum SR of the OpAmp be such that the output is not distorted?

d) Assuming that the operational amplifier has an output resistance r_{out}=100Ω, what is the resistance $\mathbf{R_{out}}$ of the amplifier?

e) What is the average value of the common mode voltage at the input of the OpAmp? Justify your answer.



R=10kΩ

$V_1(t)=1V\cdot\sin(\omega t)+2V$ A_0=100dB

$V_2(t)=1V\cdot\sin(\omega t+\pi)+2V$ ω=314krad/sec

a) The configuration realizes a subtractor. Assuming that the OpAmp is ideal, using the principle of superposition, it is possible to obtain $V_{out}(t)$. Turning off $V_2(t)$, the transfer function from $V_1(t)$ to $V_{out}(t)$ is that of an inverting configuration: $V_{out}(t)|_1=-R/R\cdot V_1(t)=- V_1(t)$. Turning off $V_1(t)$, the transfer function from $V_2(t)$ to $V_{out}(t)$ is that of a non-inverting

configuration: $V_{out}(t)|_2=R/(R+R)\cdot(1+R/R)\cdot V_2(t)=V_2(t)$. Then $V_{out}(t)=V_{out}(t)|_1+V_{out}(t)|_2=V_2(t)-V_1(t)=1V\cdot\sin(\omega t+\pi)+2V-(1V\cdot\sin(\omega t)+2V)=-2V\cdot\sin(\omega t)$.

b) Let us separately analyze the two configurations, in one of which $V_1(t)$ is active (inverting configuration) and, in the other, $V_2(t)$ is active (non-inverting configuration).

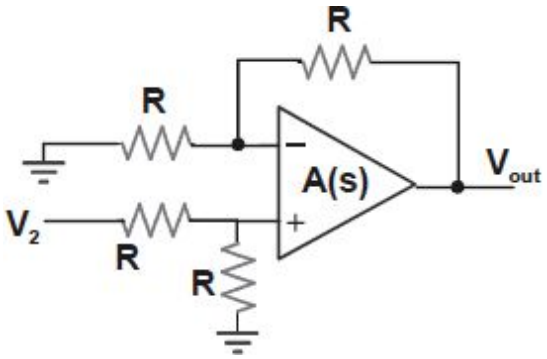
$V_1(t)$:

Ideal gain: $G_{id}=-R/R=-1$;

Loop gain: $G_{loop}=-R/(R+R)\cdot A(s)=-1/2A(s)$;

Open-loop gain: $G_{op}=-G_{loop}G_{id}=-1/2A(s)$.

On the graph, the point where G_{op} intersects G_{id} is the point at which $G_{loop}=1$. Beyond that point, G_{real} follows the trend of G_{op} and not that of G_{id}. So, at $1/2\cdot GBWP=500kHz$, there is a pole in the real gain and then the bandwidth is limited to 500kHz.



$V_2(t)$:

Ideal gain: $G_{id}=(R/(R+R))\cdot(1+R/R)=+1$;

Loop gain: $G_{loop}=-R/(R+R)A(s)=-1/2A(s)$;

Open-loop gain: $G_{op}=-G_{loop}G_{id}=1/2A(s)$.

On the graph, the point where G_{op} intersects G_{id} is the point at which $G_{loop}=1$. Beyond that point, G_{real} follows the trend of G_{op} and not that of G_{id}.

So, at 1/2·GBWP=500kHz, there is a pole in the real gain and then the bandwidth is limited to 500kHz.

c) To ensure that the output is not distorted, the speed with which the output could vary should be less than the maximum rate at which the output of the OpAmp can vary. This means that the maximum slope of the output signal must be at most equal to the SR of the OpAmp. Since $dV_{out}/dt|_{max}=|d(2V\cdot\sin(\omega t))/dt|_{max}=2V\cdot\omega=0.628V/\mu s$, the minimum SR is $SR_{min}=0.628V/\mu s$.

d) In a negative feedback circuit, the resistance seen looking from the outside toward the circuit may be increased or decreased by $(1-G_{loop})$.

The feedback stabilizes the voltage at the nodes of its loop in the sense that it tends to minimize voltage changes. Then, if you have to evaluate the resistance pertaining to a node, it tends to be very low (tends to zero for $G_{loop}\rightarrow\infty$), i.e. $R=R_{open}/(1-G_{loop})$, where R_{open} is the resistance seen when the loop is open.

In addition, the feedback stabilizes the current in the branches of the feedback loop in the sense that it tends to minimize current variations. Then, if we have to evaluate the resistance coming along a branch of the loop, it tends to be very high (tends to ∞ for $G_{loop}\rightarrow\infty$), i.e. $R=R_{open}\cdot(1-G_{loop})$, where R_{open} is the resistance seen when the loop is open.

In this case, if the feedback were ideal ($G_{loop}\rightarrow\infty$), the output resistance would tend to be very low ($R_{out}\rightarrow 0$). Then, the output node's voltage is stabilized and $R_{out}=R_{open}/(1-G_{loop})$, where $R_{open}=r_{out}\|(R+R)\approx 100\Omega$ and $G_{loop}=-A_0\cdot R/(R+R+r_{out})\approx -A_0/2=-5\cdot 10^4$. Therefore, $R_{out}\approx 2m\Omega$.

e) The common mode voltage at the input of the operational amplifier is the average of the voltages on the terminals "+" and "-" of the OpAmp: $V_{in,cm}=(V^+ + V^-)/2$. The input $V_1(t)$ has no effect on either the voltage V^+ or the voltage V^-. In fact, in assessing the effect of $V_1(t)$, we have to turn off $V_2(t)$, and then at the terminal "+", there is 0V. With ideal feedback, also at the terminal "-", there is 0V, and that node is a virtual ground, on which $V_1(t)$ has no effect. The input $V_2(t)$, instead, has an effect on both the "+" and "-" terminals. Its contribution on both terminals is $V_2(t)\cdot R/(R+R)=V_2(t)/2$.

Then the common mode voltage at the input of the OpAmp is $V_{in,\ cm}$= (V_2(t)/2+V_2(t)/2)/2=V_2(t)/2. The average value of the common mode voltage at the input of the amplifier is

Ex. 10

a) Determine V_{out} as a function of the inputs V_1 and V_2 in the case of ideal operational amplifier. Draw the magnitude Bode plot of the transfer function. What is the function of the circuit?

b) Calculate the gain static error assuming A_0=80 dB.

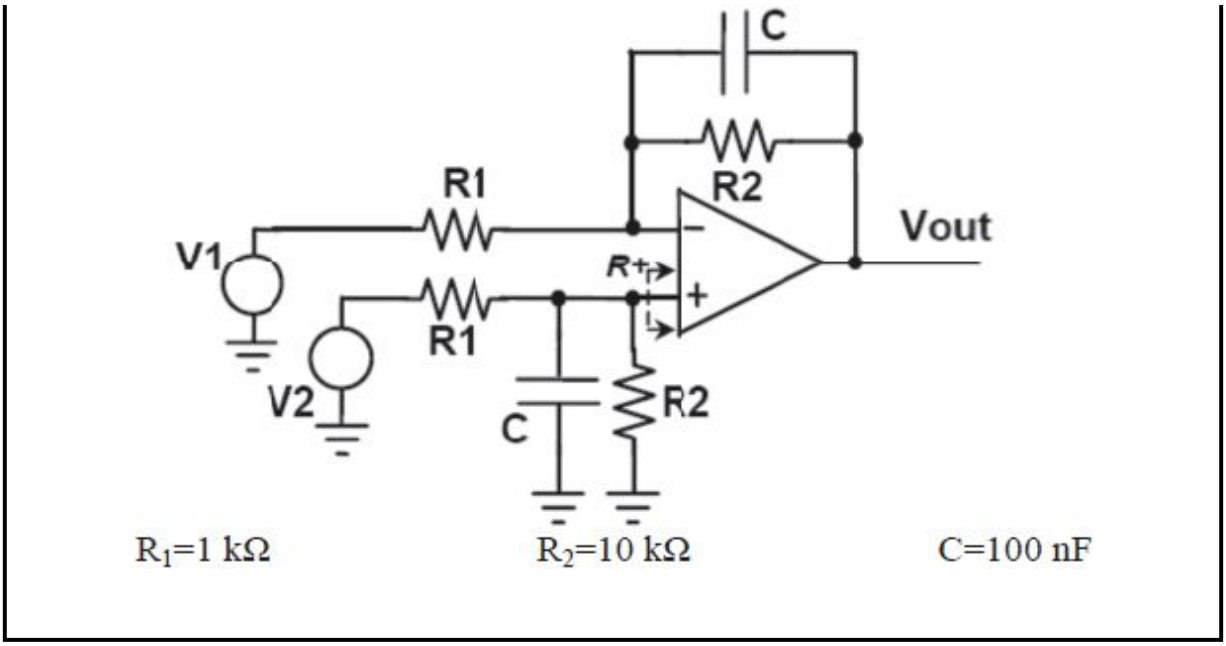
c) Knowing that V_{OS}=10mV, determine the error at the output due to the offset voltage.

d) Assuming that the OpAmp is characterized by differential input resistance R_{id}=100 kΩ and a dc gain A_0=80 dB, determine the expression and the value of the resistance R^+ seen from the positive terminal of the OpAmp in dc.

e) Determine the expression of the loop gain and draw the magnitude and phase Bode plots, assuming that the gain-bandwidth product of the operational amplifier is GBWP=1 MHz (A_0=80 dB and neglect R_{id}). Finally, determine the phase margin of the circuit.

f) Compute the response to a step voltage of 100 mV in V_2 (with V_1=0), assuming that the OpAmp is ideal and plot it.

g) What is the minimum value of the SR of the OpAmp, for which the response calculated in f) is still valid?



a) To determine V_{out} as a function of V1 and V2, we must apply the principle of superposition. Turning off V2 to evaluate the effect of V1, at the terminal "+", we have a passive network in which a current cannot flow (the ideal operational amplifier has an input resistance R^+ which tends to infinity and therefore no current can flow). Then, the voltage at the terminal "+" is 0V. The transfer from V1 to V_{out} is therefore through an inverting configuration with impedance $Z_1(s)$=R1 and $Z_2(s)$=R2∥1/sC=R2/(1+sC·R2):

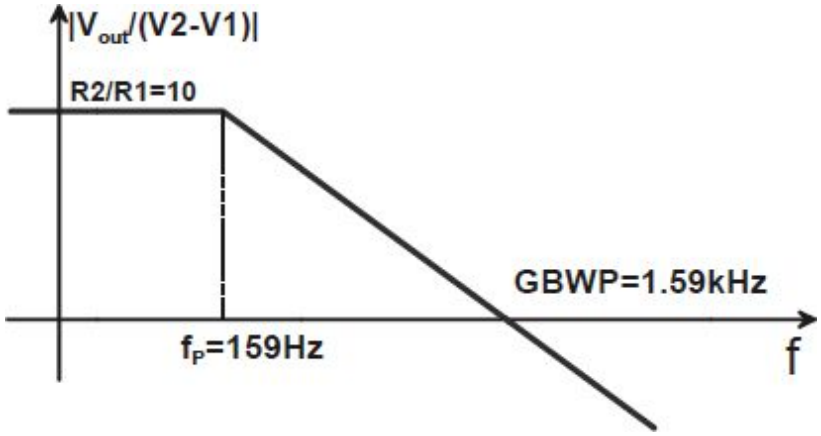
$$V_{out}\big|_{V1} = -\frac{R_2}{R_1}\frac{1}{1+s\cdot C\cdot R_2}$$

For V2, V1 is turned off. The voltage on the terminal "+" is determined by a voltage divider made with R1 and R2∥1/sC: V^+=V2·(R2∥1/sC)/(R2∥1/sC+R1). From V^+ toward the output V_{out}, the transfer is that of a non-inverting configuration with $Z_1(s)$=R1 and $Z_2(s)$=R2∥1/sC=R2/(1+sC·R2). Overall, the transfer function between V2 and V_{out} is:

$$V_{out}\big|_{V2} = \frac{R_2}{R_1+R_2}\frac{1}{1+sC\cdot R_1 \parallel R_2}\left[1+\frac{R_2}{R_1\left(1+sC\cdot R_2\right)}\right]V_2 =$$
$$= \frac{R_2}{R_1}\frac{1}{1+sC\cdot R_2}V_2$$

In conclusion, V_{out} as a function of V1 and V2 is given by the sum of $V_{out}|_{V1}$ and $V_{out}|_{V2}$:

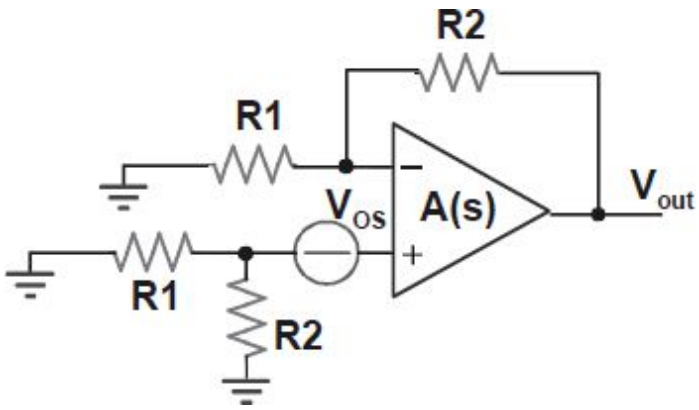
There is a pole at $f_P=1/(2\pi C\cdot R2)=159Hz$ and the dc gain is R2/R1=10. The magnitude Bode plot is the following. It is a low-pass difference amplifier.



b) The static error ε of a feedback configuration with OpAmps is the percentage difference between the real gain and the ideal one $\varepsilon=(G_{ideal}-G_{real})/G_{ideal}$. Recalling that $G_{real}=G_{ideal}/(1-1/G_{loop})$, we find the general expression $\varepsilon=1/(1-G_{loop}(0))$. In this case, $G_{loop}(0)$ is found by turning off V1 and V2 and by replacing C with an open circuit. In the feedback branch, there is now only R2 and the dc loop gain is $G_{loop}(0)=-R1/(R1+R2)\cdot A_0\approx -10^3$. Then, we find that the static error is $\varepsilon=1/(1-G_{loop}(0))=1/(1+1000)\approx 1/1000=0.1\%$.

c) The effect of the offset voltage must be evaluated in dc (C = open circuit and V1 and V2 off). The equivalent generator of the offset voltage

V_{OS} can be placed, for example, on the terminal "+", as shown in the figure below.

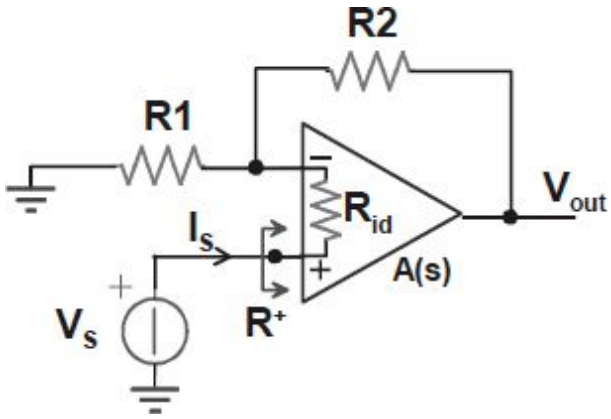


In the parallel R1∥R2, no current can flow (in the terminal "+" of an ideal OpAmp, no current can flow) and then the transfer function from V_{OS} to V_{out} is that of a non-inverting configuration. In conclusion, $|V_{out}|=(1+R2/R1)V_{OS}=110mV$. Note that V_{OS} can be either positive or negative, but the sign is not known a priori.

d) In assessing the resistance seen looking towards the terminal "+" of the OpAmp, it may be useful to consider first the asymptotic case, in which the gain A_0 of the OpAmp tends to infinity. In this case, the difference between the two terminals "+" and "-" tends to zero and then the current I_S (which develops after the application of the probe generator V_S) tends to zero. Then, we find that in the ideal case, the resistance seen $R^+=V_S/I_S$ tends to infinity.

It is understandable then that the effect of feedback is to increase the resistance R^+ seen, and therefore the expression for the resistance seen at the terminal "+" of the feedback system is $R^+=R^+_{open\ loop}\cdot(1-G_{loop}(0))$, where $R^+_{open\ loop}$ is the resistance seen when the loop is open, and $G_{loop}(0)$ is the dc loop gain. Note that the loop gain $G_{loop}(0)$ now includes also R_{id} (that is placed in parallel with R1) and, therefore, is:

The resistance $R^+_{open\ loop}$ is given by the series of R_{id} and $R1\|R2$: $R^+_{open\ loop}=R_{id}+R1\|R2$. In conclusion, $R=R^+_{open\ loop}\cdot(1-G_{loop}(0)) \approx 100k\Omega\cdot(1+10^3) \approx 100M\Omega$.



e) With $Z_2(s)=R2\|1/sC=R2/(1+sC\cdot R2)$ and $A(s)=A_0/(1+s\tau_0)$, the loop gain is:

$$G_{loop}(s) = -A(s)\frac{R_1}{R_1+Z_2(s)} = -A_0\frac{R_1}{R_1+R_2}\frac{1}{1+s\tau_0}\frac{1+sC\cdot R_2}{1+sC\cdot R_1\|R_2}.$$

Knowing the gain-bandwidth product, we find that the pole of the OpAmp is at $f_{P1} = 1/(2\pi\cdot\tau_0) = GBWP/A_0 = 100Hz$. In the Laplace transform of the loop gain, there is also a zero at $f_Z=1/(2\pi\cdot C\cdot R2)=159Hz$ and a pole at $f_{P2}=1/(2\pi\cdot C\cdot R1\|R2)=1.59kHz$. The dc loop gain is, instead, $G_{loop}(0)=-A_0(R1/(R1+R2))\approx 10^3$.

It follows that the asymptotic magnitude Bode plot of G_{loop} is as follows:

The value of G_{loop} in the range between f_Z and f_{P1} was found by using the relation between the frequency and the amplitude of a curve with a slope of -20dB/decade: $100Hz\cdot 10^3=159Hz\cdot x \rightarrow x=628$.

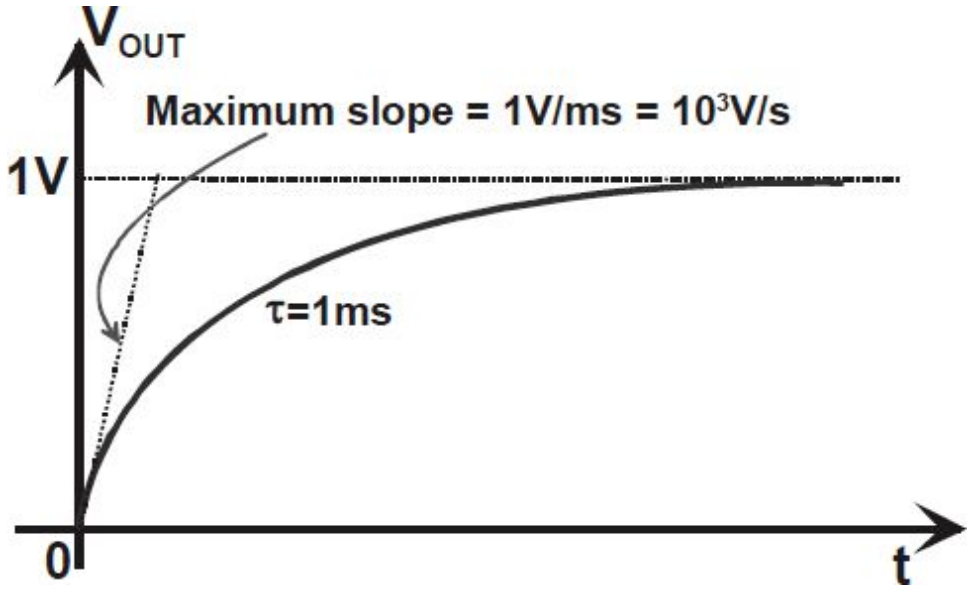
The phase Bode plot has a phase shift of -180° due to the negative sign of G_{loop}. In addition, the poles give a contribution of -90° while the zero gives a contribution of +90°. The pole f_{P1} and the zero f_Z are less than a decade

apart, and then their contribution is not complete. This means that between the first pole and the zero, the phase begins to drop by -90°, but soon rises because of the close zero.

The remaining difference between the phase of $G_{loop}(f_T)$ and -360° is the phase margin (f_T is the frequency where the loop gain is equal to one). In this case, the phase margin is $\Delta\psi\approx(-270°-(-360°))=90°$ and ensures that the feedback circuit is stable.

f) In the Laplace transform domain, the response to a step V2=0.1V/s is given by the product between V2(s)=(0.1V)/s and the transfer function

which has only one pole in $1/(2\pi\cdot C\cdot R2)=159Hz$. Then, the trend over time is that typical of the step response of a single-pole system: growing exponential with the time constant $\tau=C\cdot R2=1ms$ which tends to $0.1V\cdot 10=1V$.



g) The SR of the OpAmp does not change the behavior of V_{out} found in f) if it does not limit the maximum slope of the exponential. Then, the SR must be greater than the maximum slope that we have at the origin, i.e. SR > 10^3 V/s.

2

Circuits with OpAmps

"I don't want to lose you, this good thing
that I got 'cause if I do
I will surely,
surely lose a lot.
'Cause your love is better
than any love I know.
It's like thunder and lightning,
the way you love me is frightening.
You better knock, knock on wood, baby.

"Knock on Wood", Amii Steward

2.1. Linear Circuits with OpAmps

The circuits employing OpAmps can be sorted into two big categories: the nonlinear circuits (super diodes, Schmidt trigger comparators, oscillators, etc.) and the linear circuits. The latter (with a particular attention for negative feedback circuits) will be treated in the first part of this chapter. In these circuits, the output is brought back to the input through a feedback network. The practical effect of the negative feedback lies in maintaining the inverting input of an OpAmp at the same voltage as the non-inverting input pin. In this way, the voltage difference is null (except for a fraction of mV due to the non-infinite gain of the OpAmp).

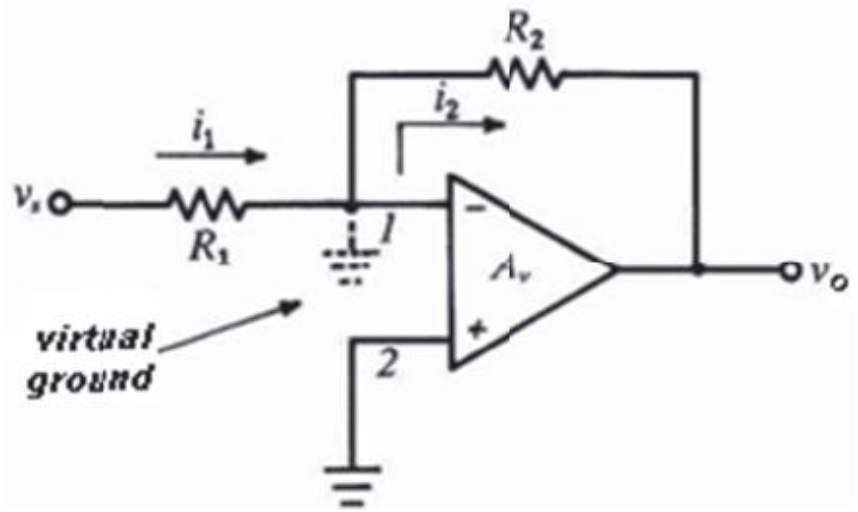


Fig. 2.1: Inverting amplifier stage.

2.1.1 Inverting stage

The inverting stage is shown in Fig. 2.1. The non-inverting input is connected to the ground, and the input signal is applied to the inverting input through the resistance R_1. Starting from the relationship $v_o=A_v\cdot(v_1-v_2)$, we obtain $A_v=v_o/(v_1-v_2)$. Assuming an ideal OpAmp (infinite A_V) and that the output signal v_o is bounded, $v_2=v_1$ must hold. In this case, $v_2=0$ and $v_1 = 0$, and it is possible to regard the inverting terminal as a "**virtual ground"**. The term *virtual* indicates that, even if the inverting terminal has a zero voltage, the current through the resistance R_1 is not short-circuited to the ground through the inverting pin as that pin has infinite impedance and cannot sink any current. Therefore, applying the first Kirchhoff principle (Kirchhoff Current Law, K.C.L.) to the node 1, it is possible to write $i_1=i_2$, i.e. the OpAmp changes the voltage of its output in such a way as to "recall" the feedback current i_2. This current suffices to bring the inverting input voltage to the same value as the non-inverting input. It is possible to write: $i_1=v_s/R_1$ and $i_2=-v_u/R_2$. Thus, $$A_{vr}=\frac{v_o}{v_s}=-\frac{R_2}{R_1}$$

This equation can be generalized, referring to the circuit shown in Fig. 2.2, in which the resistors R_1 and R_2 are replaced with generic Z_1 and Z_2

impedance values (possibly as a function of the frequency). Generally, we have:

$$A_{vr} = -\frac{Z_2}{Z_1}$$

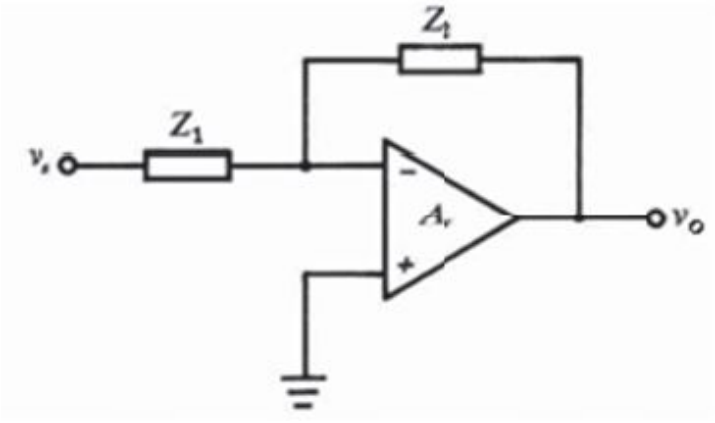


Fig. 2.2: Generic scheme of the inverting stage.

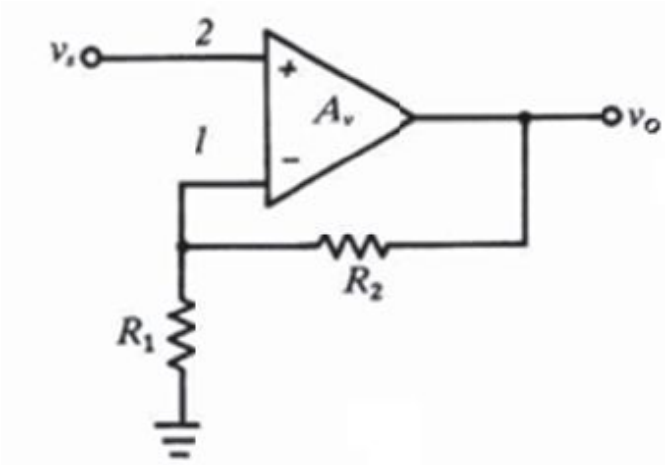


Fig. 2.3: Non inverting stage amplifier.

2.1.2 Non-inverting stage

The non-inverting stage is shown in Fig. 2.3. The inverting terminal is connected to the ground through the resistance R_1 while the input signal will be applied to the non-inverting pin. Assuming that the OpAmp is ideal, we have: $v_1 = v_2 = v_s$. Moreover, the resistors R_1 and R_2 form a voltage divider

because R_i is infinite, and the signal v_s across the resistance R_1 is: $v_s = v_o \cdot \frac{R_1}{R_1 + R_2}$. The gain is equal to: $A_{vr} = \frac{v_o}{v_s} = \frac{R_1 + R_2}{R_1} = 1 + \frac{R_2}{R_1}$

This expression can be generalized for the impedance values: $A_{vr} = 1 + \frac{Z_2}{Z_1}$

2.1.3 Inverting Adder

Consider the circuit shown in Fig. 2.4 and that the OpAmp is ideal. The K.C.L. applied to the node A gives $i = i_1 + i_2 + ... + i_n$. Regarding inverting terminal as a virtual ground, it is possible to write: $-\frac{v_o}{R} = \frac{v_1}{R_1} + \frac{v_2}{R_2} + ... + \frac{v_n}{R_n}$. From the preceding equation, it is possible to obtain: $v_o = -\left(\frac{R}{R_1}v_1 + \frac{R}{R_2}v_2 + ... + \frac{R}{R_n}v_n\right)$. If $R_1 = R_2 = ... = R_n$ If $R_1 = R_2 = ... = R_n$, we get:

$$v_o = -\frac{R}{R_1}(v_1 + v_2 + ... + v_n)$$

That shows how the circuit adds the voltages of the input terminals with a gain equal to R/R_1 (possibly an attenuation if the ratio is lower than one), and the output gives the inverted voltage (due to the sign "-"). If $R=R_1$, we will have:

$$v_o = -(v_1 + v_2 + ... + v_n)$$

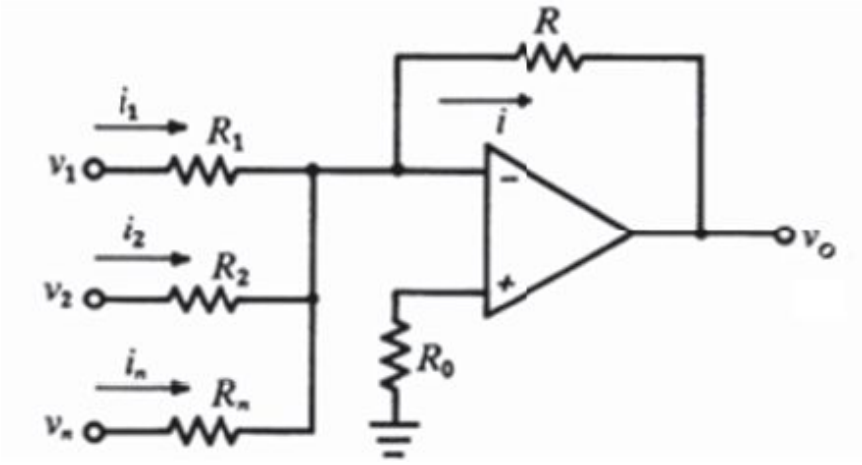


Fig. 2.4: Inverting adder.

2.1.4 Subtractor (difference amplifier)

Consider the circuit shown in Fig. 2.5. The output signal can be obtained by using the superposition principle. Imposing $v_2 = 0$, the circuit becomes that shown in Fig. 2.6, on the left. It is an inverting amplifier:

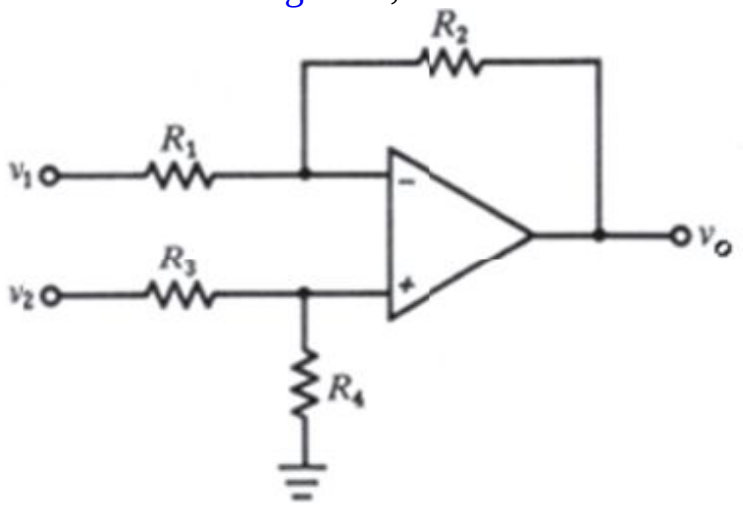


Fig. 2.5: Subtractor (or difference amplifier).

$$v_{o1} = -\frac{R_2}{R_1} v_1$$

In the same way, imposing $v_1 = 0$, the circuit becomes that shown in Fig. 2.6, on the right, which is a resistive divider built up from R_4 and R_3, followed by a non-inverting amplifier: $v_{o2} = \frac{R_4}{R_3 + R_4} \cdot \left(1 + \frac{R_2}{R_1}\right) \cdot v_2$

Due to the superposition principle, the output signal is equal to:

$$v_o = v_{o1} + v_{o2} = -\frac{R_2}{R_1} v_1 + \frac{R_4}{R_3 + R_4} \cdot \left(1 + \frac{R_2}{R_1}\right) \cdot v_2 = -\frac{R_2}{R_1} v_1 + \frac{R_4}{R_3} \frac{1 + \frac{R_2}{R_1}}{1 + \frac{R_4}{R_3}} \cdot v_2$$

Imposing the condition $\frac{R_2}{R_1} = \frac{R_4}{R_3}$, we will obtain $v_o = -\frac{R_2}{R_1} v_1 + \frac{R_4}{R_3} v_2 = -\frac{R_2}{R_1} (v_1 - v_2)$.

Assuming $R_1 = R_2 = R$, the output is $v_o = v_2 - v_1$.

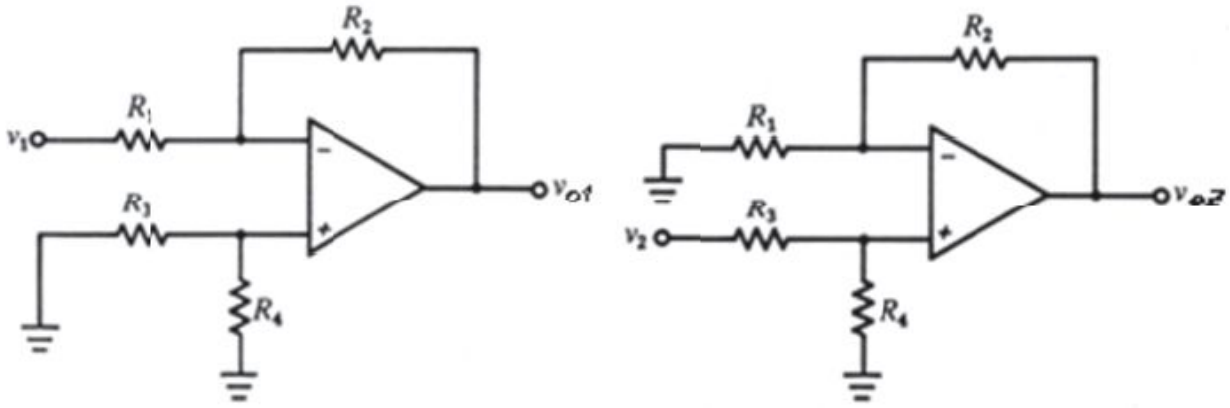


Fig. 2.6: Difference amplifier with $v_+ = 0$ (left) and with $v_1 = 0$ (right).

2.1.5 Instrumentation Amplifier

The Instrumentation Amplifier (INA) is a difference amplifier with the following specifications:

- Finite, accurate, and stable gain, typically between 1 and 1000.
- Extremely high input impedance, avoiding the attenuation of the input signal.
- Very low output impedance, maximizing the transfer to the load.
- Very high CMRR, amplifying only the input differential signal while completely neglecting the common-mode signal.

The difference amplifier can be used for obtaining an accurate and stable gain. Unfortunately, the second specification is not fully accurate, insomuch as the input impedance values, differential (Z_{id}) and common-mode (Z_{ic}), are not infinite. Moreover, the two input impedance values are not equal. Those imperfections can be eliminated by connecting two OpAmps in the INA configuration (shown in Fig. 2.7), where A_1 and A_2 form the input stage, and A_3 is the output stage.

We can analyze the whole circuit. V_{oi} and V_{o2} are the outputs of A_1 and A_2, respectively. The difference amplifier A_3 provides: $$V_o = \frac{R_2}{R_1} \cdot (V_{o2} - V_{o1})$$

Because $I_{R_3} = I_{R_G}$, starting from the Ohm's Law, we have: $V_{oi} - V_{o2} = (2R_3 + R_G) \cdot I_{R_G}$ and $I_{R_G} = \dfrac{V_1 - V_2}{R_G}$.

We obtain: $V_o = A \cdot (V_2 - V_1)$ and $A = A_I A_{II} = \left(1 + 2\dfrac{R_3}{R_G}\right) \cdot \left(\dfrac{R_2}{R_1}\right)$

where A_I and A_{II} are the first and the second stage gain, respectively. As can be seen, the gain A depends only on the resistors and can be made accurate and stable by using high quality and low tolerance (1%) resistors.

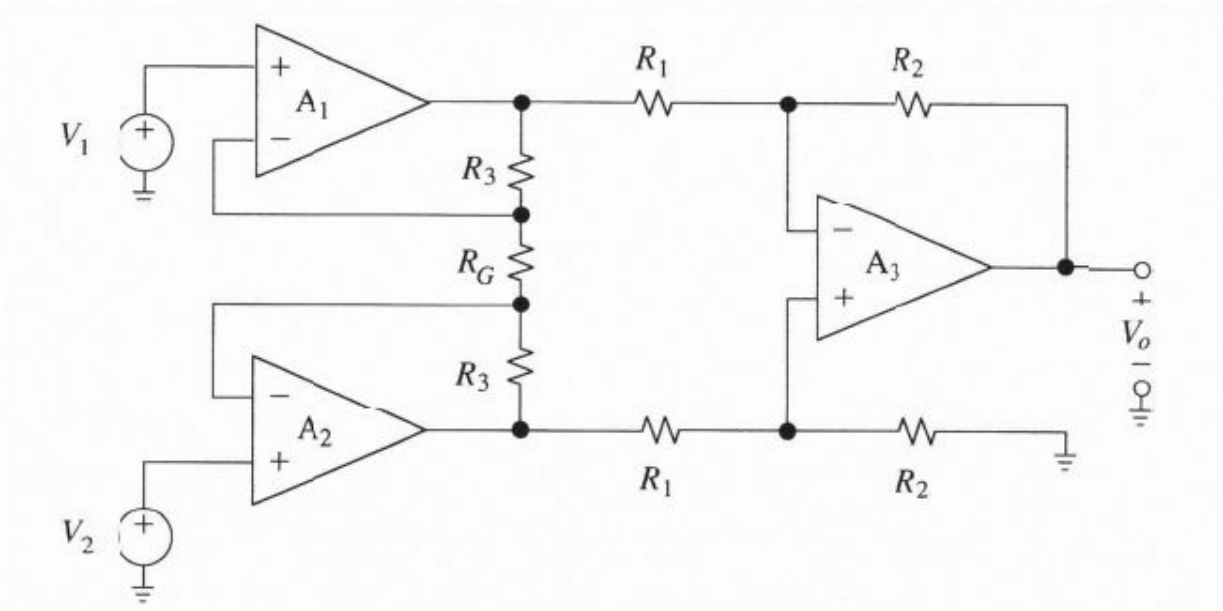


Fig. 2.7: Schematic of an Instrumentation Amplifier using three OpAmps.

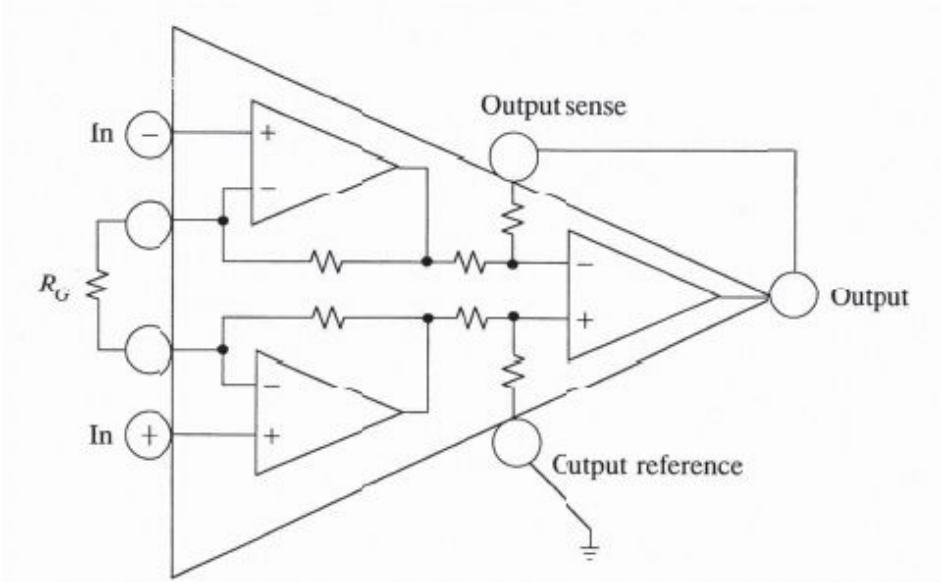


Fig. 2.8: Typical pin-out of a commercial INA.

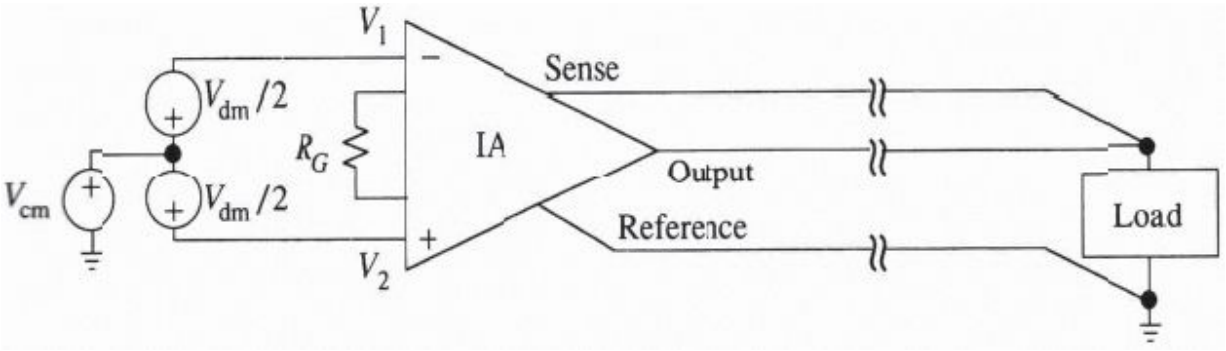


Fig. 2.9: Typical connection between an INA and a far load.

The INA built up from three OpAmps is commercially available in the form shown in Fig. 2.8. The most popular INAs are the INA101, INA102, INA104 (Burr-Brown), AD522 (Analog Devices), and LH0036 (National Semiconductors). Such devices contain all the components except the resistance R_G, which is externally connected by the designer to adjust the device gain. Unlike a normal OpAmp, the INA has some additional pins (*sense* and *reference*), as shown in Fig. 2.9. In this way, the undesired signal drop due to the resistance of the wires can be eliminated. In fact, the spurious resistance along the output connection will determine a voltage drop between the OpAmp output and the load. However, the signal supplied to the sense pin is read by the load, and the feedback of the OpAmp imposes the required

voltage (equal to the input signal multiplied by the gain of the OpAmp) between the sense and the reference pins.

2.2. Integrators and Derivators

In order to process signal time-transitions, OpAmp configurations require some reactive components in the feedback network, such as inductors and capacitors. Due to economical reasons and circuit-size constraints, usually only capacitors are used.

2.2.1 Ideal Integrator

The circuit is shown in Fig. 2.10, on the left. It can be thought of as an inverting amplifier with the feedback resistance R_2 replaced with the capacitance C.

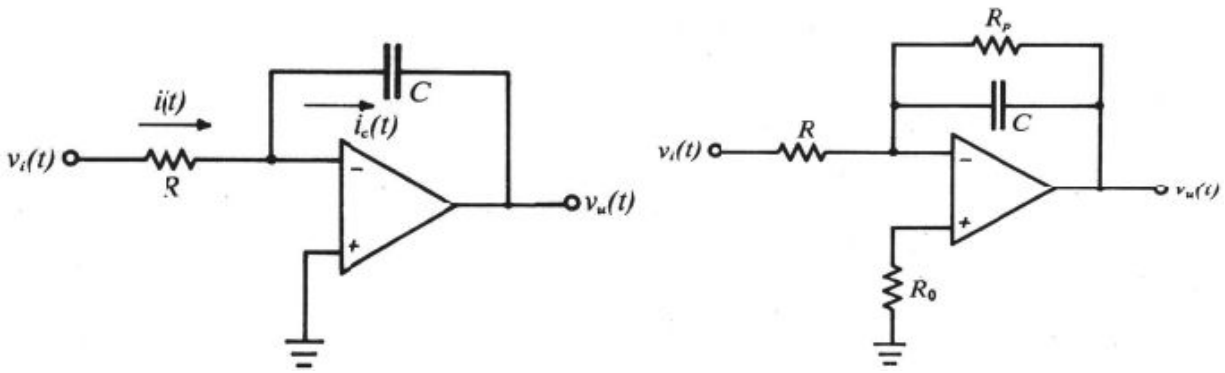


Fig. 2.10: Ideal (left) and real (right) inverting integrator.

Considering the OpAmp ideal, we can write $i(t) = i_c(t)$. Moreover, we have $v_i(t) = R \cdot i(t)$. Thus, the output signal is equal to:

$$v_o(t) = -\frac{1}{C}\int i(t)dt = -\frac{1}{RC}\int v_i(t)dt$$

Defining the time constant of the circuit as $\tau = RC$, we have

$$v_o(t) = -\frac{1}{\tau}\int v_i(t)dt.$$

Applying the Laplace Transform, we obtain:

$$A_v(s) = \frac{V_o(s)}{V_i(s)} = -\frac{1}{s\tau}$$

which highlights the temporal integration of the input signal, which is performed by the circuit. The integrator is used in waveform generators (triangular, saw tooth, etc.), active filters, A/D converters, and analog control systems.

2.2.2 Real integrator

The circuit shown in Fig. 2.10, on the left, has only academic purposes as it is sensitive to OpAmp offset and bias currents that cause a divergent output that will saturate the OpAmp. To avoid this problem, the resistance R_P is connected in parallel to the capacitance C (Fig. 2.10, on the right), thus limiting the gain at low frequencies.

The amplification, a function of the complex variable s, is:

$$A_v(s) = -\frac{Z_p(s)}{R}$$

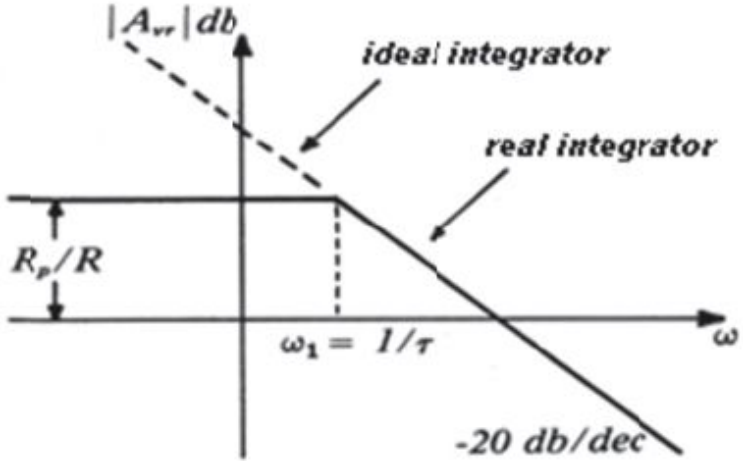


Fig. 2.11: Bode diagrams of transfer function $A_v(j\omega)$ for the ideal and real integrators.

where the impedance $Z_P(s)$ is given by the parallel branch built with the resistance R_P and C's reactance:

$$Z_p(s) = \frac{R_p \frac{1}{s \cdot C}}{R_p + \frac{1}{s \cdot C}} = \frac{R_p}{1 + s \cdot R_p \cdot C}$$

Starting from the last two equations, we can find:

$$A_v(s) = -\frac{R_p}{R} \cdot \frac{1}{1 + s \cdot R_p \cdot C} = \frac{A_{vo}}{1 + s\tau}$$

where A_{vo}= -R_P/R is the low frequency gain, and $\tau = R_p \cdot C$ is the time constant of the feedback branch. One can observe the presence of a pole s_1= -1/τ, and for this reason, the Bode diagram of the gain A_v (Fig. 2.11) will be characterized by a slope variation happening at an angular frequency ω_1=1/τ.

In fact, assuming s=jω, we have: $$A_v(j\omega) = \frac{-\frac{R_p}{R}}{1 + j\omega\tau}$$

Fig. 2.11 demonstrates that the circuit shown in Fig. 2.10, on the right, presents the behavior of an integrator (as if the resistance R_p were not there) only for frequencies higher than: $$f_1 = \frac{1}{2\pi \cdot R_p \cdot C}$$

Indeed, for frequencies f<f_1, the behavior of the circuit can be approximated with that of an inverting amplifier (as if the capacitance C were not there) with a gain: $$A_{v0} = -\frac{R_p}{R}$$

The resistance R_P is usually chosen to be equal to 10·R while the resistance R_0 is equal to the parallel resistance made from R_P and R so as to nullify the effects attributable to the bias currents of the OpAmp. The time constant $\tau = R_p \cdot C$ must be chosen starting from the frequency of the signal that has to be integrated.

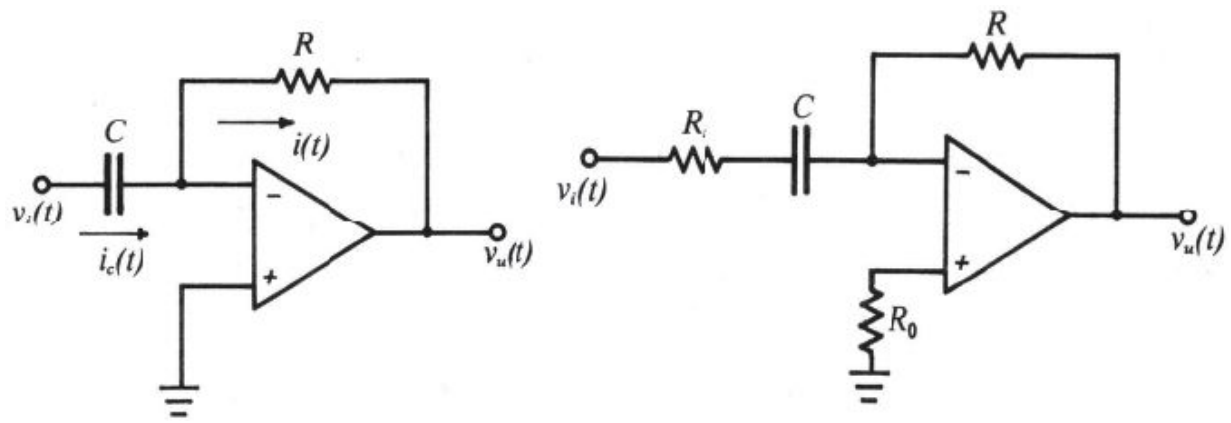


Fig. 2.12: Ideal (left) and real (right) inverting derivator.

2.2.3 Ideal derivator

The derivator is shown in Fig. 2.12, on the left. It can be thought of as an inverting amplifier with the input resistance R_1 replaced with the capacitance C.

Assuming that the OpAmp is ideal, one can write $i(t) = i_c(t)$. As:

$$i_c(t) = C \cdot \frac{dv_i(t)}{dt} \qquad \text{and} \qquad i(t) = -\frac{v_o(t)}{R},$$

the output signal is $v_o(t) = -\tau \cdot \frac{dv_i(t)}{dt}$ Assuming $\tau = R \cdot C$, we can write:

$$v_o(t) = -RC \cdot \frac{dv_i(t)}{dt}$$

Applying the Laplace Transform, we can reach the following solution:

$$A_v(s) = -s\tau$$

which demonstrates that the behavior of the stage is the derivative function.

2.2.4 Real derivator

The derivator shown in Fig. 2.12, on the left, which is theoretically valid, is excessively sensitive to high frequency noise (at high frequencies, the reactance of C will be very low). In fact, if we impose s=jω in the gain equation, we will have $A_v(j\omega) = -j\omega\tau$, i.e. a voltage gain that increases with

frequency. To avoid this drawback, as shown in Fig. 2.12, on the right, the resistance R_S is added in series with the capacitance C, which limits the high frequency gain.

Now, the amplification as a function of the complex variable s is given by:

$$A_{vr}(s) = -\frac{R}{Z_s(s)}$$

where Z_S (s) denotes the impedance of the series R_S and C, i.e.:

$$Z_s(s) = R_s + \frac{1}{SC} = R_s \cdot \left(1 + \frac{1}{s \cdot R_s \cdot C}\right)$$

Therefore, we can obtain:

$$A_{vr}(s) = \frac{A_{vr0} \cdot s\tau}{1 + s\tau}$$

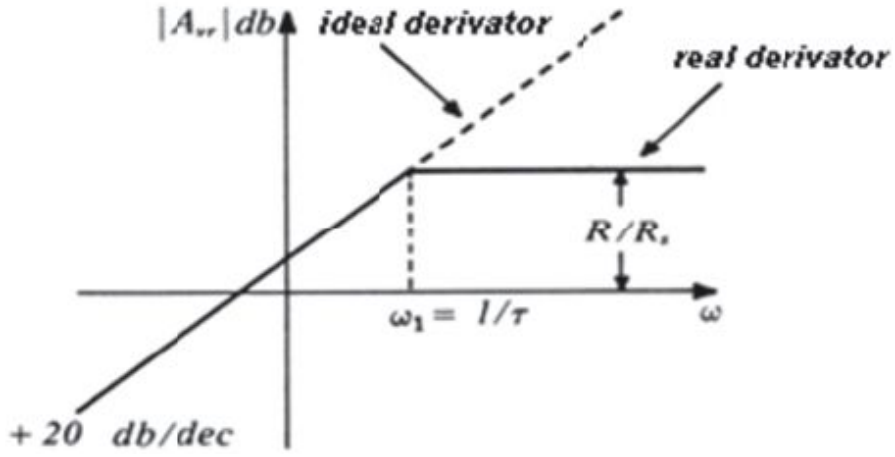


Fig. 2.13: Bode diagram of absolute value of the real derivator gain.

where A_{vro}= -R/R_s denotes the high frequency gain (when C is similar to a short circuit), and τ = R_s C denotes the time constant of the input branch. The transfer function shows the presence of a zero at null frequency and of a pole at a frequency equal to s_1= -1/τ; therefore, the Bode diagram will be the one shown in Fig. 2.13. Observe that the circuit shown in Fig. 2.12, on the right,

demonstrates the behavior of a derivator (as if the resistance R_S were not there) only for frequencies lower than: $$f_1 = \frac{1}{2\pi \cdot R_s \cdot C}$$
On the other hand, for frequencies $f>f_1$, the behavior of the circuit can be approximated with an inverting amplifier (as if the capacitance C were not there, which behaves as a short circuit) having a gain equal to: $$A_{vr0} = -\frac{R}{R_s}$$
The resistance R_S is normally sized as R/10 while the resistance R_0 is chosen to be equal to the parallel resistance of R_s and R. The time constant $\tau=R_sC$ is chosen according to the period or duration of the signal to derive.

2.3. Current and Voltage Converters

Current and voltage converters are circuits used to linearly convert a voltage signal into a current signal and vice-versa.

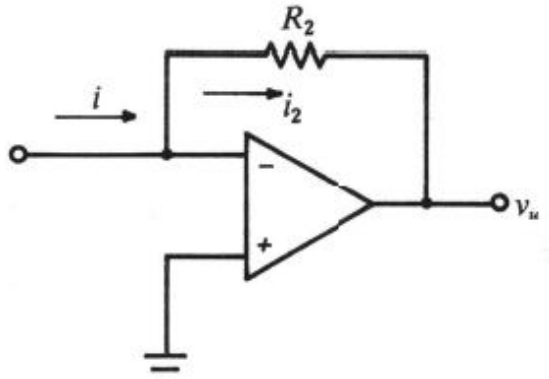


Fig. 2.14: Current-voltage converter.

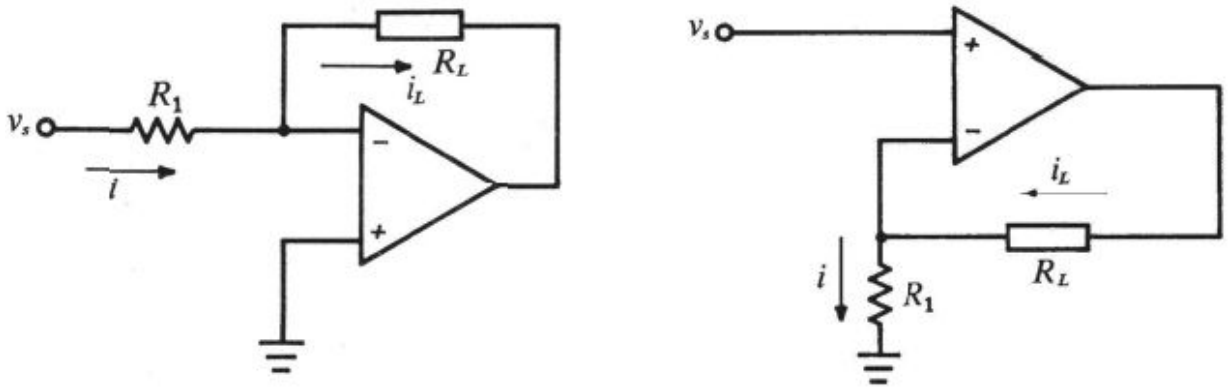


Fig. 2.15: Voltage-current converter with inverting (left) and non-inverting (right) stages.

2.3.1 Current-Voltage Converter

Fig. 2.14 shows a simple circuit that converts an input current signal to its corresponding output voltage signal. In the ideal case, the output is: $v_o = -R_2 \cdot i_2 = -R_2 \cdot i$

Observe that the input impedance is extremely low (the virtual ground of the OpAmp is always around 0V), and therefore the circuit is a good current reader. Moreover, the output impedance is also extremely low, which makes the circuit a good voltage source, too.

2.3.2 Voltage-Current Converter

The circuit that performs the inverse function with respect to the previous one (i.e. voltage to current) can be designed as shown in Fig. 2.15, on the left. Considering the OpAmp ideal, we can write: $i_L = i = \frac{v_s}{R_1}$

The current flowing through the load resistance R_L is independent of the load itself and is a function of the applied input voltage.

The same converter can also be constructed by employing a non-inverting configuration, as shown in Fig. 2.15, on the right.

Unfortunately, both circuits must have a floating load connected to nodes with a voltage different from the ground. There are many applications in which the load must be connected to the ground. In such applications, it is possible to use the circuit shown in Fig. 2.16. We can write the node equations for the inverting and non-inverting input pins, respectively as:

$\frac{v_s - v_1}{R_1} + \frac{v_o - v_1}{R_2} = 0$ and $\frac{v_o - v_2}{R_4} - \frac{v_2}{R_3} - i_L = 0$ Obtaining v_o and taking into account that $v_1 = v_2$, we have:

$$v_o = R_2 \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right) \cdot v_1 - \frac{v_s}{R_1} \cdot R_2 = \left(1 + \frac{R_2}{R_1}\right) \cdot v_1 - \frac{R_s}{R_1} \cdot v_s$$

and $i_L = \frac{v_o}{R_4} - v_2 \cdot \left(\frac{1}{R_4} + \frac{1}{R_3}\right) = \frac{v_1}{R_4} \cdot \left(1 + \frac{R_2}{R_1}\right) - \frac{R_2}{R_1 \cdot R_4} \cdot v_s - v_1 \cdot \left(\frac{1}{R_4} + \frac{1}{R_3}\right)$

Imposing $\frac{R_2}{R_1} = \frac{R_4}{R_3}$, we obtain: $i_L = -\frac{R_2}{R_1 R_4} \cdot v_s$ i.e. the current flowing in the load R_L is proportional to the input voltage v_S while that current is independent of the load itself.

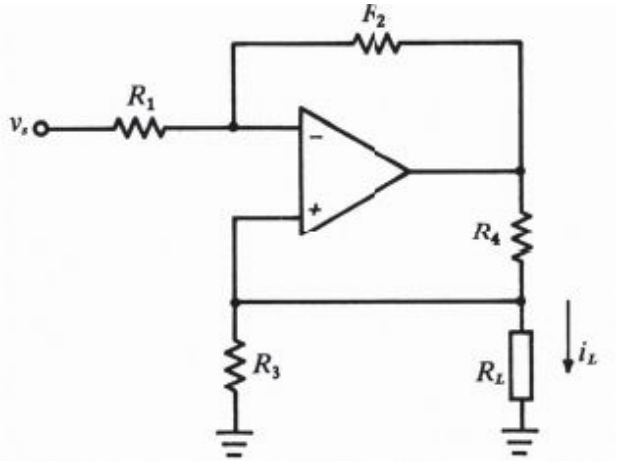


Fig. 2.16: Voltage-current converter with load connected to ground.

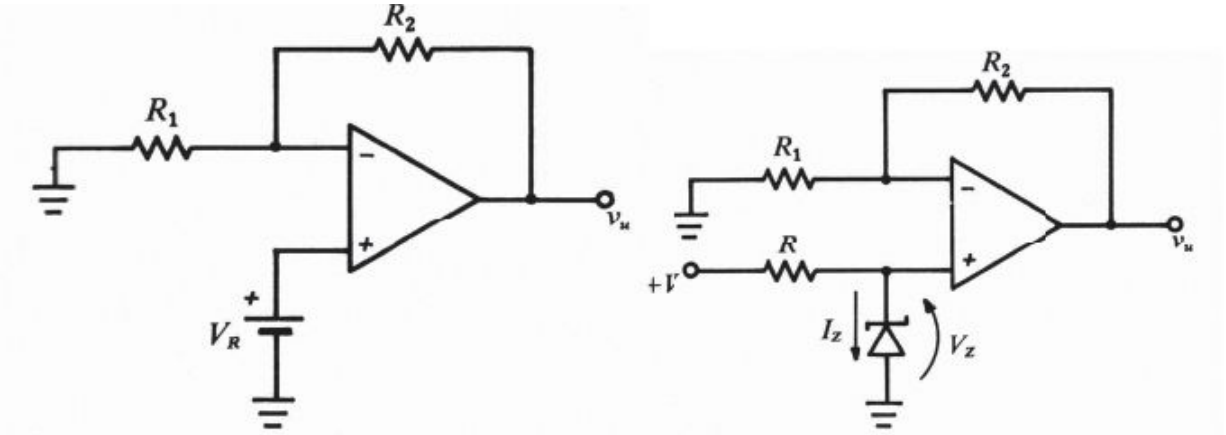


Fig. 2.17: Principle scheme of a reference voltage source (left) and reference voltage source with a Zener diode as stabilizer (right).

2.4. VOLTAGE REFERENCE

A voltage reference source must supply a current to a load without changing the output voltage. Therefore, the output resistance must be extremely low. Such circuits find use in power sources.

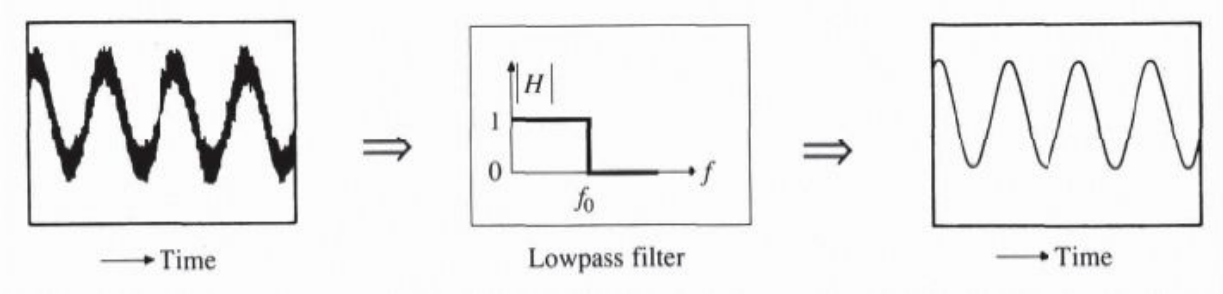
In principle, a voltage reference can be obtained by using the circuit shown in Fig. 2.17, on the left, in which a source V_R is applied to the non-inverting amplifier as an input. In this circuit, the OpAmp insulates the load form the source V_R.

Practically, the source reference V_R can be manufactured by means of a Zener diode as a stabilizer, as depicted in Fig. 2.17, on the right. In this way, we have $v_o = V_z \cdot \left(1 + \frac{R_2}{R_1}\right)$.

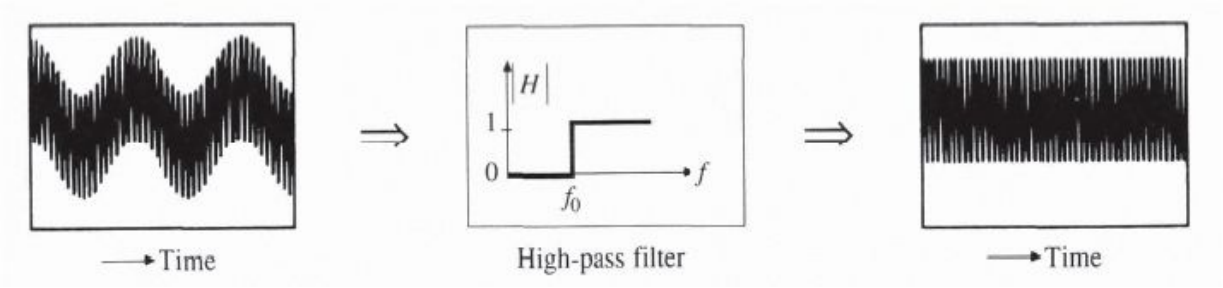
.

2.5. FILTERS

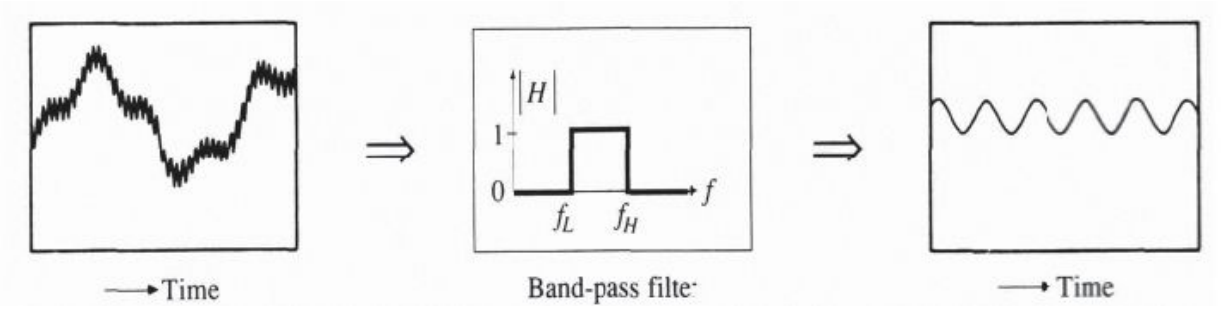
Depending on the frequency response (H), the filters are classified as follows: low-pass, high-pass, band-pass, and band-reject (or notch). The low-pass filter's frequency response (shown in the following) is characterized by the frequency f_0, named the cut-off frequency, such that at $f<f_0$, we have $|H|=1$ while at $f > f_0$, the response is $|H|=0$. This means that an input signal (i.e. sinusoids) whose frequency is less than f_0 will not change in amplitude while signals whose frequency is larger than f_0 will be totally attenuated. The typical application is the high frequency noise-canceling.



The frequency response of a high-pass filter is complementary to the low-pass one: sinusoids with $f>f_0$ will pass through unchanged while sinusoids with $f<f_0$ will be completely attenuated.

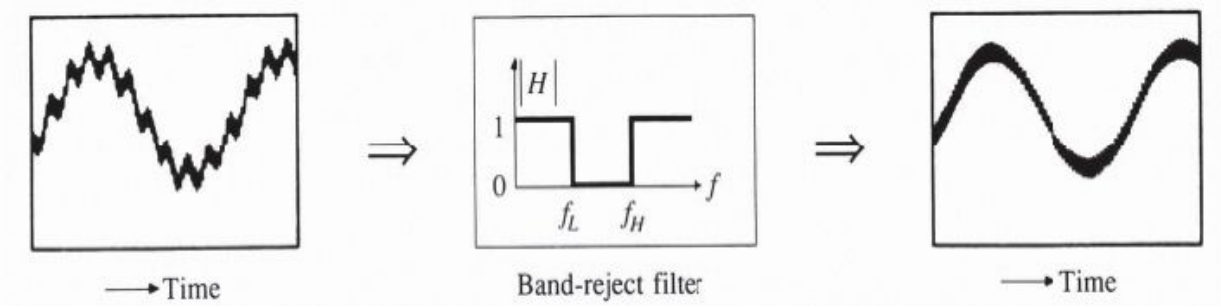


The frequency response of a band-pass filter is characterized by a frequency band $f_L<f<f_H$ in which sinusoids will pass through the filter without an amplitude alteration while those whose frequency is outside the band will be softened. The difference f_H-f_L is named *bandwidth*, and the central point is called the *center frequency*. A very common example for a band-pass filter is the tuning circuit for an ordinary AM/FM application that allows users to select a certain station and block the others.



The frequency response of a band-reject filter is complementary to that of a band-pass filter because it blocks the components within the band $f_L < f <$

f_H and leave the other frequencies unchanged. When the band is narrow enough, the response is named *notch*.



A fifth kind of response (not shown) is the so called *all-pass* filter, which shows a response with constant amplitude, but with a phase that varies linearly with frequency.

A filter can be analyzed in both the time domain and the frequency domain, as shown in Fig. 2.18 (the frequency spectrum of the input signal is shown on the left while the corresponding frequency spectrum of the output is shown on the right). Analyzing, for example, the low-pass filter, it can be noticed that the input signal is the sum of four different components (four frequency lines in the figure). The output signal is made from the single line that lies below the frequency f_0 while the other components are removed because the filter has a null gain ($|H(f)|=0$) at their respective frequencies. Often, filters are not employed to remove some frequency components, but to emphasize or de-emphasize frequencies. One common example can be found in a hi-fi system to control the tones.

Filters are classified according to their *order* which corresponds to the grade of the filter transfer function H (i.e. the order of the differential equation that binds output variables to input variables).

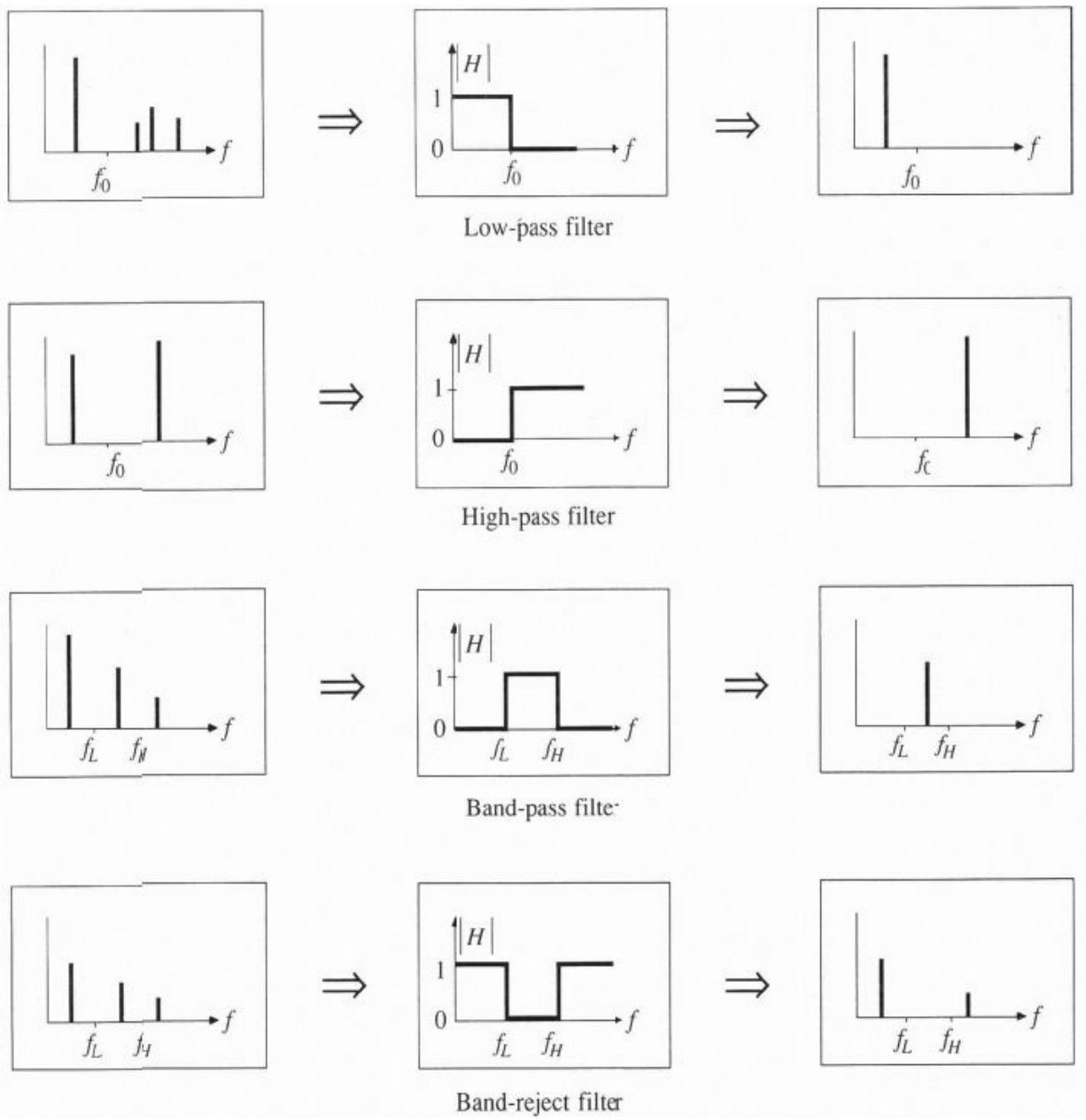


Fig. 2.18: Ideal responses (frequency domain).

2.6. First Order Active Filters

The simplest active filters can be obtained by replacing one or more resistors with reactive elements in an OpAmp inverting stage. In doing so, a frequency-dependent gain is achieved.

2.6.1 Inverting integrator

As seen before, if the feedback resistor is replaced with a capacitor in an inverting stage, the new configuration is an integrator (shown in Fig. 2.19). The transfer function is: $$H(s) = -\frac{Z_C}{Z_R} = -\frac{1/(sC)}{R} = -\frac{1}{RsC}$$

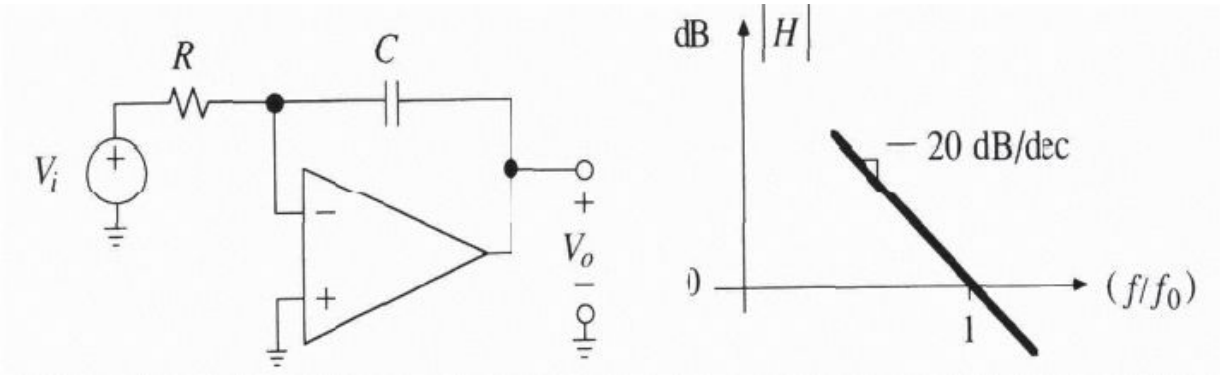


Fig. 2.19: Inverting integrator (left) and module of its transfer function (right).

Imposing s = jω and assuming that the pole is at $f_0 = \frac{1}{2\pi \cdot RC}$, it is possible to obtain the amplitude and the phase of the filter transfer function: $$|H| = \frac{1}{f/f_0} \qquad \text{and} \qquad \angle H = +90^\circ$$

Assuming $|H|_{dB}=20\cdot\log[1/(f/f_0)]=-20\cdot\log(f/f_0)$, the Bode diagram of the amplitude of H as a function of $\log(f/f_0)$ is a straight line, e.g. $y=-20\cdot x$, as shown in Fig. 2.20. This means that the integrator belongs to the low-pass filter class. Signals whose frequency is higher than f_0 are attenuated whereas those whose frequency is lower than f_0 are even amplified ($|H|>1$).

This behavior is not surprising since $|Z_C|=1/(2\pi f_0 C)$ depends on the frequency. At high frequencies, $|Z_C|<R$, so $|H|<1$, and the circuit attenuates the signals applied. At low frequencies, $|Z_C|>R$, and thus the circuit amplifies the signals applied. The borderline is the frequency f_0 such that $|Z_C|=R$, i.e. $1/(2\pi f_0 C)=R$. Obviously, the frequency f_0 is named the *unity gain frequency* of the integrator. Moreover, the integrator introduces a phase delay of 90° between the output and the input, regardless of the frequency. For this reason, integrators are used in quadrature oscillators (to simultaneously generate sine and cosine signals). If the input is a sinusoid with a frequency f_0, the output

will be a sinusoid with the same frequency and the same amplitude, but with a phase delay of -90°.

Because of the very high gain at low frequencies, integrators are rarely used alone because they tend to saturate. This can be avoided simply with a feedback connection that maintains the OpAmp in its linear zone.

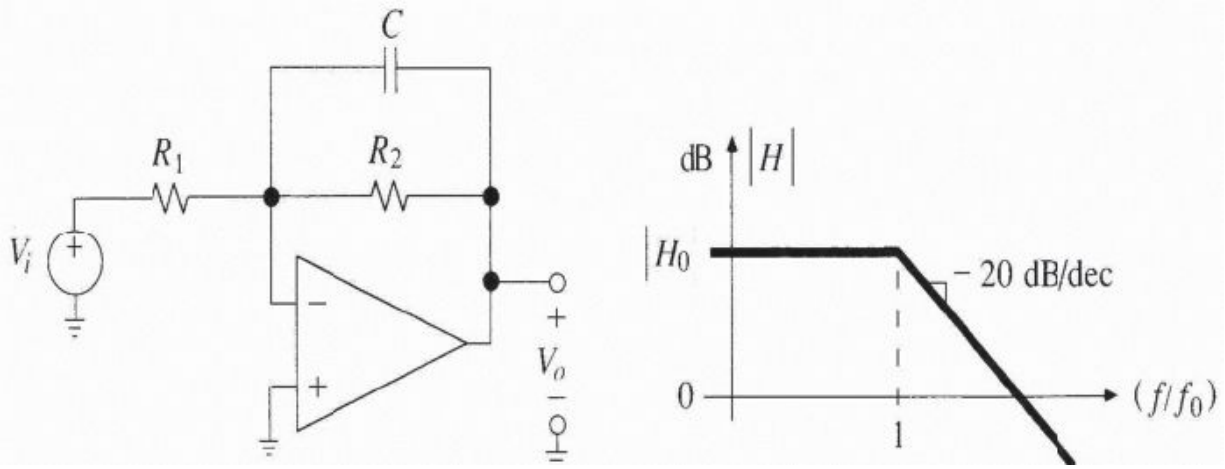


Fig. 2.20: Low-pass filter with finite gain.

2.6.2 Low-pass filter with finite gain

As seen before, a real integrator is obtained by adding a parallel resistor to the feedback branch of an inverting integrator. The real integrator can be regarded as a low-pass filter with a finite gain (Fig. 2.20).

The transfer function is simply derivable from $H=-Z_2/R_1$ with $Z_2=R_2//[1/(j\omega C)]=R_2/(1+j\omega R_2 C)$. Remembering $\omega = 2\pi f$, it is possible to write H in the following form:

$$H = H_0 \frac{1}{1+j(f/f_o)}, \qquad \text{where } H_0 = -\frac{R_2}{R_1} \text{ and } f_0 = \frac{1}{2\pi \cdot R_2 C}$$

In order to draw the Bode diagram of the amplitude of H, it is useful to notice that for $(f/f_0)<<1$, we have $H \cong H_0$. In fact, at low frequencies, $Z_2 \cong R_2$ holds because the capacitance can be regarded as an open circuit. Therefore, the stage is an inverting amplifier with a gain equal to $-R_2/R_1$. Obviously, H_0 is named the *filter DC gain*.

On the other hand, for $(f/f_0)>>1$, we have $H = H_0 \frac{j(f/f_0)}{1+j(f/f_0)}$. This indicates that the asymptote at high frequencies is similar to that of an integrator with a gain $|H_0|$ proportional to f/f_0. This happens because at high frequencies, the impedance Z_C of the capacitance is very small, and R_2 can be neglected. This circuit has the same behavior of the inverting integrator.

This circuit approximates the behavior of an integrator only for frequencies above f_0. At frequency f_0, we have the transition from the amplifier behavior to the integrator. For $(f/f_0)=1$, we have $H=H_0/(1+j)$, thus $|H|=|H_0|/\sqrt{2}$ or, equivalently, $|H|_{dB}=|H_0|_{dB}$-3dB. For this reason, f_0 is named the *–3dB frequency* or the *corner frequency*.

2.6.3 High-pass filter with finite gain

It is possible to obtain this filter by placing a capacitor in series with the input resistance in the classical inverting stage. The real derivator has already been demonstrated. With simple mathematical steps, one can demonstrate that: $H = H_0 \frac{j(f/f_0)}{1+j(f/f_0)}$

where $H_0= -R_2/R_1$ and $f_0=1/(2\pi R_1 C)$.

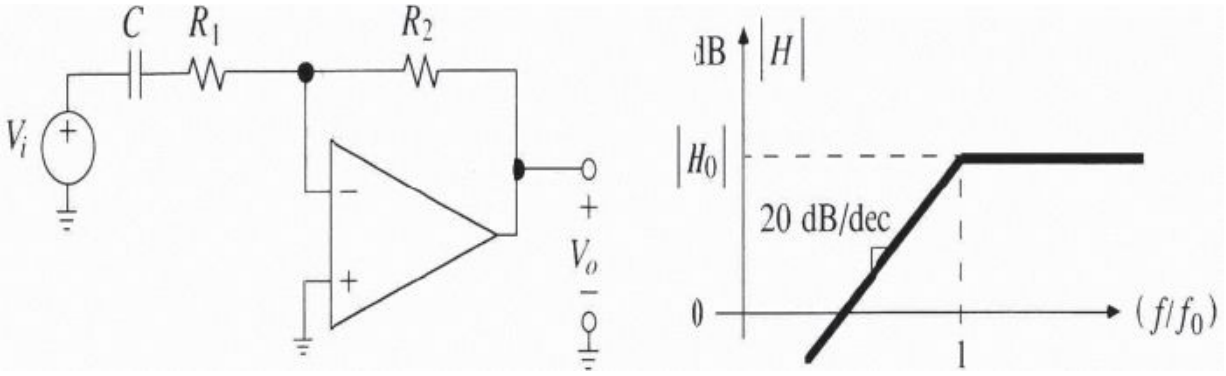


Fig. 2.21: High-pass filter with finite gain.

2.6.4 Band-pass filter

The circuits in Fig. 2.20 and Fig. 2.21 can be fused together, obtaining the band-pass filter shown in Fig. 2.22. The input impedance $Z_1=R_1+1/(j\omega C_1)$ is the high-pass filter with the corner frequency $f_1=1/(2\pi R_1 C_1)$ while the

feedback impedance $Z_2=R_2//[1/(j\omega C_2)]$ is a low-pass filter with the corner frequency $f_2=1/(2\pi R_2C_2)$. The input frequency within the band $f_1 \le f \le f_2$ will pass through the filter unchanged while those outside of this band will be attenuated by the first or the second stage. The transfer function will be $H=-Z_2/Z_1$, i.e.:

$$H = H_0 \frac{j(f/f_1)}{[1+j(f/f_1)][1+j(f/f_2)]}$$

where $H_1=-R_2/R_1$, $f_1=1/(2\pi R_1C_1)$ and $f_2=1/(2\pi R_2C_2)$. The Bode diagram for the module of the transfer function is shown in Fig. 2.22.

This filter is used in audio applications, where the signal components inside the vocal band (from 20Hz to 20kHz) need to be amplified whereas those at low frequencies (for example, the very low fluctuation due to the thermal drift) and the noise at high frequencies need to be blocked. The filter can be made very selective by making f_1 and f_2 very close to each other. Band-pass filters with a ratio f_2/f_1 equal to or more than 2 are classified as *wide-band* filters while those with f_2/f_1 less than 2 are classifies as *narrow-band* filters. Such classification is applied to filters with a notch response.

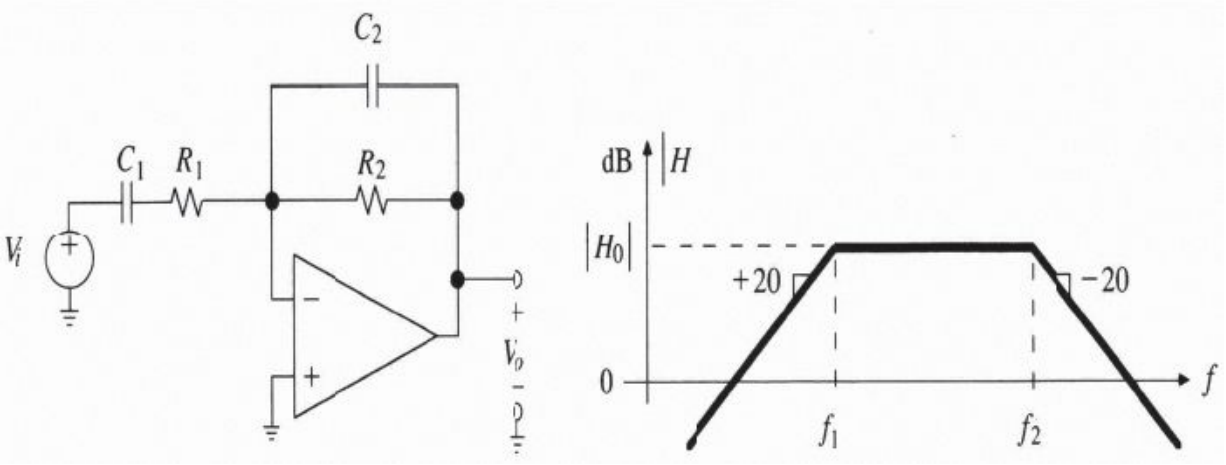


Fig. 2.22: Band-pass filter (left) and its module Bode diagram (right).

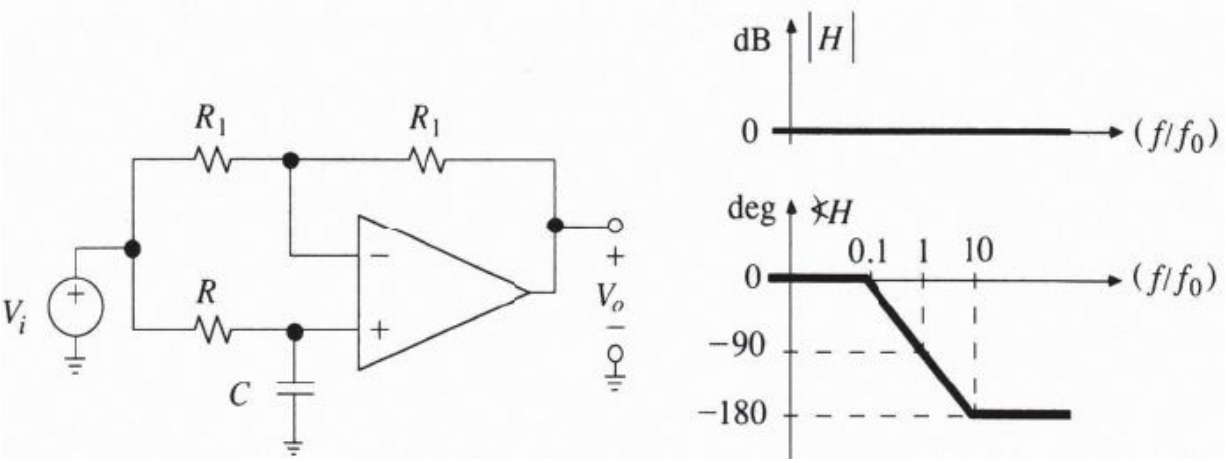


Fig. 2.23: Phase shifter (left) and its module and phase diagrams (right).

2.6.5 Phase Shifter

This kind of filters introduces a phase delay between its input and its output, maintaining the signal amplitude constant. Filters analyzed thus far are not a part of this family because they also modify the signal amplitude. Fig. 2.23 shows the typical implementation of a phase shifter.

It can easily be demonstrated that $H = \frac{1 - j(f/f_0)}{1 + j(f/f_0)}$, where $f_0 = \frac{1}{2\pi \cdot RC}$.

For this reason, we have:

$$|H| = \frac{\sqrt{[1+(f/f_0)^2}}{\sqrt{[1+(f/f_0)^2}} = 1 \qquad \angle H = \tan^{-1}[-(f/f_0)] - \tan^{-1}(f/f_0) = -2\tan^{-1}(f/f_0)$$

While the modulus of the frequency response is constant and equal to 1 (from which the name *all-pass filter* derives), the phase gradually changes from 0° to 180° (90° for f =f_0).

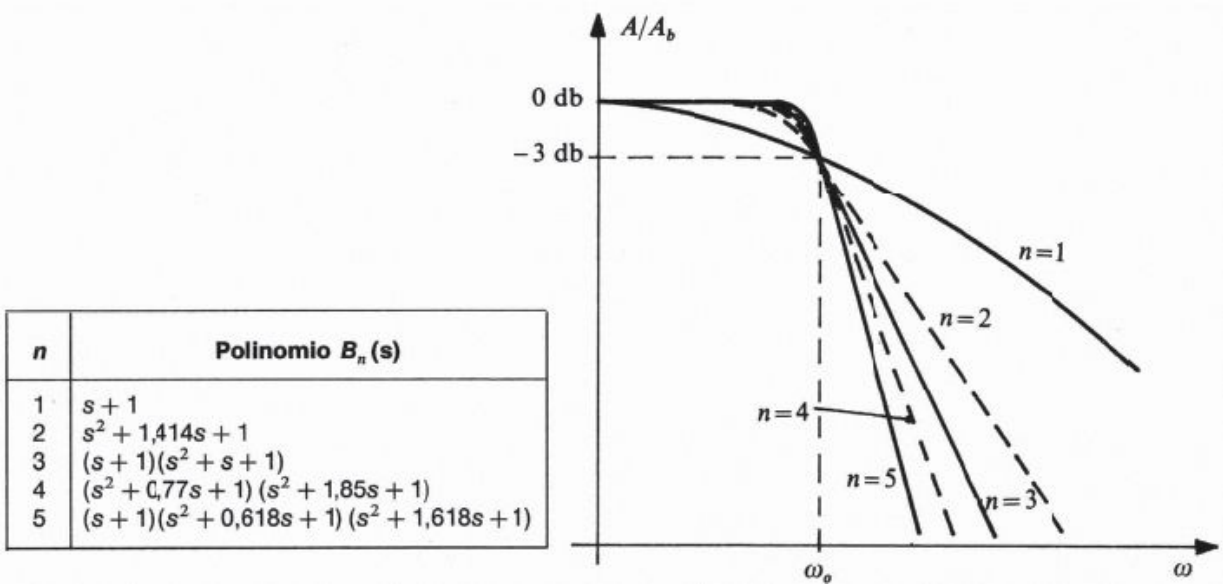
2.7. Butterworth Filters

In the following paragraphs, filters of an order higher than one will be described.

First of all, the Butterworth filters will be analyzed. The ideal response of such filters can be approximated with polynomials named Butterworth polynomials. Because the mathematical treatment is slightly complex and lies outside of the target of this book, the generic transfer function is quoted here:

$$A(s) = \frac{v_o}{v_i} = \frac{A_b}{B_N(s)}$$

where $s = j\frac{\omega}{\omega_0}$ is normalized with the angular cut-off frequency ω_0, A_b is the DC static gain, and B_N is the N^{th} order Butterworth polynomial. The higher N is, the more selective the frequency response becomes, approximating the ideal trend (with "vertical" transition), as shown in Fig. 2.24.



n	Polinomio B_n (s)
1	$s+1$
2	$s^2+1{,}414s+1$
3	$(s+1)(s^2+s+1)$
4	$(s^2+0{,}77s+1)(s^2+1{,}85s+1)$
5	$(s+1)(s^2+0{,}618s+1)(s^2+1{,}618s+1)$

Fig. 2.24: Typical low-pass responses for Butterworth Filters.

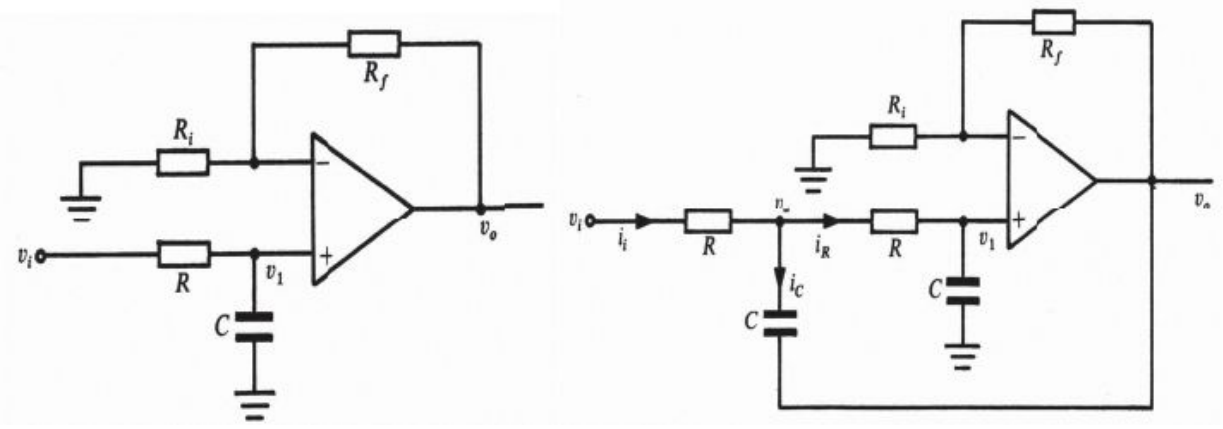


Fig. 2.25: First order (left) and second order (right) low-pass Butterworth filters.

The peculiarity of these filters lies in the absence of ripples within its bandwidth. The gain is reduced by -3dB at ω_0 for every order N of the polynomial, after which the gain changes with a slope of –20·N dB/dec. Subsequently, the low-pass typology will be treated as all other typologies can be obtained simply from that typology, as will be shown later.

2.7.1 First order low-pass Butterworth filter

The first order low-pass Butterworth filter is shown in Fig. 2.25, on the left. Because the OpAmp does not sink any current, the voltage v_1 of the non-inverting input pin is: $v_1 = \frac{v_i}{R+\frac{1}{j\omega C}}\frac{1}{j\omega C} = \frac{v_i}{1+j\omega CR} = \frac{v}{1+j\frac{\omega}{\omega_0}}$, where $\omega_0 = 1/RC$. The non-inverting gain is $G = \frac{v_0}{v_1} = 1+\frac{R_f}{R_i}$. For this reason, we get:

$$H = \frac{v_o}{v_i} = \frac{v_o}{v_1}\frac{v_1}{v_i} = \left(1+\frac{R_f}{R_i}\right)\frac{1}{1+j\frac{\omega}{\omega_0}}$$

2.7.2 Second order low-pass Butterworth filter

The second order low-pass Butterworth filter is shown in Fig. 2.25, on the right, in which a second passive RC network establishes feedback between the output and the non-inverting input pin.

Because the configuration is non-inverting, we have $v_o = Gv_1 = \left(1+\frac{R_f}{R_i}\right)v_1$. And because the OpAmp does not sink any current, it is possible to write $i_R = \frac{v_a}{R+\frac{1}{j\omega C}}$ and $i_C = \frac{v_a - v_o}{\frac{1}{j\omega C}} = j\omega C(v_a - Gv_1)$. The input current is $i_i = i_R + i_C$ and:

$$v_a = v_i - Ri_i = v_i - \frac{j\omega CRv_a}{1+j\omega CR} - j\omega CR(v_a - Gv_1) = \frac{v_i + j\omega CRGv_1}{1+j3\omega CR-(\omega CR)^2}(1+j\omega CR)$$

The non-inverting input voltage can be computed as:

$$v_1 = \frac{v_a}{R + \frac{1}{j\omega C}} \frac{1}{j\omega C} = \frac{v_a}{1 + j\omega CR} = \frac{v_i + j\omega CRGv_1}{1 + j3\omega CR - (\omega CR)^2}$$

Remembering that $v_1 = v_o/G$, we obtain:

$$\frac{v_o}{G} = \frac{v_i + j\omega CRv_o}{1 + j3\omega CR - (\omega CR)^2}$$

Finally:

$$H = \frac{v_o}{v_i} = \frac{G}{1 + j\omega CR(3 - G) - (\omega CR)^2} = \frac{G}{1 + j\frac{\omega}{\omega_o}(3 - G) - \left(\frac{\omega}{\omega_o}\right)^2} \qquad (\omega_0 = 1/RC)$$

and imposing $s = j\omega/\omega_0$: $H = \frac{G}{s^2 + (3 - G)s + 1}$

Because the denominator must coincide with the second order Butterworth polynomial, $B_2(s) = s^2 + 1,414s + 1$, we need to choose 3-G=1,414, i.e. $G=1+(R_f/R_i)=3-1,414=1,586$.

2.7.3 Nth order low-pass Butterworth filters

To build Butterworth filters of an order higher than 2, it is necessary to cascade a number of filters with lower orders so as to attain the required order. The global bandwidth gain is equal to the product of single filters' gains.

Filters of an odd order greater than 1

One can build this kind of filters by the cascade connection of even and odd order filters to satisfy the corresponding Butterworth polynomial. For example, a filter of an order n=5 is made with a cascade of one 1st order filter and two 2nd order filters, as shown in Fig. 2.26.

Because $B_5 = (s + 1)(s^2 + 0,618s + 1)(s^2 + 1,618s + 1)$, the followings must hold: • filter A: the gain G can assume any value while $f_{0A}=f_H$ (the desired cut-off frequency);

• filter B: because 3-G=0,618, the bandwidth gain is G=2,382 and $f_{0B}=f_H$;

• filter C: because 3-G=1,618, the bandwidth gain is G=1,382 and $f_{0C}=f_H$.

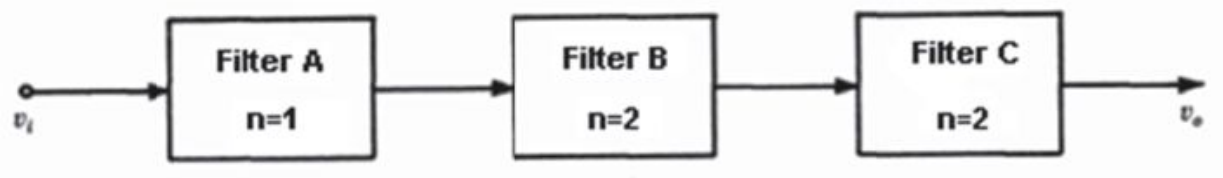


Fig. 2.26: Fifth order Butterworth filter.

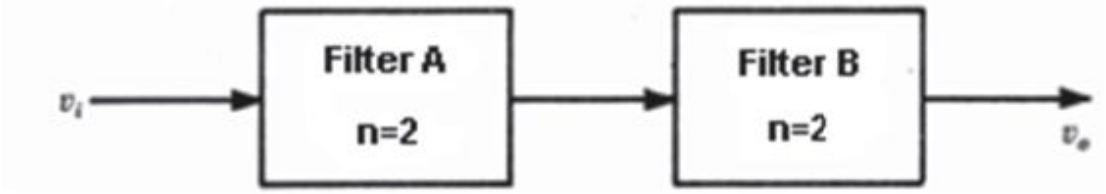


Fig. 2.27: Fourth order Butterworth filter.

Filter of an even order greater than 2

To build Butterworth filters of an even order higher than 2, it is necessary to establish a cascade connection of a number of filters with lower orders so as to attain the required order. For example, a filter of an order n=4 is made with a cascade connection of two 2nd order filters (Fig. 2.27). Because $B_4 = (s^2 + 0,77s + 1)(s^2 + 1,85s + 1)$, the followings must hold: • filter A: because 3-G=0,77, the bandwidth gain is G=2,23 and $f_{oA}=f_H$;

• filter B: because 3-G=1,85, we need G=1,15 and $f_{0B}=f_H$.

2.7.4 High-pass, Band-pass and notch Butterworth filters

All the typologies of Butterworth filters can be obtained by exchanging positions between R and C elements, both of which define the cut-off frequency, in a low-pass filter, as shown in Fig. 2.28.

Setting up a cascade connection of a low-pass filter with a high-pass filter (provided that $f_H>f_L$), we can obtain a band-pass filter with wide bandwidth, in which the overall gain at the center-band is equal to the product of single filters' gains.

Provided that f_H is less than f_L, by connecting a high-pass filter in parallel with a low-pass filter (Fig. 2.29) and by adding their outputs with an adder circuit with an equal gain for the two inputs, a notch filter is obtained.

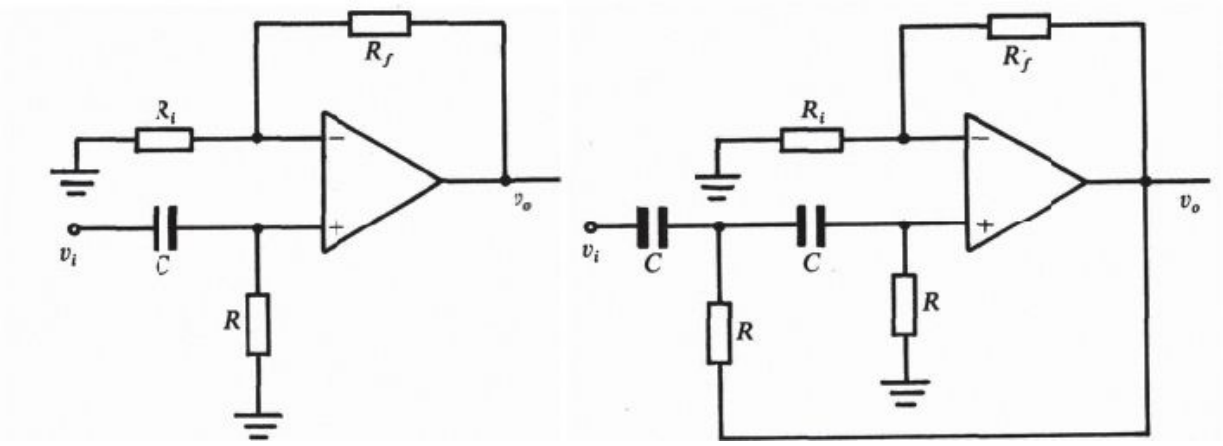


Fig. 2.28: First order (left) and second order (right) high-pass filters starting from the low-pass stage, exchanging the capacitances and resistances positions.

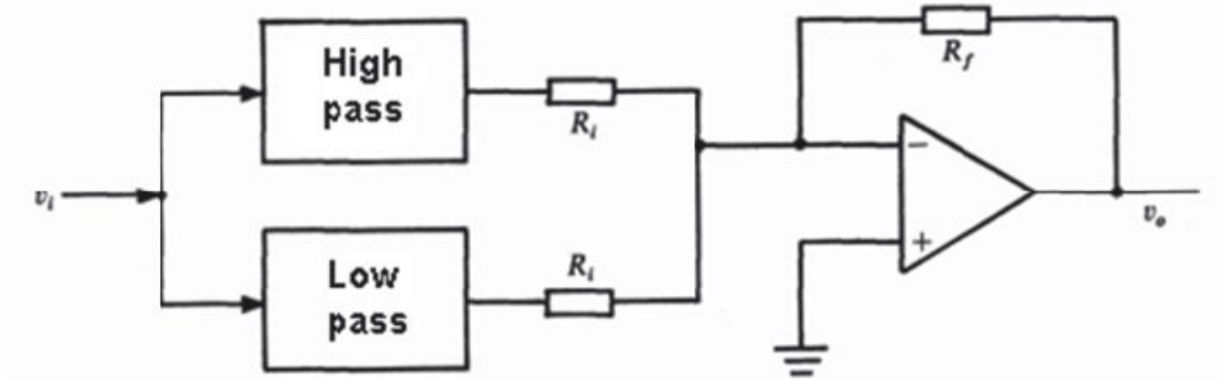


Fig. 2.29: Notch filter, starting from different topologies filters.

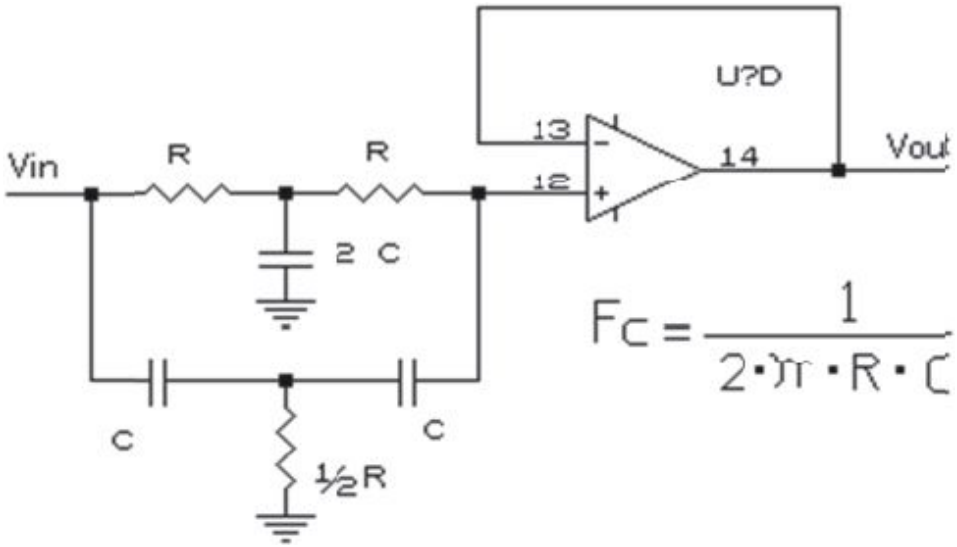


Fig. 2.30: Double-T notch filter.

For example, consider the double-T notch filter shown in Fig. 2.30. The filter frequency f_c is expressed in the same figure. If we want to reduce the power noise at f_c=50Hz, imposing C=47nF, we will have 2C=94nF (this is not a commercially preferred value, so the capacitance is built with two 47nF capacitors connected in parallel) with R=67.6kΩ and ½R=67.7/2=33.8kΩ. To have a good attenuation centered at 50Hz, the components must have high thermal stability with a degree of tolerance less than 0.5%.

To make the notch filter more selective, we can connect the 38.85kΩ resistance and the 94nF capacitance to the central pin of a voltage divider powered by the output signal, as shown in Fig. 2.31. Thanks to the trimmer, a fine calibration can be made.

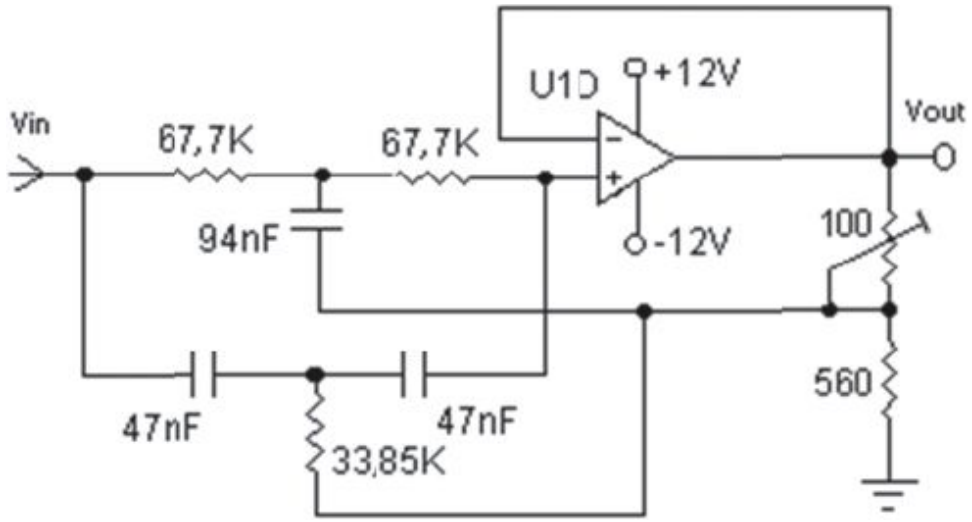


Fig. 2.31: Notch 50Hz filter sizing.

2.8. Chebyschev Filters

In some applications, a brief transition between the pass-band and the out-band attenuation is of the utmost importance. This requirement can be satisfied by Chebyschev filters, which, on the contrary of Butterworth filters, have the in-band ripple. For a given n^{th} order, the higher the tolerable ripple, the shorter the filter transition band. The attenuation depends only on the

filter order, as for Butterworth filters. For the same attenuation, a Chebyschev filter is cheaper and less complex than its Butterwoth counterpart.

	$\alpha = -0{,}5$ dB (0,9441)		$\alpha = -1$ dB (0,8913)	
n	N_n	Polynomial C_n	N_n	Polynomial C_n
1	2,863	$s + 2{,}863$	1,965	$s + 1{,}965$
2	1,431	$s^2 + 1{,}426s + 1{,}516$	0,983	$s^2 + 1{,}098s + 1{,}103$
3	0,716	$(s + 0{,}626)(s^2 + 0{,}626s + 1{,}142)$	0,491	$(s + 0{,}494)(s^2 + 0{,}494s + 0{,}994)$
4	0,357	$(s^2 + 0{,}351s + 1{,}063)$ $(s^2 + 0{,}847s + 0{,}356)$	0,246	$(s^2 + 0{,}674s + 0{,}275)$ $(s^2 + 0{,}279s + 0{,}986)$
5	0,179	$(s + 0{,}362)(s^2 + 0{,}224s + 1{,}036)$ $(s^2 + 0{,}586s + 0{,}477)$	0,123	$(s^2 + 0{,}289)(s^2 + 0{,}468s + 0{,}429)$ $(s^2 + 0{,}179s + 0{,}988)$

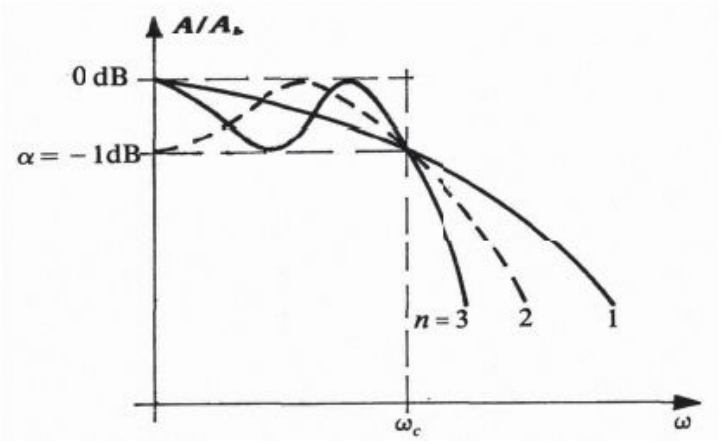


Fig. 2.32: Chebyschev polynomial (above) and typical responses for an implementation (below) varying the filter order and the α admitted ripple.

The transfer function for the low-pass filter is: $A(s) = \frac{v_0}{v_i} = \frac{N_n}{C_n(s)} A_b$

where $s=j\omega/\omega_n$ is normalized with the angular natural frequency ω_n which is coincident with the angular corner frequency ω_c (at which the response definitively leaves the ripple zone). N_n represents, for an odd n, the value of the denominator at DC and, for an even n, the denominator lowered by the ripple α. In this equation, A_b is the maximum bandwidth gain, and C_n is the n^{th} order Chebyschev polynomial. With the same order n, for the C_n polynomial, we can implement various solutions based on the maximum tolerable ripple α.

2.9. Bessel Filters

Bessel filters introduce a phase delay proportional to the frequency. If the phase delay is linear with frequency, the main effect is the introduction of a time delay on the signal. Vice versa, for a non-linear phase delay, input harmonic components with different frequencies will pass through the filter suffering from different delays (causing signal distortion). This phenomenon can be observed by applying a square wave to the filter. In the case of a linear response, the output will only be a simple delayed replica of the input signal. On the other hand, in the case of a non-linear response, the output will present a significant ringing.

Bessel filters, also called Thompson filters, maximize the time delay within the band of the filter and, as for the Butterworth filters, present a response which is not affected by the ripple effect, therefore, very flat. The result is a transfer function with a phase characteristic linear within the bandwidth and a transition slope not very abrupt.

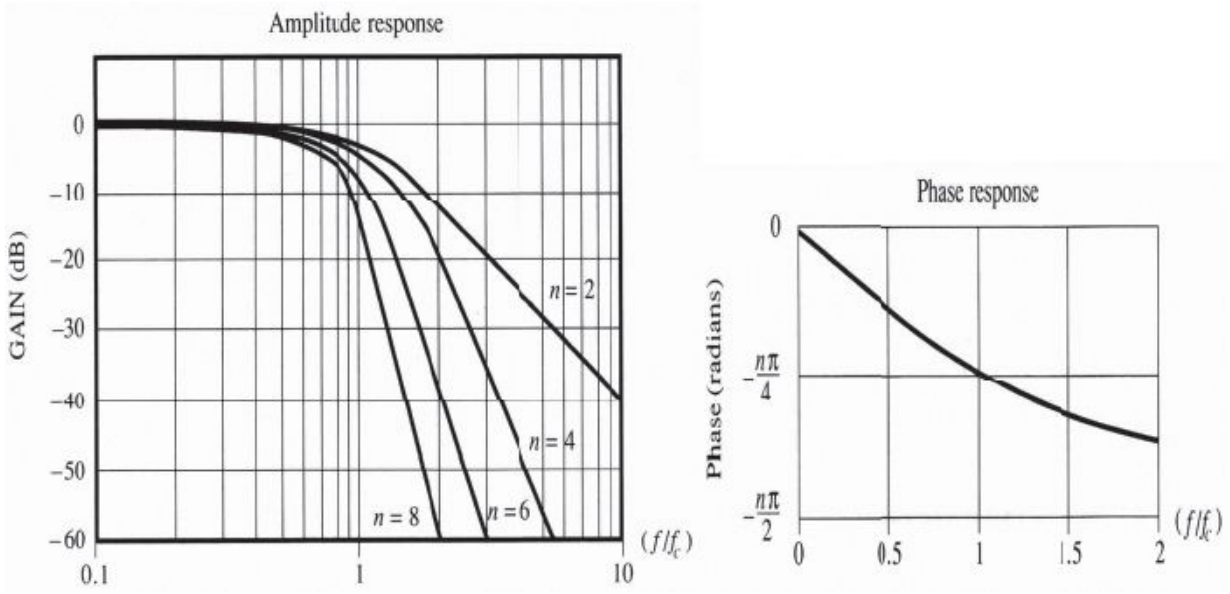


Fig. 2.33: Typical responses for Bessel filters.

2.10. Elliptical Filters (Cauer)

This last filter typology is characterized by a frequency transfer function with a ripple both inside and outside the band, thereby ensuring a more

abrupt transition outside the bandwidth (Fig. 2.34). The concept is the following: a notch filter is connected in cascade with a low-pass filter for a more selective response. Obviously, the notch has to be very narrow to ensure the correct work of this solution; thus, a second notch is placed, the curve is brought down, and the process is repeated until the envelope of the curve outside the band is below the minimum attenuation required.

Compared with all other filter typologies, with the same selectivity, the elliptical filters require the lowest order and thus less implementation complexity.

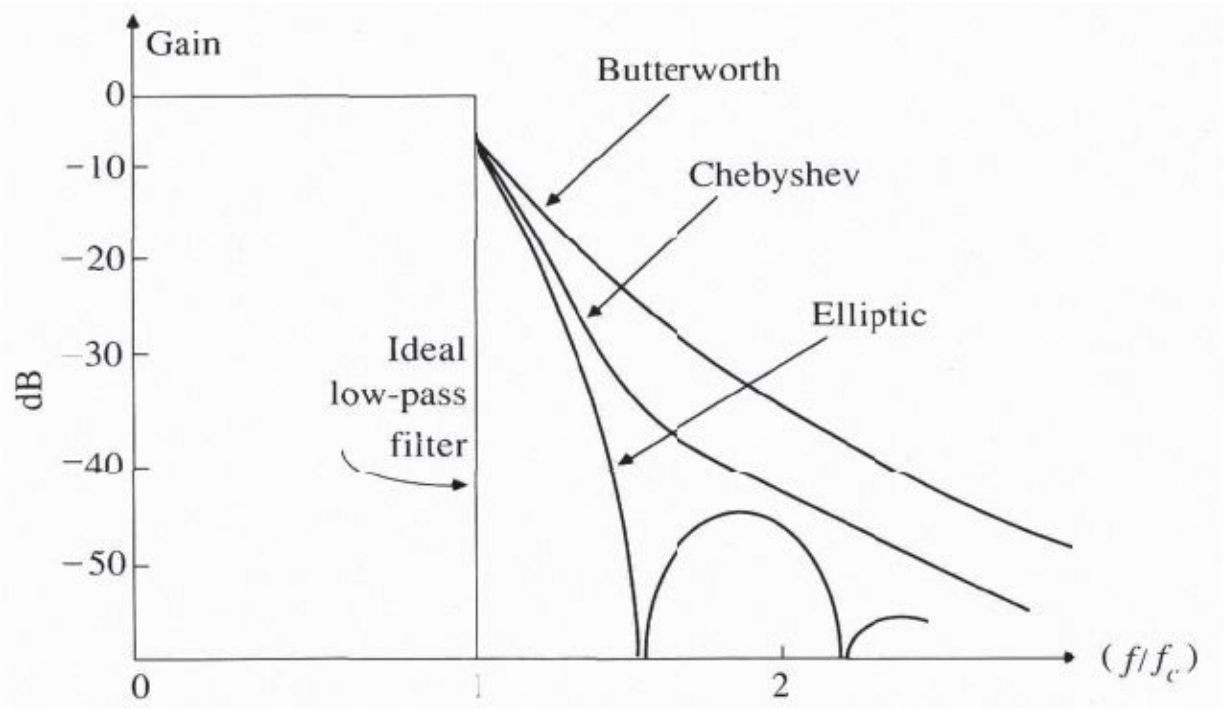
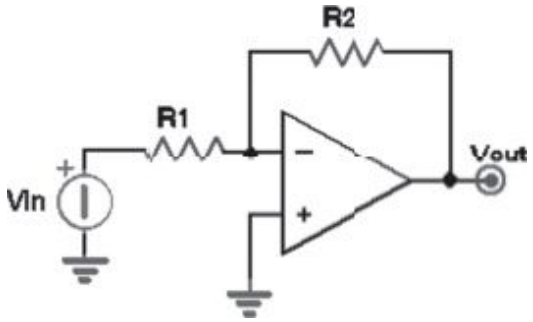


Fig. 2.34: Comparison between different filters with same order.

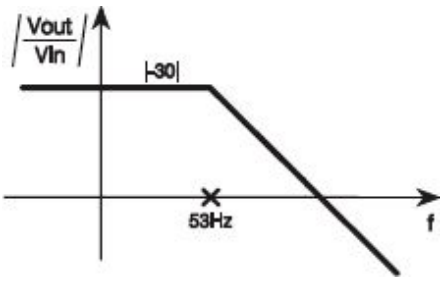
2.11. EXERCISES: LINEAR CIRCUITS WITH OPAMP

Ex. 1

Starting from the circuit on the right, with an ideal OpAmp:
a) What is the input impedance seen by the signal *Vin?*
b) With *R1=10kΩ*, size *R2* in order to have a gain equal to *Vout/Vin=-30*;
c) Introduce a capacitor *C2=10nF* in parallel to *R2*; draw the quoted Bode diagram.



a) The input impedance is *R1* because the OpAmp is in inverting configuration.
b) It is enough to choose *R2=30·10kΩ=300kΩ*.
c) The pole is at: $pole = \frac{-1}{R_2 \cdot C_2} = -53Hz$

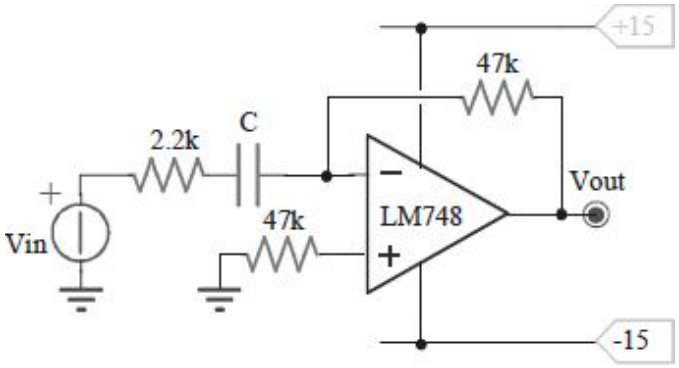


Ex. 2

An approximated derivator with a non-compensated LM748 (A_0=*105dB*, the second pole at *1MHz*, and A_{min}=*20dB*) is shown.

a) Determine the pole frequency.

b) Draw the Bode diagram.

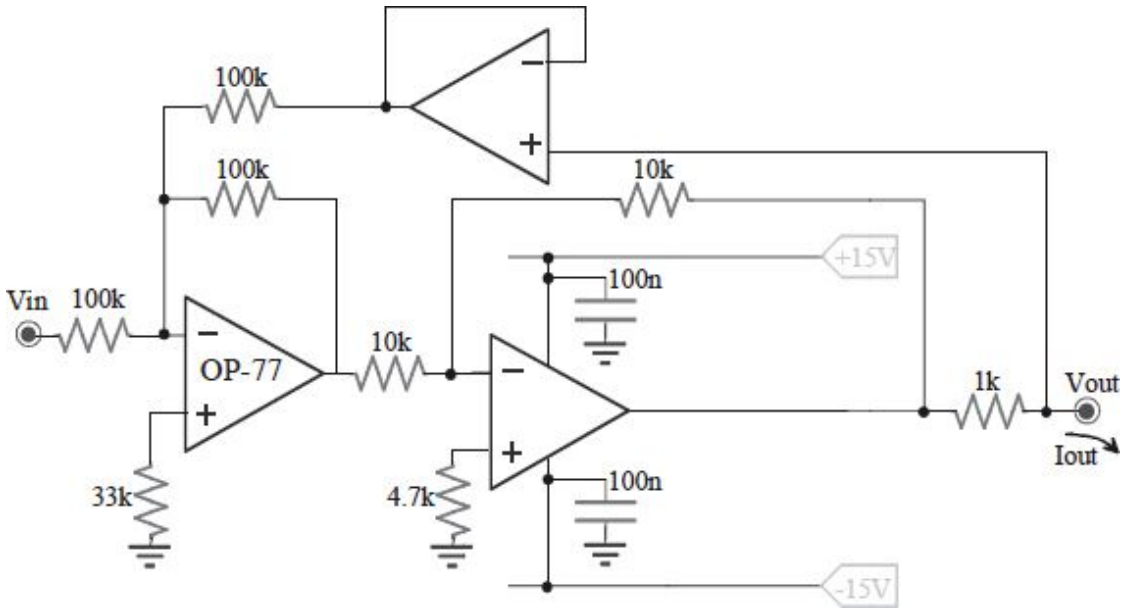


Ex. 3

The circuit uses three OpAmps with $A_0=120dB$. Compute:

a) the relationship between V_{out}, I_{out}, and V_{in} and the ranges of variability for V_{out} and I_{out}.

b) Calculate the small signal output impedance.

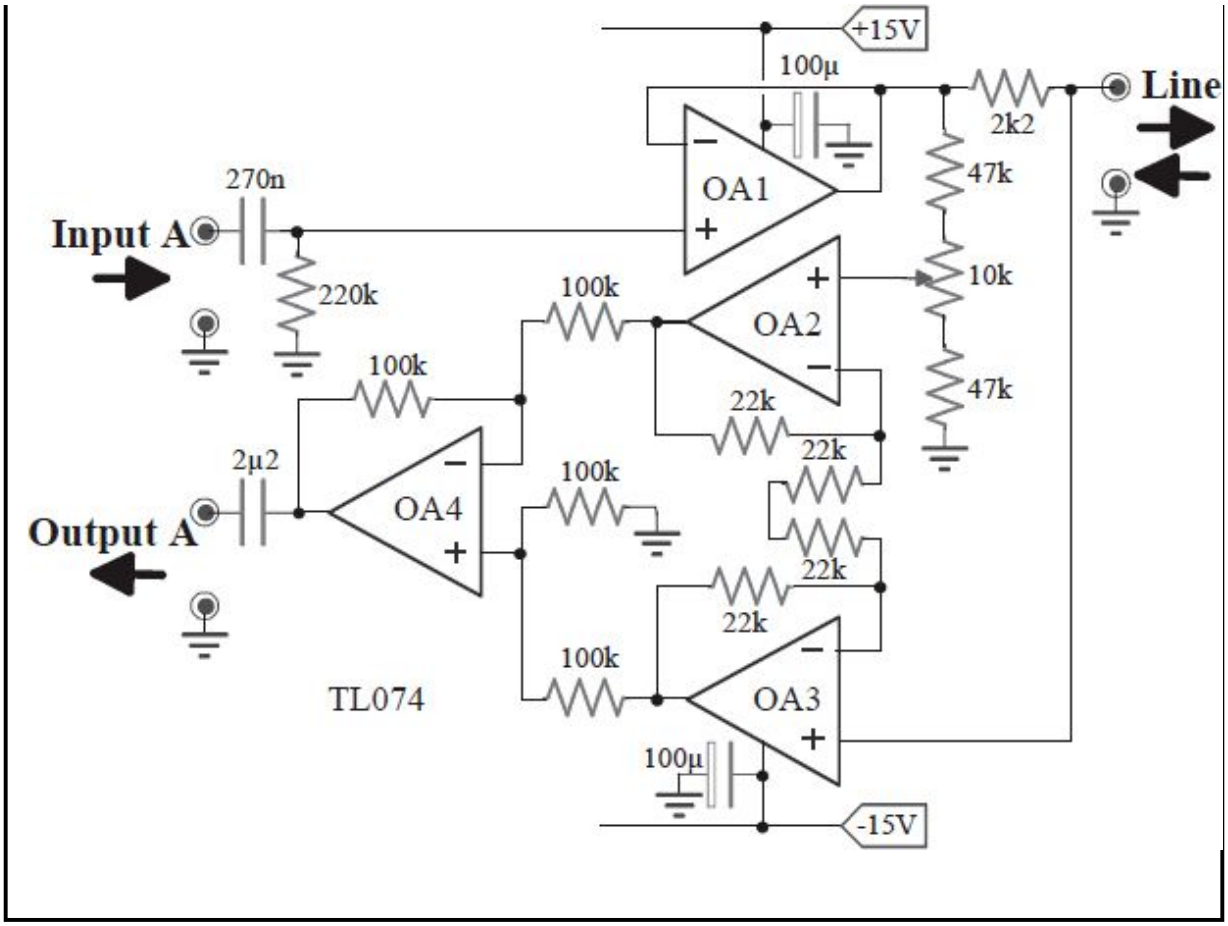


Ex. 4

The circuit is a "*Duplex Audio Link*", capable of using the *Line* for both transmitting the *Input A* signal from the left to the right and receiving a signal from the remote user from the right to the left via sending it to the terminal *Output A*. The remote circuit is identical to this (notice the impedance matching).

a) Extract all the relationships input/output, i.e. *Line/Input A*, *Output A/Input A*, and *Output A/Line* when the trimmer is correctly calibrated.

b) Compute the residual crosstalk on the *Output A* in *dB* due to the signal *Input A*, which is caused by a degree of resistive tolerance of *1%*.

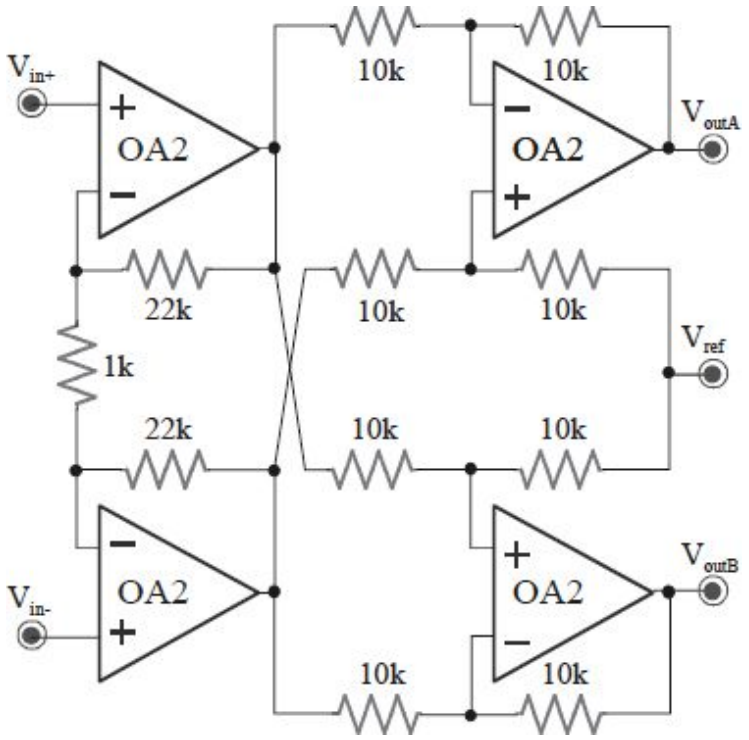

+15V
100µ
Line
2k2
47k
10k
47k
OA1
OA2
OA3
OA4
270n
Input A
220k
100k
100k
22k
22k
2µ2
Output A
100k
22k
22k
100k
TL074
100µ
-15V

Ex. 5

A quad OpAmp is supplied with $\pm 12V$ and with the pin $V_{ref}=+5V$.

a) Extract the relationship between V_{outA} and V_{outB} as a function of the input differential signal.

b) Assuming an offset current of $I_{OS}=10nA$, determine the worst case for the output offset over-imposed to V_{outA} and V_{outB}.



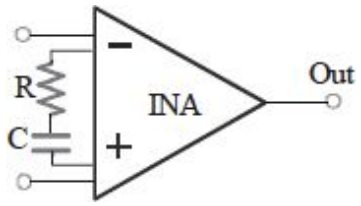
Ex. 6

The INA has all of its internal resistance values equal to $20k\Omega$. Moreover, $R=200\Omega$ and $C=47\mu F$.

a) Draw the internal structure and the quoted Bode diagram (with the analytical expression of poles and zeros introduced by R and C and for

the DC gains at medium and high frequencies) of the modulus of the gain $V_{Out}(s)/V_{diff,\ In}(s)$.

b) Determine the minimum *GBWP* that the OpAmps of the INA must have in order to have the cut-off frequency over *100kHz*.

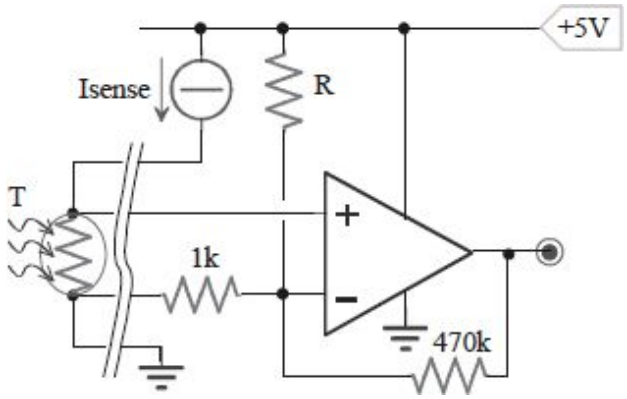


Ex. 7

The OpAmp has $A_0=100dB$, $CMRR=90dB$, and $I_B=500nA$. For $T=0\div100°C$, we want $V_{Out}=0\div4.7V$. The thermo-resistance is $200\Omega+1\Omega/°C$.

a) Size I_{sense} and R to ensure $V_{Out}=0$ @ $T=0°C$. For these values, compute the residual error for V_{Out}.

b) Starting from the chosen value of R, compute V_{Out} for $T=0°C$ and $T=100°C$, explaining the reasons of the deviation of the ideal straight line with a slope $4.7V/100°C=47mV/°C$ and suggesting possible solutions.



Ex. 8

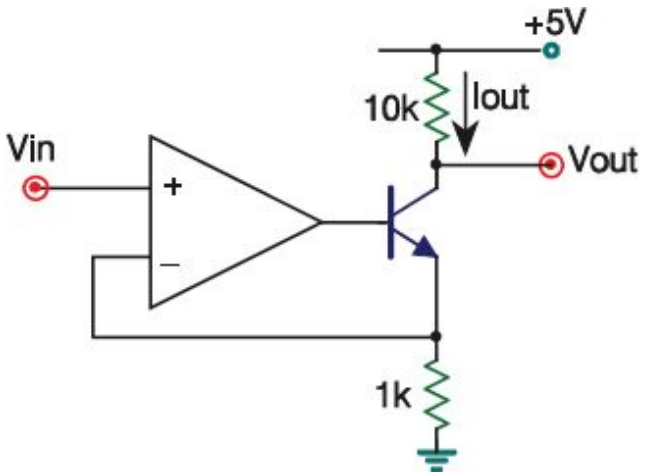
Given the circuit on the right side:

a) Extract the relationship between I_{out} and V_{in};

b) Obtain the relationship between V_{out} and V_{in};

c) What is the function performed by the OpAmp, the BJT, and the 1kΩ resistance?

d) What is the maximum value for I_{out} above which the BJT does not work properly? Why?



a) $I_{out} = \frac{V_{in}}{1k}$

b) $V_{out} = 5 - V_{in} \cdot \frac{10k}{1k} = 5 - V_{in} \cdot 10$

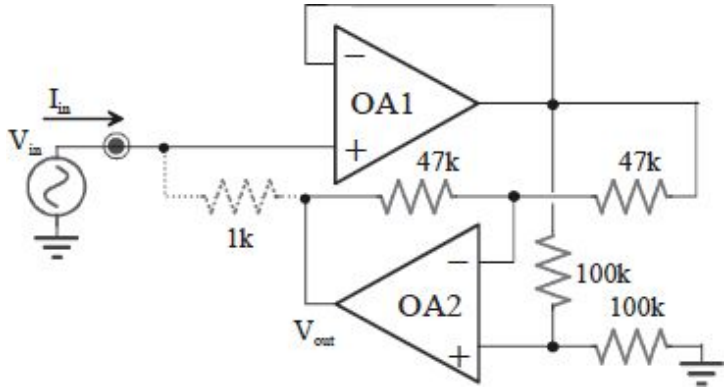
c) voltage-current converter.

d) When the BJT saturates, i.e. with $V_{CE,sat}=0.2V$, we have $5\text{-}10k\cdot I_{max}\text{-}1k\cdot I_{max}=V_{CE,sat}=0.2V$. Thus $I_{max}=436\mu A$.

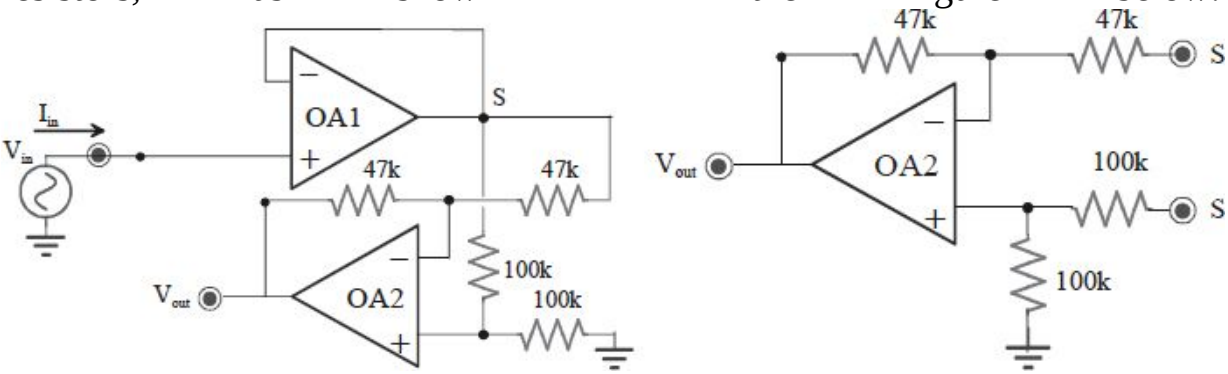
Ex. 9

a) Without the $1k\Omega$ resistance, compute the relationship between the input voltage V_{in} and the output voltage V_{out}.

b) Inserting the *1kΩ* resistance, obtain the relationship between input voltage and input current.



a) Analyzing the circuit, we can notice that the OA1 amplifier is a Buffer configuration, and OA2 makes a difference amplifier. At the node S, we find the input voltage V_{IN}, thus V_{OUT} can be thought of as the superimposition of effects of the two equal input voltages V_{IN} on the first 100kΩ and 47kΩ resistors, as shown in the figure below:



The V_{OUT} voltage is: $S^{-}\left(-\frac{47k}{47k}\right)+S^{+}\left(\frac{100k}{100k+100k}\right)\cdot\left(1+\frac{47k}{47k}\right)=0$ and the ratio $\frac{V_{OUT}}{V_{IN}}=0$.

b) Connecting the 1kΩ resistance, the output voltage V_{OUT} is 0V (OA2 is a difference amplifier). The input current, which flows through the 1kΩ branch, assuming OA1 ideal, is thus: The ratio $\frac{V_{IN}}{i_{IN}}$ is equal to 1kΩ.

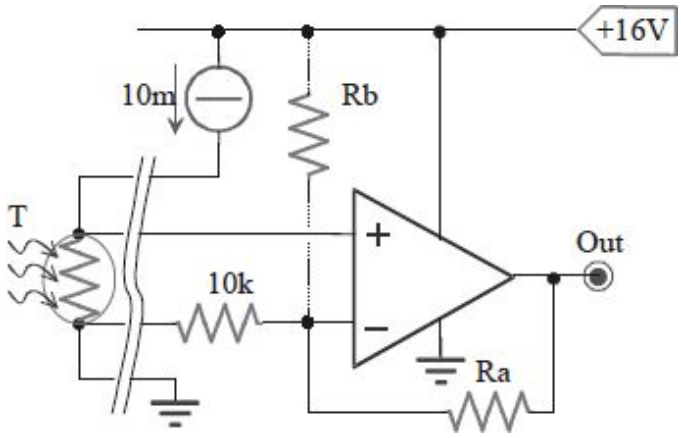
Ex. 10

Assume that the OpAmp is ideal. We want to measure the temperature *T=0÷600°C* through a thermo-resistance that is *100Ω* at 0°C and that grows by *0.5Ω/°C*.

a) Without *Rb*, size *Ra* in order to have V_{Out}*=12V* at *600°C*.

b) Introduce *Rb* and size it in order to have V_{Out}*=0* at *0°C*.

c) Modify *Ra* to satisfy the condition of a).



a) The thermo-resistance is R=100Ω+0.5Ω/°C·T(°C). At 600°C, we have $R_{@600}$=100Ω+0.5Ω/°C·600°C=400Ω. In these conditions, the input voltage is V_{IN}=10mA·400Ω=4V, thus the required gain is V_{OUT}/V_{IN}=12V/4V=3.

Because the stage is a non-inverting configuration, the gain is equal to: , from which we can obtain R_a=20kΩ.

b) At 0°C, the thermo-resistance is 100Ω. Thus, the input voltage is equal to 10mA·100Ω=1V. Because of the OpAmp gain, the output voltage must be equal to 1V·3=3V. This is an undesired baseline that reduces the signal dynamic range, and it will hence be better to avoid it. To lower the output of 3V, it will be possible to inject a current into the virtual ground node.

The inverting input pin presents the same voltage as the non-inverting one, i.e. 1V. The current i_b flowing through the resistance Rb is equal to (16V-

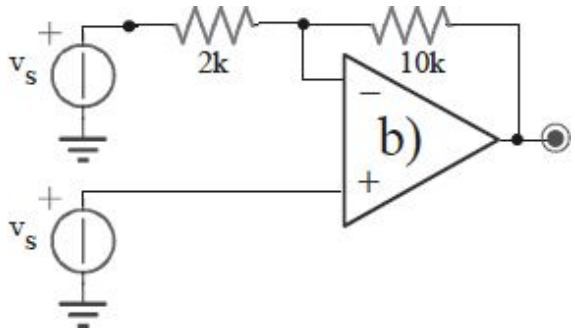
1V)/Rb. The current in the 10kΩ branch is equal to (1V–0V)/10kΩ=100μA. The current flowing through the Ra branch needs to be such that it fixes the output voltage to 0V, thus the current should be (1V-0V)/20kΩ=50μA. It is possible to obtain: $\frac{15V}{Rb} - 100\mu A = 50\mu A \rightarrow Rb = 100k\Omega$

c) Repeating the procedure of b), we can find that the input voltage corresponding to 600°C is equal to 4V. The current flowing through the resistance Rb is equal to (16V–4V)/100kΩ=120μA. The current flowing through the 10kΩ resistance is (4V–0V)/10kΩ=400μA, thus the current flowing through Ra is 400μA–120μA=280μA. In order to have the output equal to 12V, we should size: $Ra = \frac{(12V - 4V)}{280\mu A} = 28{,}6k\Omega$ .. In this way, the baseline is away from 0°C. So, how can we size the circuit to ensure an output between 0V and 12V when the temperature is between 0° and 600°C?

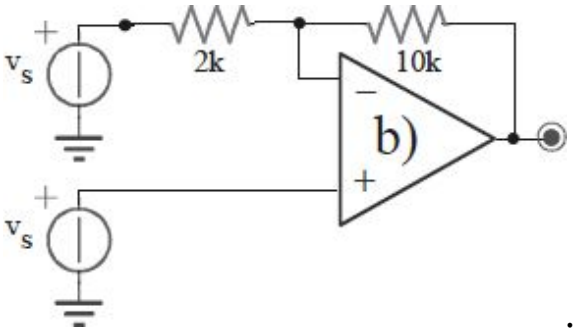
Ex. 11

A_0=100dB, i_B=100nA, v_{OS}=5mV.
Calculate:
a) the gain v_{out}/v_s and the impedance seen by v_s.
b) the error due to v_{OS}.
c) the error due to i_B.

a) The circuit A is a buffer, and the gain is 1. The input resistance is the resistance of the OpAmp non-inverting input pin, which is ideally infinite, plus the effect of the negative feedback.
The circuit B can be thought of as follows:



For the superimposition effect, the output V_{OUT} is:



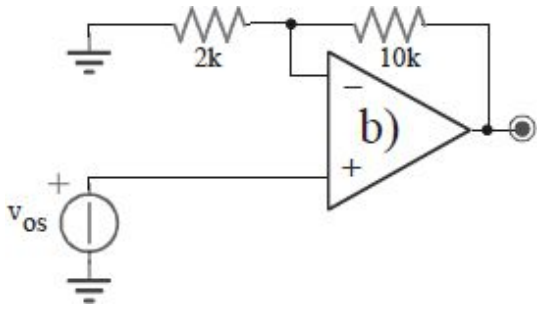
In this case, the gain is 1. It is possible to reach the same result by noticing that there is no voltage drop across the OpAmp inputs, thus there is no current flowing through the 2kΩ resistance. For this reason, the stage is effectively a buffer.

The impedance seen by the input node is reduced by (1-G_{loop}*). Notice that G_{loop}* is the gain from the positive input pin and the negative input pin of the OpAmp. It is NOT the G_{loop} of the stage, which is equal to A(s) of the OpAmp multiplied by β=2kΩ/(2kΩ+10kΩ)=0.167. At first, the input impedance will be reduced (because it is divided by G_{loop}*). Actually, the feedback is positive, although it is not enough to make the circuit oscillating. If the OpAmp were ideal, G_{loop}*=+1 would hold. The actual value is G_{loop}*≈+0.999. Thus, the impedance seen by the input is the same without feedback (2kΩ+10kΩ=12kΩ), which is reduced by the factor 1-(+0.999)=0.001, i.e. incremented by a factor of 1000 (it becomes 12MΩ).

Circuit C is a buffer because the 2kΩ resistance is shorted between the two input pins and the virtual ground. The 10kΩ resistance is in series with the impedance treated in the preceding step, i.e. 2kΩ/(1-G_{loop}*). In the ideal case, it is infinite, otherwise with G_{loop}*≈+0.999, it is around 2MΩ

Circuit D is a buffer as before. The input impedance is ideally infinite.

b) For all cases, we can model the offset as a voltage generator connected to the non-inverting input. In the circuit A (buffer), we have V_{OUT}=V_{OS}=5mV. The circuit B is a normal non-inverting configuration.



Therefore, even though the stage is a buffer for the signal, for the offset, it has a gain of 5, determining: $V_{OUT} = V_{OS} \cdot \left(-\frac{10k\Omega}{2k\Omega}\right) = -5V_{OS} = -25mV$.

For the circuit C, the offset is applied between the two OpAmp input pins, thus there is a signal current on the 2kΩ resistance. This current determines

an output voltage equal to $V_{os}/2k\Omega \cdot (2k\Omega+10k\Omega)=6 \cdot V_{os}=30mV$. Instead, circuit D is a buffer for the offset, too. In fact, the capacitance does not allow the current to flow and $V_{out}=V_{os}=5mV$.

c) For circuit A, there is resistance on neither the non-inverting branch nor the feedback branch, thus there are not bias contributions.

For circuit B, the non-inverting pin bias current gives no contribution while the inverting pin current flows through the 10kΩ resistance and is absorbed by the OpAmp, which causes $V_{out}=i_B \cdot 10k\Omega=1mV$.

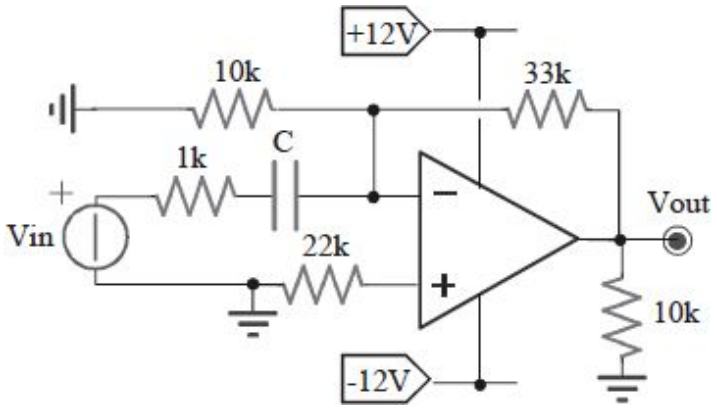
In circuit C, the output voltage is equal to $100nA \cdot 10k\Omega=1mV$. In fact, the bias current flows through the 10kΩ resistance because the 2kΩ resistance is between the two equipotential nodes. Finally, the current i_B for circuit D has the same behavior as the current in circuit C, which causes an error of 1mV on the output.

Ex. 12

The OpAmp has $V_{OS}=3mV$, $I_B=200nA$, and $I_{out}=20mA$.

a) Calculate the transfer function $V_{out}/V_{in}(s)$.

b) Determine the output static error, when $V_{in}=0V$.



a) The 10kΩ resistance seems to form a partition on the input signal. Instead, if the stage has good feedback, the input voltage on the inverting input is the virtual ground; therefore, the 10kΩ resistance is between the two ground pins, and, in this branch, there is not any current flowing. In fact, the stage's inverting gain is:

independently of the presence of the 10kΩ resistance. There is a zero in DC and a pole at $1/2\pi 1k\Omega \cdot C$, and the stage is therefore an approximate derivator.

In order to make more precise calculations, we should compute the stage G_{loop}, which is the product of A(s) of the OpAmp and the feedback block β(s). The 10kΩ resistance intervenes by lowering β(s) and slightly reducing G_{loop}. However, if A(s) is very high, then $G_{loop}>>1$. Thus, the stage has good feedback, and the behavior can be considered ideal.

b) The static error sees an open capacitor C. It would see a buffer stage if the 10kΩ resistance were not there. Instead, the offset voltage (the voltage source in series with the non-inverting input pin) undergoes an amplification of

G=1+33kΩ/10kΩ=4.3. Therefore, we have an output of 3mV·4.3=13mV. The I_{B+} flowing through the 22kΩ resistance undergoes an amplification of 4.3, too. For the current, we have:

$$V_{out} = I_{B+} \cdot 22k\Omega \cdot \left(1 + \frac{33k\Omega}{10k\Omega}\right) - (I_{B-} \cdot 33k\Omega) = 19mV - 6.6mV = 12.3mV \cdot$$

Ex. 13

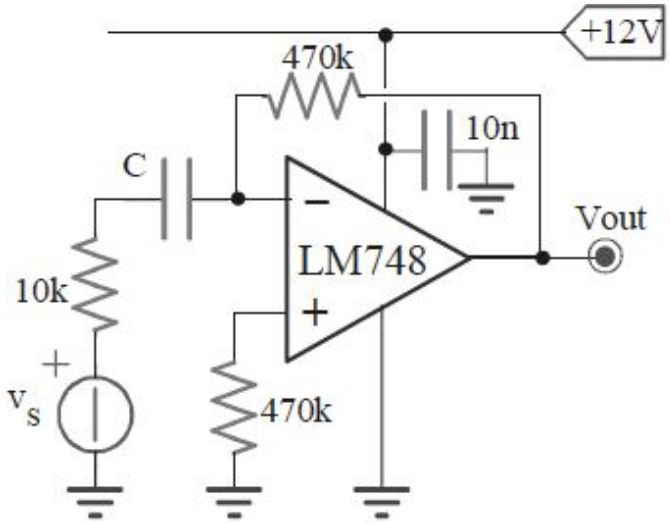
I_B=500nA, V_{OS}=3mV, I_{OS}=100nA, A_0=100dB, and *I_{outmax}=1mA.*

a) Calculate all the output static errors.

b) Size C in order to have *20Hz* within the bandwidth and draw the Bode diagram of $|v_{out}/v_s(f)|$.

c) Compute the range of input v_s for a stage "working well".

d) We want a bandwidth of *1MHz*. Compute the required *GBWP* for the OpAmp.



a) Thanks to the insertion of the 470kΩ resistance connected to the non-inverting input pin, the static error due to the bias currents is zero because the impedance values seen from the input pins (inverting and non-inverting) have the same value (ignoring the degrees of tolerance). The offset voltage is transferred to the output with unity gain.

The offset current determines the following error:

In the worst case, this voltage will be added to the offset voltage, and the total error is 47mV+3mV=±50mV.

b) We should ensure a correct transmission for the 20Hz signal, thus the bandwidth must be at least a decade less than that, i.e. it must be around 2Hz. The pole is given by the product 10kΩ·C. In fact, the transfer function is:

$$G = -\frac{470k\Omega}{10k\Omega + 1/sC} = -\frac{470k\Omega \cdot sC}{1 + 10k\Omega \cdot sC}$$

which represents a zero in DC and a pole at the frequency $1/2\pi 10k\Omega \cdot C$. Thus, the capacitance is $C=(2\pi \cdot 10000\Omega \cdot 2Hz)^{-1}=7.96\mu F$ and a good value is 10μF. For higher frequencies, the gain is –R2/R1=-47.

c) Assuming that the OpAmp is rail-to-rail (it works in the linear regime for the whole dynamic), the output range is 0÷12V. Remembering that the amplification is INVERTING equal to -47, we deduce:

$V_{OUT,min}=0 \rightarrow V_{IN,max}=0/(-47)=0V$

$V_{OUT,max}=12V \rightarrow V_{IN,min}=12/(-47)=-255mV$

Moreover, another constraint is due to the I_{outmax} that the OpAmp can supply or sink. This limit is 1mA, and then |Vin| will be less than $I_{outmax} \cdot 10k\Omega=10V$, well above of the preceding limit.

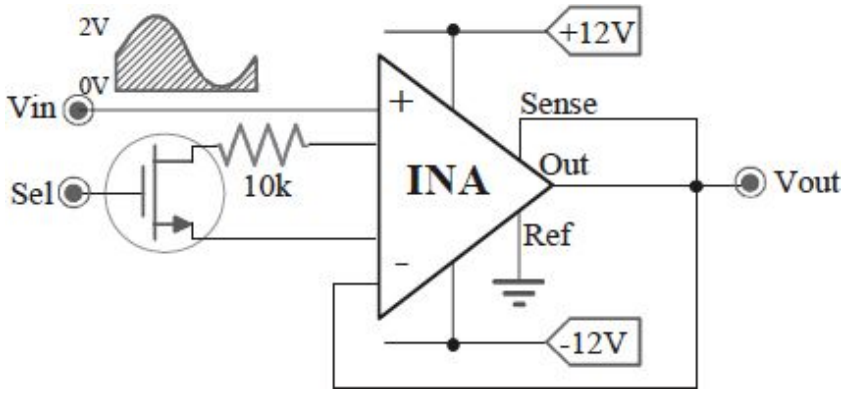
d) Remembering that for a compensated OpAmp, the Gain-BandWidth Product is equal to the frequency at which A(s) is 1, for a band of 1MHz and with a 1/β (usually equal to the non-inverting gain) of 48, we need an OpAmp with GBWP=1MHz·48=48MHz, as shown in the graph above.

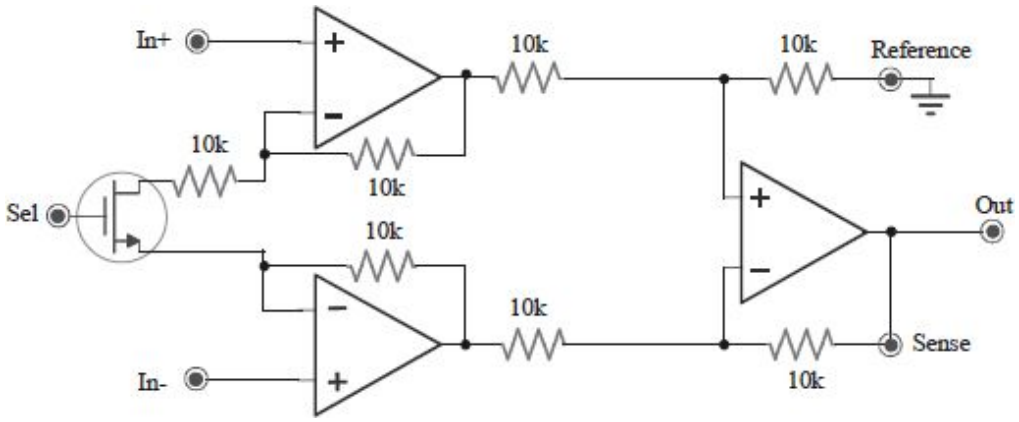
Ex. 14

The circuit uses an INA (internal resistance values of *10kΩ*).

a) Calculate V_{out}/V_{in} with **Sel**=*low*.

b) Calculate the **Sel** *high* level which allows the correct closing of the MOS (V_T=*2V*).





a) The circuit of concern uses an Instrumentation Amplifier (INA) with a feedback between the output and the inverting input pin. To better understand the behavior of the circuit, consider the internal scheme of the INA.

Whether the internal resistance values of 10kΩ is connected or not depends on the MOS state (on or off) and therefore on the *Sel* value. Assuming a common mode signal centered on 0V, when Sel=Low, the MOS does not

conduct, and the two OpAmps are two decoupled buffers. When Sel=High, the MOS conducts (ideally the resistive channel between the drain and the source has negligible resistance with respect to the 10kΩ resistance), and the differential gain is that of a classical INA stage:
The third OpAmp is a differential amplifier. This device reads the differential input voltage and gives it without any gain as output (all the resistors have the same value).
The commercial INAs have two extra pins. The *Reference* pin (usually connected to the ground) can be connected to any voltage setting the output baseline. The *Sense* pin represents the feedback branch of the differential amplifier and is connected to Out and as close as possible to the load so as to avoid unwanted voltage drop across connection wires.
With respect to the circuit under scrutiny, we can write the following equations:

$$V^- = Vout$$

$$\frac{Vout}{V^+ - V^-} = \left(1 + \frac{2 \cdot 10\,k\Omega}{\infty}\right) = 1$$

$$Vout = 1 \cdot (V^+ - V^-) = Vin - Vout \Rightarrow Vout = \frac{Vin}{2}$$

The gain is ½, i.e. the stage does not amplify, but is a 1:2 divider. Why not use a common resistive divider with two resistors? Because, with this circuit, the gain is exactly ½ and does NOT depend on the resistors' degrees of tolerance, the thermal drift, or devices matching. The question spontaneously arises within the reader: "so, how much do the non-idealities of the INA weight?" The answer of the author is: "these are your exercises! Do them". Notice that the author has been spending many weekends to write this book while the reader is skiing or going to the seaside.
b) In order to be well closed, the MOS must work in the triode mode with a low R_{ON}. Such a condition is satisfied if $V_{DS}<V_{GS}-V_T$, i.e. for $V_{GD}>V_T$. If R_{ON} is effectively negligible, we have $V_D \approx V_{in}$ and $V_S \approx V_{out}$. However, in this case, V_{out} is not equal to $V_{in}/2$ because when the MOS is closed, the INA gain is 3. With similar calculations, it is possible to obtain $Vout = \frac{3}{4} \cdot Vin$. Then, $V_G > V_T + V_D = V_T + Vin = 2V + 2V = 4V$, considering the worst case, i.e. the

drain has higher potential (V_{in}) than the source (V_{out}) and the maximum V_{in} is equal to 2V.

Ex. 15

The INA is made with ideal OpAmps.

a) Draw the frequency response $(v_{pos}-v_{neg})/(IN_{pos}-IN_{neg})$ for the first stage.

b) Compute the static error on v_{out} due to $I_B=100pA$ and $V_{os}=2mV$.

a) In DC, we can consider all the capacitors open. In such a case, both OpAmps are buffers, and the transfer function is equal to 1. At medium frequencies, C_{g1} must be shorted (i.e. the pole must have a frequency less than the minimum frequency of the bandwidth) for a gain equal to:

Indeed, the high frequencies have to be attenuated. Thus, when C_a and C_b are shorted, the first stage can have a double buffer-behavior and therefore a gain of 1. The 470µF capacitance introduces a pole with a frequency $f_{low}=1/2\pi\cdot R_g\cdot C_g=0.072Hz$. The feedback capacitance of the OpAmps of the first stage introduces a pole at $f_{high}=1/2\pi\cdot 1M\Omega\cdot 620pF=256Hz$. It is possible to graphically extract the two zeros: the lower at a frequency $f_{Zlow}=f_{low}/426=170\mu Hz$ and the higher at a frequency $f_{Zhigh}=f_{high}\cdot 426=110kHz$.

b) The first stage OpAmps' bias currents do not give any contribution because they are equal to each other and cause a common mode voltage that is rejected by the INA's second stage. Concerning the second stage OpAmp, as the resistance values seen by the two input pins are equal, the error due to bias currents is zero for the second stage, too.

As the offset voltages' signs are unknown (opposite to the bias currents that are always sign concordant), in the worst-case, the contributions of the two first stage OpAmps will be added on top of each other. The DC transfer is equal to 1, and thus on the INA output, we will have the worst case error of $2 \cdot V_{OS}$ due to the first stage OpAmp offsets. The second stage offset voltage will be multiplied to attain the required gain:

Ex. 16

The OpAmp has $A_0=100dB$, $V_{OS}=5mV$, and $I_B=500nA$. At $0°C$, the thermo-resistance is 100Ω, and the variation is $0.5\Omega/°C$. At $500°C$, we want $V_{out}=10V$.

a) Choose R and calculate the output variation for $T=0\div500°C$.

b) Size $I_{out,max}$ of the OpAmp.

c) Modify the circuit to have a threshold around $100°C$ with a hysteresis of $\pm10°C$.

a) The circuit shown in figure is a simple system that allows temperature measurements through a thermo-resistor: a 5mA current generator injects a constant current into the sensor which modifies its resistance according to the temperature (and the voltage across itself, thanks to the Ohm's law).

The thermo-resistance value at 0°C is 100Ω and grows 0.5Ω for every Celsius degree. Sometimes, instead of the resistance increment expressed in Ω/°C, the sensor thermal coefficient is provided. This coefficient is normally marked with α [$°C^{-1}$]. In this case, it is possible to calculate the resistance at certain temperature with the following formula: $R_T = R_0(1 + \alpha\Delta T)$,

where R_0 is the resistance at reference temperature T_0 (for example 0°C) and ΔT is the thermal increment with respect to T_0.

At 500°C, the thermo-resistance is equal to R_T=100Ω+0.5Ω/°C·500°C=350Ω. The voltage of the OpAmp non-inverting input pin is equal to V^+=350Ω·5mA=1.75V. The OpAmp has negative feedback in a non-inverting configuration. In order for *Out* to have a voltage of +10V, it will be necessary to size the stage as follows:

Such a value is commercially available. At 0°C, the output voltage is: V_{out}=5mA·100Ω·G=2.85V.

The OpAmp offset voltage introduces a “readout” error of 2°C; in fact, the 5mV offset determines an output signal with the same amplitude as a 1Ω variation. The effect of the bias currents is lower than the one due to the offset voltage; in fact, I_{B+} introduces the maximum error, which is 500nA·350Ω=175μV at 500°C. Instead, I_{B-} causes an error of 500nA·4.7kΩ=2.35mV.

b) Assuming that the chosen OpAmp is not rail-to-rail, i.e. it cannot provide output values equal to the supply voltages (positive and negative); we can assume that 10V is the output voltage limit that guarantees a linear operation of the device.

Thus, the maximum current that the OpAmp must provide is $I_{out,max}=V_{out,max}/R_{tot}$ (without load), where the resistance is the one seen between the terminal and the ground, i.e. $R_{tot}=4.7k+1k=5.7k\Omega$. Thus, it results that $I_{out,max}=1.75mA$.

c) A possible implementation is the following:

As you can notice, positive feedback is used to implement the hysteresis. R_2 is used to translate the threshold level up to 0.75V, corresponding to an input of 100°C. In order to minimize the average consumption of the circuit, it is possible to choose a very high R_2, for example 100kΩ. For a hysteresis of ±10°C, corresponding to 25mV on the inverting pin, R_1 must be chosen so that: , from which R_1= 395kΩ can be obtained. The closest commercial value is 390 kΩ.

Ex. 17

The OpAmp has $V_{OS}=3mV$, $I_B=200nA$, and $I_{offset}=10nA$.

a) Draw the modulus Bode diagram for the circuit transfer function V_{out}/V_{in}.

b) Determine the output static error when $V_{in}=+5V$.

a)

b) V_{OS}=0.3V; Vi_{BIAS}=201mV; Vi_{OS}=10mV.

2.12. Non Linear Circuits with OpAmps

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.

2.12.1 Precision rectifier (Super Diode)

With a diode in feedback, as shown in Fig. 2.35, the threshold is reduced to the value V_γ/A, where A is the open loop OpAmp gain. It is possible to obtain rectifying thresholds in the order of micro volts, thanks to equal potentials of the OpAmp input terminals (the concept of virtual ground).

Considering a differential voltage $v_d = v_i - v_L \neq 0$, we have $v_0 = v_L + V_\gamma = Av_d$, from which: $v_L = \frac{A}{1+A}\left(v_i - \frac{V_\gamma}{A}\right) \cong v_i - \frac{V_\gamma}{A}$ can be obtained. The above equation demonstrates that the diode threshold is reduced by the factor A.

We can see that if v_i is positive and the diode is conductive, then the OpAmp behaves as a voltage buffer and thus $v_L = v_i$. The current flows from the OpAmp output pin to the load R_L through the conductive diode. On the other hand, if we assume that the diode is conductive also with a negative v_i, then the voltage across R_L should be negative, thus the OpAmp should recall the load current. However, this is not allowed by the diode which is inversely polarized. Therefore, for negative v_i, the diode opens the OpAmp feedback loop, and the output is V_L=0V. The waveforms of the super diode are shown in Fig. 2.35.

If we consider now a capacitance C in parallel to R_L, we will obtain a *peak detector* (Fig. 2.36) able to store the maximum value of the input signal. In fact, the load voltage v_L will follow the input v_i only as long as v_i is greater than v_L (the diode D is conductive, and it charges C) while if $v_L > v_i$ D is off, the capacitance is discharged through the resistance R_L with a time constant $R_L C$. The OpAmp is saturated with the negative output (because the feedback is interrupted). If one wants to reduce this voltage to $-V_\gamma$, it will be enough to insert the diode D_1, as shown in Fig. 2.36.

Fig. 2.35: Precision Diode, using an OpAmp.

Fig. 2.36: Peak detector.

Fig. 2.37: Half-wave rectifier with an inverting configuration OpAmp.

Fig. 2.38: Half-wave rectifier with a non-inverting configuration OpAmp.

2.12.2 Half-wave rectifier

In the circuit shown in Fig. 2.37, the OpAmp is used in the inverting configuration. When v_i is positive, the output v_o is negative; the diode D_1 is in interdiction (as the feedback through R_f) while D_2 conducts (the feedback through D_2 introduces a virtual ground on the inverting input, and the OpAmp cannot saturate) and thus v_L=0. On the other side, when v_i is negative, the output v_o is positive; the diode D_2 is off while D_1 conducts (R_f is connected in the feedback, and the OpAmp works as an inverter) and results $v_L=-(R_f/R_i)\cdot v_i>0$.

Fig. 2.39: Full-wave rectifier.

Fig. 2.40: Full-wave rectifier for $v_i>0$.

Consider now the circuit in Fig. 2.38, in which the input is connected to the non-inverting input pin. When v_i is positive, the output v_o is positive, too; D_2 is in interdiction while D_1 is conductive, and thus we have $v_L=\left(1+\frac{R_f}{R_i}\right)v_i$

When v_i is negative, the output v_o is negative; D_1 is in interdiction, D_2 conducts, and the OpAmp works as a voltage follower with $v_2 = v_i = v_o+0.7V$. Thus, the voltage on the load is: which coincides with v_i if $R_L>>R_f$.

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.

2.12.3 Full-wave rectifier

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). The circuit is made with two cascaded inverters (as exemplified in Fig. 2.40), and the output voltage is: (positive for $v_i > 0$).

When v_i is negative, and the output v_o is consequently positive, the diode D_1 is in interdiction while D_2 conducts (as shown in Fig. 2.41). The terminals of the OpAmp A2 have the same potential, thus A1 (that works as an inverter) presents feedback resistance equal to $(R_{f1}+R_{i2})//R=2/3R$, and therefore we have:

Because the OpAmp A2 does not sink any current on the input, we have The output is (positive for $v_i<0$).

Thus, the output voltage is always positive and then "rectified" while the gain is controlled by the sole R_{i1} resistance.

Fig. 2.41: Full-wave rectifier for $v_i<0$.

Fig. 2.42: Schematic diagram of a comparator and the output characteristic.

Fig. 2.43: An open loop (without feedback) OpAmp used as comparator.

2.13. COMPARATORS

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 shown in Fig. 2.44. The two Zener clamp the output $v_{\grave{o}}$ to the value $V_o = \pm (V_Z + V_\gamma)$, where V_γ is around 0.7 V and is the voltage drop on the direct biased Zener while the series R on the output limits the current flowing through the Zener diodes.

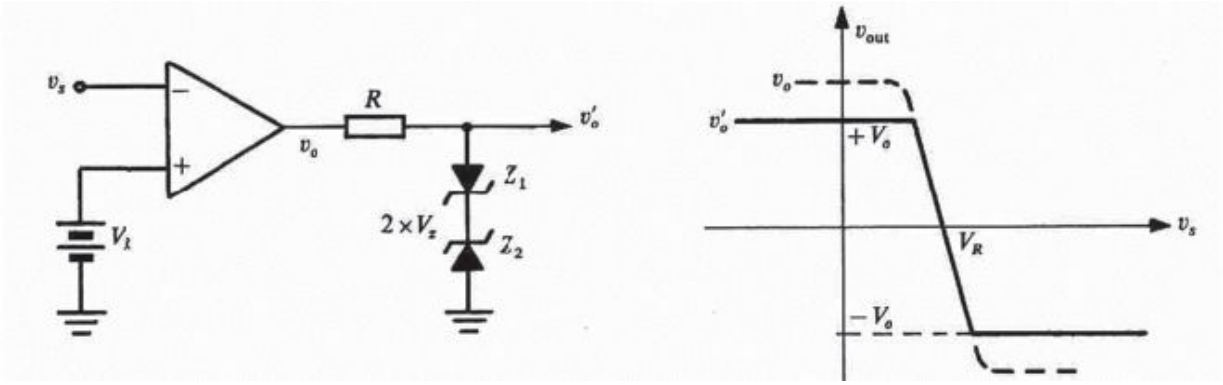


Fig. 2.44: Inverting comparator with output voltage limited by Zener diodes.

Fig. 2.45: Output voltage limited comparator with feedback Zener diode.

Another system to limit the output voltage is the introduction of a negative feedback network with non-linear elements. In this case, the resistor R_i is needed on the inverting input to bring the signal v_s independent of the reference V_R, as shown in Fig. 2.45. The current in the limiting circuit is:

When $v_s > V_R$, the Zener is directly biased, and the output voltage is $V_R - V_\gamma$ while, when $v_s < V_R$, the Zener will work in the Zener region, and the output will be $V_R + V_Z$.

2.13.1 Zero-crossing detector

If the voltage reference is $V_r = 0$, the output will change whenever the input passes the zero (zero-crossing detector), as shown in Fig. 2.46. When the input signal is $v_s > 0$, the current $i_i = v_s / R_i$ flows through the feedback network. The Zener diode Z_1 is biased directly while Z_2 is biased indirectly. Because the inverting terminal is connected to a virtual ground, the output voltage will be $V_{OL} = -(V_Z + V_\gamma)$. When the input signal is $v_s < 0$, the current i_i is inverted, and, obviously, we will have $V_{OH} = +(V_Z + V_\gamma)$.

Fig. 2.46: Zero-crossing detector.

Fig. 2.47: Level detector.

Observe how the Zener diodes fix the output value preventing the OpAmp saturation, increasing the switching speed and the speed of response for high frequency input signals.

2.13.2 Level detector

If the V_r reference voltage is very far from zero, it is possible to detect when the input signal v_s reaches a preset voltage value. Considering the circuit shown in Fig. 2.47, we observe how, thanks to the virtual ground of the inverting input, the current on the feedback branch is equal to:

If , the current $i_f > 0$ flows through the output, Z_1 is direct biased while Z_2 inverse biased. Therefore, the output voltage will be $v_o = V_{OL} = -(V_Z + V_\gamma)$. If, instead, $v_s <$, $i_f < 0$ holds, and the output voltage v_o is thus inverted with respect to the preceding case and equal to $V_{OH} = +(V_Z + V_\gamma)$.

The reference value at which the comparator switching occurs is equal to and can be varied by acting on the ratio R_i/R_r without forgetting that R_i determines also the input resistance for the signal.

2.13.3 Window comparator

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).

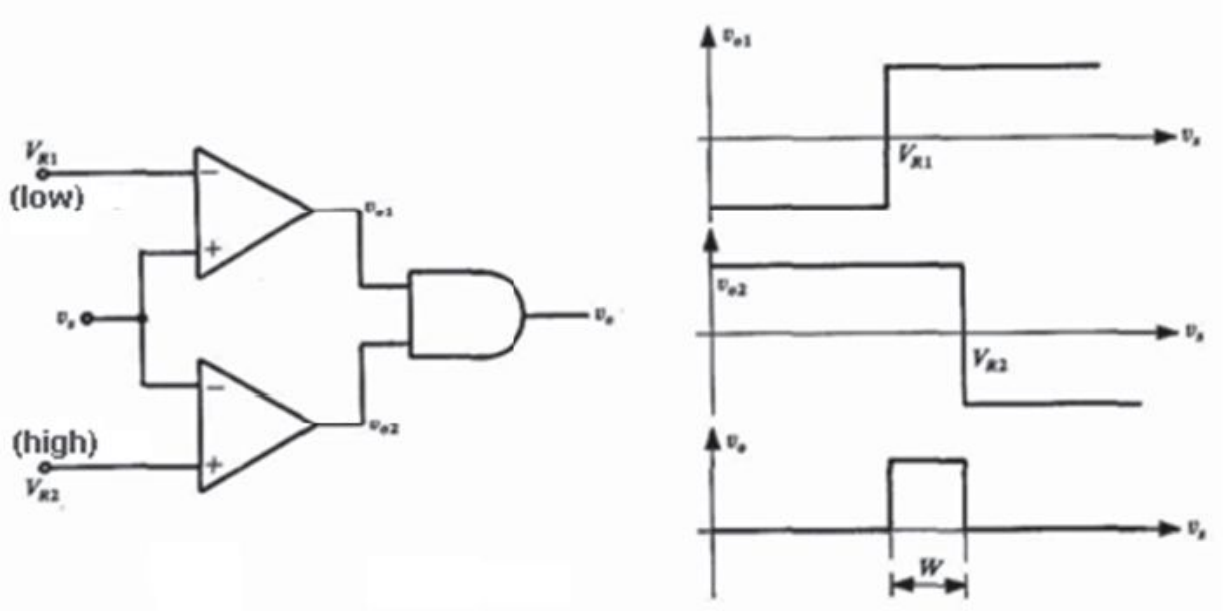


Fig. 2.48: Window comparator and its characteristic.

Fig. 2.49: Non-inverting comparator with hysteresis.

Fig. 2.50: Inverting Schmitt trigger (when the input is above the high threshold, the output commutates towards the low value).

2.14. SCHMITT TRIGGER WITH HYSTERESIS

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.

2.14.1 Inverting Schmitt Trigger

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):

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:

Fig. 2.51: Non-inverting Schmitt Trigger (when the input exceeds the upper threshold, the output switch high).

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:

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

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 $R_r << R_f$. In this case, we have:

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:

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

2.14.2 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:

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:

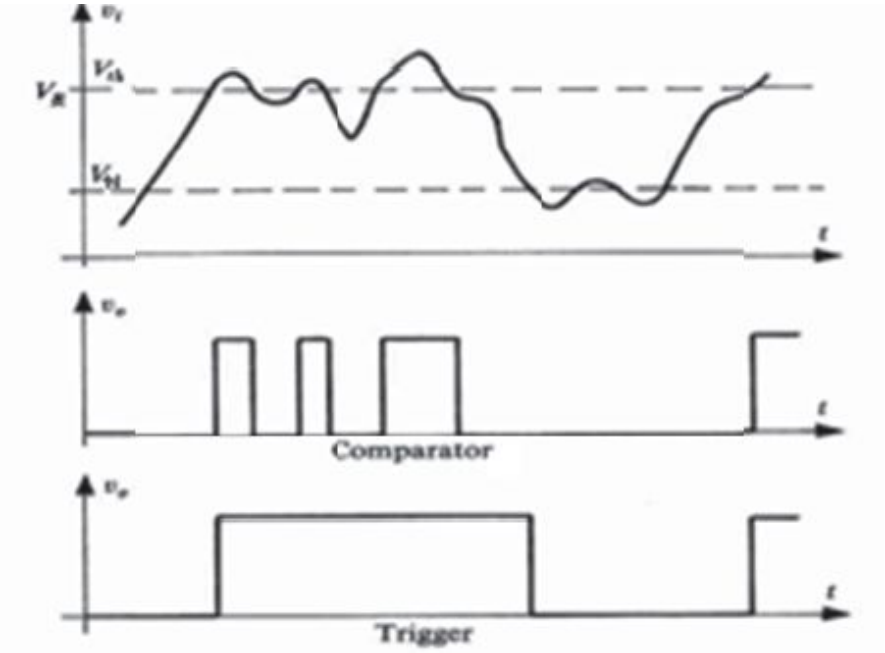


Fig. 2.52: Input (above) and output signal for a comparator (center) and for a trigger (below).

and thus 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 and thus with The hysteresis is:

2.14.3 Comparison between Trigger and Comparator

As seen so far, the comparators with and without hysteresis have different characteristics of operations that make them suitable for particular jobs.

The comparator switches every time the input voltage v_i is greater or lower than the reference voltage V_R. For this reason, the comparator is very sensitive to the voltage noise. Another problem is the uncertainty of the commutation, as pointed out in Fig. 2.52.

On the opposite side, the Schmitt Trigger provides a stable output when the input voltage v_i exceeds the corresponding threshold (rising or falling), because the threshold changes when the output changes. In this way, the system is insensitive to the noise with a peak-to-peak value less than the hysteresis V_h (last diagram depicted in Fig. 2.52).

2.15. Waveform Generators

The waveform generators can be classified according to the base of the waveform that they generate. Sinusoid and square-wave are the most common waveforms. Other kinds of waveform are the triangular and the saw-tooth. For a sinusoidal waveform, the generators are named the oscillator while for a square-wave, the circuit is named the *relaxation oscillator* or the *astable multivibrator.*

Fig. 2.53: Loose (βA<1), self-supporting (βA=1) and divergent (βA>1) positive feedback.

2.15.1 Sinusoidal Oscillators

Sinusoidal oscillators are circuits that generate a sinusoidal voltage. Consider the linear amplifier with a gain A, which is shown in Fig. 2.53, with an external signal v_i applied as an input (at node 1). The amplifier output signal v_o is the input signal v_i, but can be different in amplitude and phase

delay from the latter. With a feedback network β, it is possible to bring the signal (v_f=βv$_o$) back to the node 2, which is proportional to the output v_o=Av_i. The loop gain is equal to G_{loop}=v_f /v_i=βA.

While in the circuits employing OpAmps that were seen before, the most used is the negative feedback (βA<0); in such a case, we will use the positive one (βA>0). The different types of positive feedback are shown in Fig. 2.53. If the circuit is designed to satisfy the condition G_{loop}=βA=1 (*Barkhausen criterion*), v_f=v_i will hold. Then, it is possible to remove the external signal v$_i$ and replace it with the feedback signal v_f (closing the loop and connecting the node 1 and 2): the circuit is self-sustaining.

The stages' gains depend on the frequency because of the reactive components. The Barkhausen criterion can be translated into two conditions: $|G_{loop}|$=1 and φ_{Gloop}=0°, i.e. the signal must travel through the loop and go back to the starting point (the node 2) with the starting phase and the amplitude (present at node 1). Because G_{loop}(2πf) varies with frequency f, the conditions can both be satisfied for a definite frequency f$_o$, which is the oscillating frequency of the circuit. The triggering of the oscillation is possible, thanks to the presence of the thermal noise of the devices within the circuit.

In actual oscillators, the gain loop βA must be greater than 1 to ignite the oscillations while at steady-state, to maintain the oscillations, the loop gain must be brought back to unity when the output reaches the desired amplitude. This is possible if, once the oscillations are ignited, the value of A or β is reduced; the automatic reduction is possible, thanks to the intrinsic non-linearities of the amplifier A (the saturation or the interdiction of transistors), or introduced by non-linear networks of the feedback-block β.

2.15.2 Phase-shift oscillator

Analyzing the phase-shift oscillator in Fig. 2.54, it is possible to highlight both the amplifier A (comprised of an OpAmp) and the feedback block β (comprised of three identical RC networks). In the hypothesis in which the amplifier A is in inverting configuration (phase-shift of 180°), the feedback network must produce an additional phase shift of 180° and attenuation in order to give rise to oscillation.

Fig. 2.54: Phase-shift oscillator.

An RC network introduces a progressive phase-shift starting from 90° (low frequency) up to 0° (high frequency). It is clear that we need a cascade of at least three RC cells to obtain a phase delay of 180°. Having assigned the values of R and C, for only one value of the frequency, the phase delay will be exactly 180°.

With a feedback network with three identical RC cells, the feedback gain is:

where $\alpha=1/(2\pi fRC)$. The phase delay between v_o and v_f is 180° only if the imaginary part of the denominator is null, i.e. for $\alpha^2=6$ and thus at the frequency $f_o=1/(2\pi RC\sqrt{6})$., we have a feedback of $\beta=1/29$. Therefore, in order to have oscillations at this frequency, satisfying the Barkhausen criterion, the amplifier should have a gain $|A|=R_f/R_i=29$, simply attained by the right values of R_f and R_i.

Practically, we have sinusoidal oscillations with low distortion only if $|G_{loop}|$ is slightly higher than 1, and adjusting the gain A is possible by using, for example, a trimmer instead of R_f.

2.15.3 Wien bridge oscillator

Thanks to the high frequency stability, the Wien bridge oscillator is used in circuits operating between few Hz and more than 1MHz, where the latter limit is due to the OpAmp slew rate. The scheme is shown in Fig. 2.55, on the left. The bridge is supplied by the OpAmp output voltage v_o, and the OpAmp is supplied by the bridge voltages v_1 and v_2. Highlighting the feedback β network (positive) and the OpAmp A with negative feedback (Fig. 2.55, on the right), we have the feedback factor:

where $\gamma=2\pi f\cdot RC$ if $R=R_1=R_2$ and $C=C_1=C_2$. The amplifier is non-inverting, thus the phase shift between v_o and v_f must be null, and this is verified only if the imaginary part of β is null, i.e. for $\gamma^2=1$ and for the frequency

$f_o=1/(2\pi RC)$, where $\beta=1/3$ holds. Therefore, in order to have oscillations at this frequency, the amplifier must have a gain A=3 to satisfy the condition: $R_f=2R_i$.

Fig. 2.55: Wien bridge oscillator.

The ignition of the oscillations is possible only if $|\beta A|$ is slightly larger than 1, and thus $R_f>2R_i$ must initially hold. Practically, it is possible to use non-linear elements: for example, thermistors that vary the resistance according to the temperature, NTC type (Negative Temperature Coefficient) instead of R_f, or PTC (Positive Temperature Coefficient) instead of R_i. In such a case, the condition $|\beta A|>1$ is initially satisfied, but once the oscillation is reached, we automatically obtain $|\beta A|=1$. In fact, while the output oscillation amplitude increases, the power dissipation on the PTC increases (increasing also the resistive value), reducing the gain A and definitely reducing the oscillation amplitude.

2.15.4 Three-points oscillators

The high frequency oscillators use a parallel LC tank circuit in the feedback network. It oscillates with its natural frequency because the energy stored in the capacitance is discharged on the inductor. The inductor then returns the energy to the capacitance, and so on. The amplifier is needed for compensating the losses due to the stray series resistance of inductor and the stray parallel resistance of capacitance, which try to dump the oscillations.

Fig. 2.56: Prototype structure of a three points oscillator.

The general scheme of an LC oscillator, also known as *three points*, is shown in Fig. 2.56. The tank is built up with three reactive elements (Z=jX, where $X=\omega L$ for the inductor and $X=-1/(\omega C)$ for the capacitance) which must have infinite impedance at the resonance natural frequency (null admittance) for the signal v_o that must assume the maximum value.

for which $X+X_f+X_i=0$ results. We can deduce that the three reactive elements cannot be of the same type. Calculating the loop gain, we have:
Because $X_f+X_i=-X$, to satisfy the Barkhausen criterion, we have from which we can deduce that, as the gain A is negative, the reactance values X_i and X must have the same sign, i.e. both must be either capacitive or inductive. Finally, because $X+X_i=-X_f$, we can conclude that the reactance values X and X_i must be capacitive and that X_f must be inductive, or vice-versa. The oscillation frequency is determined by the resonance condition of the tank circuit that results to be purely resistive ($X+X_f+X_i=0$) while the oscillation ignition is possible only if $|G_{loop}|$ is slightly larger than 1. An output voltage with a higher value without distortion is obtained by experimentally changing the feedback factor β.

Consider now the two different schemes shown in Fig. 2.57. For the Colpitts (on the left) and for the Hartley (on the right), the oscillation frequency can be obtained from:

$$-\frac{1}{\omega C}-\frac{1}{\omega C_i}+\omega L_f=0 \qquad \omega L+\omega L_i-\frac{1}{\omega C_f}=0$$

and are equal to and respectively. The oscillation ignition occurs for and thus for Therefore, we have: and , neglecting any mutual coupling between L_i and L.

Fig. 2.57: Colpitts (left) and Hartley (right) oscillators.

Fig. 2.58: Rectangular signal generator (astable multivibrator).

2.15.5 Astable multivibrator

The scheme of a rectangular wave generator is shown in Fig. 2.58. It consists of an OpAmp used as a hysteresis comparator and of an integrator built up with a capacitor C and a resistor R. The output is limited by using

Zener diodes connected opposite to one another. The resistance R' is introduced to limit the output current to a value , where V_Z is the Zener voltage assumed equal for both devices, and V_{uM} is the OpAmp saturation output voltage.

The oscillation frequency can be extracted starting from the charging voltage of the capacitance C, as shown in Fig. 2.59. Assuming that for t= 0, the output signal can have a value of $+V_Z$, the voltage is brought back to the non-inverting terminal through the resistive divider R_1, R2 equal to:

$$V_Z \frac{R_2}{R_1 + R_2} = H \cdot V_Z \qquad \text{where } H = \frac{R_2}{R_1 + R_2}.$$

Assuming that the OpAmp is a comparator, the voltage v_c across the capacitance C must be less than $H \cdot V_Z$ and V_Z. The voltage difference across the resistance R is due to the current that charges the capacitance, increasing the voltage v_c with an exponential trend through the value V_Z. When the voltage v_c reaches the comparator threshold voltage, which is the level $H \cdot V_Z$ on the non-inverting terminal, we have the output commutation, and the value abruptly changes from $+V_Z$ to $-V_Z$. Correspondingly, also the reference level on the non-inverting terminal switches from the value $H \cdot V_Z$ to the value $-H \cdot V_Z$. On this condition, the capacitance C, previously charged to the value $H \cdot V_Z$, stops charging and begins the exponential discharge through the value $-V_Z$. When the lower threshold $-H \cdot V_Z$ is reached, a new output commutation from the value $-V_Z$ to the value $+V_Z$ occurs, and so on. We can deduce that the time T_1, equal to the half-period of the output signal, is the time necessary for charging v_c from the value $-H \cdot V_Z$ to the value $+H \cdot V_Z$. We can demonstrate that:

The form for T_1 shows that the signal frequency depends on the time constant and on the division ratio H. It is correct that the value of H (between 0 and 1) is not so close to 1 in order to prevent the threshold level from being too close to the asymptotic value of the

Fig. 2.59: Voltages vs. time for the astable multivibrator

exponential signal ($+V_Z$ during the charge phase). In that case, the signal would cross the threshold value with a reduced slope, determining a great switching inaccuracy and thus also an inaccuracy in the oscillation frequency.

Assuming that the waveform is symmetrical ($T_1=T_2$), the oscillation frequency is equal to:

If H=1/3, we have $T_1=\tau\cdot\ln 2=0.693\cdot\tau$ and then $f=1/1.4\tau$. In order to change with continuity the oscillation frequency, we can make the value of H variable by replacing R_1 and R_2 with a potentiometer in which the slider is connected to the non-inverting OpAmp input.

2.15.6 Rectangular pulse generator (monostable multivibrator)

A circuit able to provide rectangular pulses with well-defined amplitude and duration can be obtained from the rectangular waveform generator previously analyzed by adding a diode D in parallel to the capacitance C, as shown in Fig. 2.60. This circuit is named the *monostable multivibrator* because it works at rest in a stable condition. Correspondingly, in this stable state, the voltage across the capacitance C is equal to the diode threshold voltage $V_\gamma=0.7V$. This avoids further charging of the capacitance while the output voltage has the maximum level V_Z, and the non-inverting terminal voltage, equal to $H\cdot V_Z$, maintains the output equal to the high level because $H\cdot V_Z>V_\gamma$.

To allow the commutation of the circuit from the stable to the unstable state, a negative pulse is externally applied with adequate amplitude through the network formed by R_3, C_1, and the diode D_1. The negative pulse on the non-inverting terminal determines an abrupt reduction of the comparator threshold voltage below the value V_γ present at the inverting terminal. Consequently, the output voltage commutation happens, and the value of this terminal quickly passes from the value V_Z to the value $-V_Z$, as shown in Fig. 2.61.

At this point, also the non-inverting threshold voltage assumes the value $-H\cdot V_Z$. Hence, the capacitance C starts to charge in the opposite direction towards the value $-V_Z$ through the resistance R. When the voltage v_c reaches the threshold voltage $-H\cdot V_Z$, another commutation starts towards the value V_Z, and the circuit is reset back to initial conditions.

Fig. 2.60: Rectangular pulse generator (monostable multivibrator).

Fig. 2.61: Voltages of the monostable.

Fig. 2.62: Triangular signal generator.

With a negative pulse on the input, we obtain a rectangular pulse with the time duration:

and amplitude $V_{PP}=2V_Z$. If $V_\gamma/V_Z<<1$, we have $T=\tau\cdot\ln[1/(1-H)]$.

Fig. 2.61 shows that the next input pulse can be applied only after a time delay T_1 needed to bring the circuit back to the stable state. The time T_1 determines the maximum frequency at which the input pulses can be applied. This time can be computed as follows:

2.15.7 Triangular signal generator

Fig. 2.62 shows the circuit able to generate the triangular waveform, which uses two OpAmps: the first device works as a comparator and the second one as an integrator. To explain the working principle at the time t=0, the comparator output signal v_1 has the value V_Z, defined by the Zener Z_1 (suppose $V_{Z1}=V_{Z2}=V_Z$). The resistive divider R_1, R_2 applies the voltage v_2 to the comparator non-inverting terminal. The signal v_1 is applied at the integrator input, and it is translated to a ramp with a negative slope, which is brought back to the comparator non-inverting input, determining a linear decrease of the voltage v_2. When such a voltage reaches the value 0V, we have a commutation of v_1, which switches from the value V_Z to the value $-V_Z$, as shown in Fig. 2.63.

The Fig. 2.63 shows that in the point A, a rectangular wave with a period T is present. For this reason, the circuit can simultaneously provide a triangular waveform and a rectangular signal with the same oscillation frequency. In order to determine the characteristics of the output waveform and obtain the circuit sizing equations, consider the range 0<t<T/2, where we have:
where V_{oH} denotes the maximum output voltage value. From the previous equation, we have: . Imposing t=T/2, we have , and then:

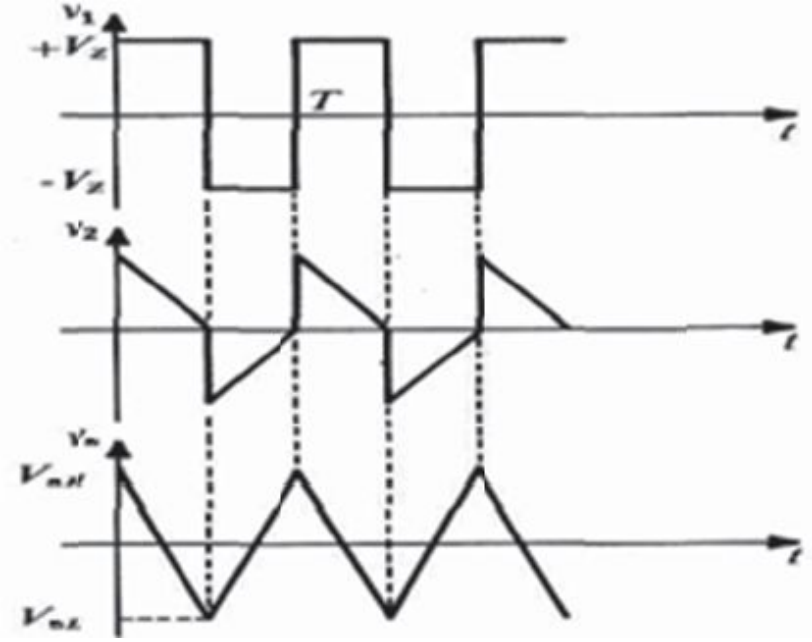


Fig. 2.63: Triangular signal generator waveforms.

where $V_{pp}=V_{uH}$ -V_{uL}. The voltage v_2 on the comparator non-inverting terminal is: $v_2 = V_Z \frac{R_2}{R_1+R_2} + v_o \frac{R_1}{R_1+R_2}$
The minimum value V_{uL} that the output voltage reaches can be obtained from the previous equation by imposing v_2=0. We can deduce: Similarly, the maximum value of the output voltage can be computed: The output amplitude V_{pp} is given by: $f = \frac{1}{T} = \frac{V_Z}{2RC} \frac{1}{2V_Z(R_2/R_1)} = \frac{R_1}{R_2} \frac{1}{4\tau}$ The frequency is given by:

2.16. EXERCISES

Ex. 1

Starting from the circuit in the figure, with an ideal OpAmp:
a) Compute *vout/vin* without the diode;
b) Assuming that the diode is ideal, plot *vout(t)* when *vin(t)* is a sinusoid with a peak value of *1V*; c) How would *vout(t)* be if the diode is considered real?
d) Compute the impedance seen by *vin* when *vin* is around *+5V* and *-5V*;
e) How can an integrator for the signal *vin* be built up with the same circuit?

a) $\frac{v_{out}}{v_{in}} = \frac{-200k}{10k+10k} = -10$

b)

c) To conduct, the real diode needs at least *0.7V* for direct polarization between the anode and the cathode. The two resistors of *10k*Ω halve the voltage. If the maximum amplitude of v*in* is less than *2·0.7=1.4V*, the diode does not conduct, there is thus no cutting of the waveform, and the *vout* will be as depicted in the figure. Instead, whenever *vin* exceeds *+1.4V*, the output *vout* will be limited to *+1.4·(-10)=-14V*.
d) For v*in* around *+5V*, the diode conducts, and the impedance is *10k*Ω. Instead, for *vin* around *-5V*, the diode is in interdiction, and the impedance is *10k+10k=20k*Ω.
e) Using a C instead of the *200k*Ω.

Ex. 2

I_B=100nA, V_{OS}=5mV, output swing=$(-V_{EE}+2V)\div(+V_{CC}-1V)$, A_{min}=10dB, and f_{high}=10MHz.

a) Study the operation of the circuit that has a sinusoidal input with a peak of 100mV.

b) Determine the output voltage when Vin=0V is caused by the OpAmp errors.

Ex. 3

Starting from the circuit in the figure, with an ideal OpAmp:
a) Compute *vout/vin* without the diode;
b) Assuming that the diode is ideal, plot *vout(t)* when *vin(t)* is the waveform depicted in the figure;
c) What would the difference be if the diode were considered real?
d) Connecting a capacitor *C* in parallel to the *50kΩ* resistance, how would *vout(t)* be with varying *C*?
e) If the diode were replaced with a capacitor C, what would *vout(t)* be?

$$\frac{v_{out}}{v_{in}} = 1 + \frac{40k}{10k} = +5$$

a)
b) See the figure on the side.
c) The actual diode determines a voltage drop of *0.7V* that must be subtracted from *vin* before the multiplication by the gain *+5*. See the figure on the side.
d) The capacitance and the diode are a peak extender with a time constant equal to $50k\Omega \cdot C$, as shown on the side.
e) The stage will be a derivator with a time constant of $50k\Omega \cdot C$, as depicted on the side.

Ex. 4

The LM324 has I_B=100nA, V_{OS}=5mV, output voltage swing=($-V_{EE}$)÷($+V_{CC}$-1.5V), and common-mode input voltage range=($-V_{EE}$)÷($+V_{CC}$-2V).

a) Study the function of the circuit, supplied with ±5V, and plot the output as a function of the ±5V range input signal, taking into account the voltage limits.

b) Determine the maximum output error due to the OpAmp errors.

c) Identify the two OpAmps that work with the lowest phase margin.

Ex. 5

Consider the circuit in the figure.

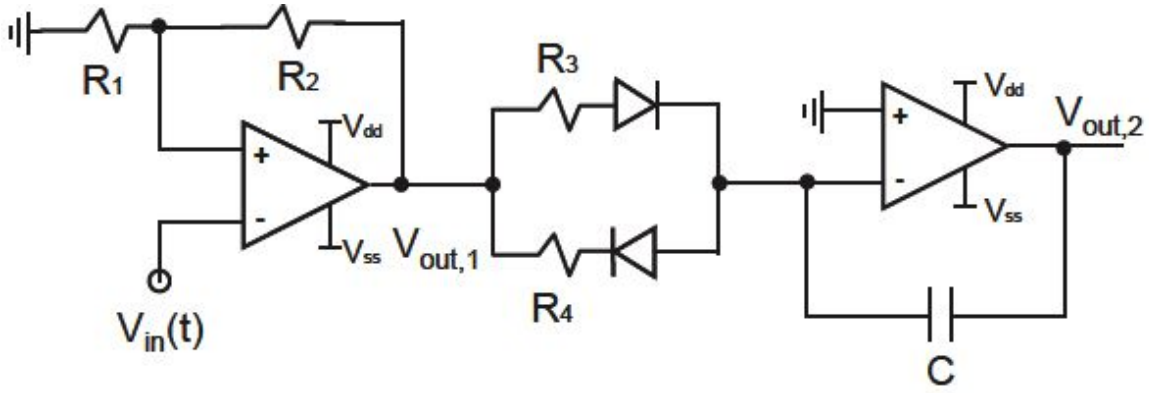
a) Plot the trigger characteristics (V_{in}-V_{out1}), specifying the commutation thresholds.

b) Apply the sinusoid $V_{in}(t)=2V\sin\omega t$ with $f=\omega 2\pi=1kHz$ to the input. Determine the output V_{out1} vs. time.

c) Determine the output V_{out2} in the response to the signal of the point b), assuming $V_D=0.7V$ for both diodes.

d) The resistance R_3 is increased to 12kΩ. Discuss the effect of the V_{out2} on the response of the sine of the point b).

e) Trace the evolution of V_{out2} in the condition of the point d), when it reaches the steady-state.



$R_1=10k\Omega$, $R_2=20k\Omega$, $R_3=R_4=10k\Omega$, $V_{DD}=5V$, $V_{SS}=-5V$, $C=50nF$

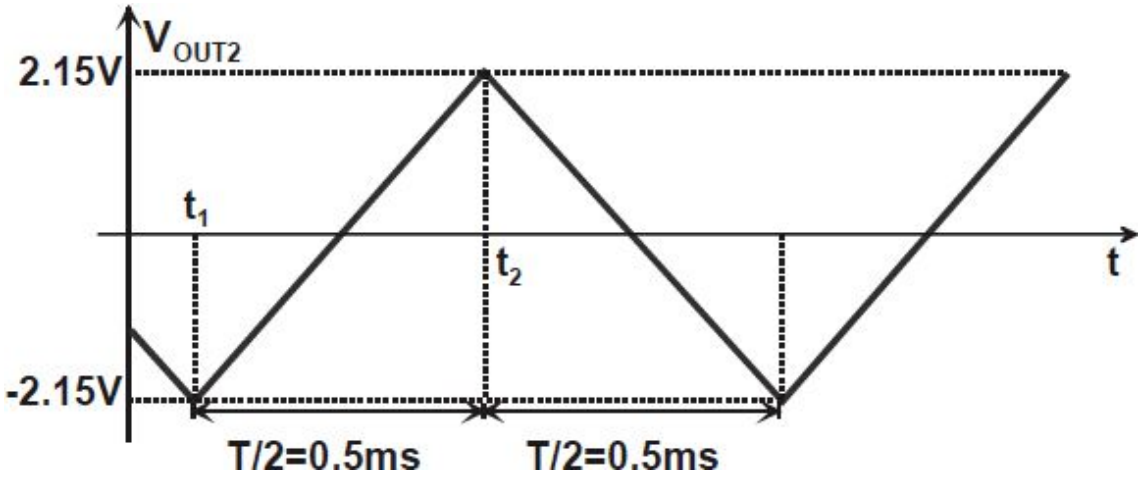
$R_1=10k\Omega$, $R_2=20k\Omega$, $R_3=R_4=10k\Omega$, $V_{DD}=5V$, $V_{SS}=-5V$, $C=50nF$

a) The circuit is an inverting Schmitt trigger, thus the characteristics is the following: Thresholds: $V_L= - R_1/(R_1+R_2)\cdot V_{DD}=-1.66V$ $V_H= R_1/(R_1+R_2)\cdot V_{DD}=1.66V$

b) The Schmitt trigger changes the sinusoidal wave to a square wave. The square wave has the same frequency as the input sinusoid. You can calculate the times when V_{in} exceeds the trigger threshold: $1.66V=2V\cdot\sin(2\pi ft_1)$ $t_1=155\mu s$ $-1.66V=2V\cdot\sin(2\pi ft_2)$ $t_2=655\mu s$

Because the Schmitt trigger is symmetrical, the output signal's duty-cycle is 50%.

c) When $V_{out1}=5V$, the diode in series to R_3 conducts, and on R_3 flows a current equal to $I_-=(5V-0.7)/R_3=0.43mA$ that discharges C in a linear way (constant current). When $V_{out1}=-5V$, the diode in series to R_4 conducts, and on R_4 flows a current $I_+=(-5V+0.7)/R_4=-0.43mA$ that discharges C in a linear way (constant current). In both cases, the slope of the voltage variation across the capacitance is $|dV_C/dt|=|I_+/C|= I_-/C=8600V/s$. In a half period, the variation in the voltage is $\Delta V=(dV_C/dt)\cdot T/2=4.3V$. Because the circuit is completely symmetrical, the voltage across the capacitance is on average null, and thus $V_{out2}(t)$ will vary around 0V.



d) Because the current that charges the capacitance $I_+=(-5V+0.7)/R_4=-0.43mA$ is greater than the current that discharges the capacitance $I_-=(5V-0.7)/R_3= 0.36mA$, the capacitance on average is discharged, i.e. the mean value of V_{out2} results greater than 0V. At the steady-state, the discharge phase will start from V_{out2} saturated to $V_{DD}=5V$. While I_- employs all the half-period to discharge C from 5V to the voltage V_{low}, I_+

employs a time ΔT_{up} lower than T/2 to bring V_C to 5V. In the remaining time ($T/2-\Delta T_{up}$), the voltage across C remains constant due to the fact that it is limited by the supply voltage and that the current is null.

e) Slope during charge: $(dV_C/dt)_{up}=((5V-0.7V)/R_4)/C=8600$ V/s;

Slope during discharge: $(dV_C/dt)_{down}=((5V-0.7V)/R_3)/C=7166$ V/s;

Voltage variation during discharge of C: $\Delta V=(dV_C/dt)_{down}\cdot T/2=3.58V$;

Charging time: $\Delta T_{up}= \Delta V/(dV/dt)_{up}=416\mu s$.

Ex. 6

The quad-OpAmp LM324 has $V_{OS}=3mV$.

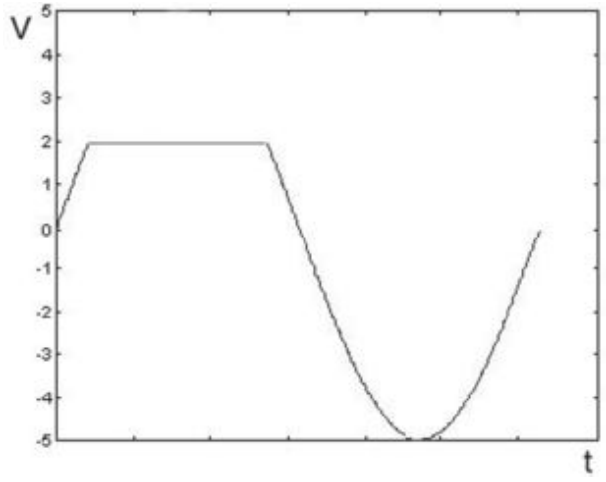
a) Explain the behavior for $V_B=2V$ and V_A sinusoidal of $\pm 5V$.

b) Determines V_{out} for $V_A=+2V$ and $V_B=+5V$.

a) The circuit in the figure employs three of four OpAmps contained in the package LM324. The chip has a dual supply of ±15V. The first OpAmp, OA1, is a buffer configuration and transfers the voltage V_A to the output. The same work is done by OA3, which picks up the voltage on the non-inverting terminal and brings it to the output with unity gain. Therefore, the whole stage can be studied starting from the behavior of OA2 and the electric network around it when V_B is maintained fixed to 2V, and the OA1 output varies with a sinusoidal fashion between ±5V. It can be noticed that the signal range is below the supply voltages; thus, there is no problem of OpAmp saturation.

Assuming to start from V_A=0V, it can be noticed that the OA2 output will saturate at +15V because of the positive V_B voltage. As a result, the OA2 output diode will be inverse biased. Starting from this situation, the feedback loop of OA2 is open, and the OpAmp has NOT negative feedback (V_B is not reported to the output). In such condition, V_A is on the non-inverting terminal

of OA3 through a resistor. This situation is valid for every V_A<2V; it is as if OA2 acted as a comparator. This is true also when V^-_{OA2} exceeds 2V. In this condition, the output will be -15V, directly biasing the diode which will start to conduct. However, the output is not fixed to -15V because the feedback loop is closed: the feedback is such that the output of OA2 will lead to a voltage such that the two input terminals are at the same voltage (ideally), that is 2V. It is exactly what happens when V_A exceeds 2V. In this situation, V_A does not give any contribution on the output because it is as if it "saw" an inverting configuration with a gain -R2/R1, where R2 = 0. The role of the 10k is now evident. Ultimately, the waveform V_{out} can be summarized as follows:



b) As for the previous point, since $V_B > V_A$, OA2 works with an open loop. The output signal V_{out} will be equal to V_A = 2V. In reality, we must also consider the effect of the offset voltage of the three OpAmps, schematized as a voltage generator at the input of each OpAmp with an unknown sign (± VOS). Considering the worst case and this particular operating condition of the circuit, it can be concluded that $V_{OS,OA1}$ and $V_{OS,OA3}$ are present on the output without any change while $V_{OS,OA2}$ does not give any contribution because it works in an open loop.

3

Frequency compensation

“Let’s talk about me for a minute

Well how do you think
I feel about what’s been going on

Let’s talk about me for a minute

Well how do you think
I feel about what's gone wrong
Let's talk about dreams
I never learned to read the signs

Let's think about what it all means
I never seem to have the time
Let's talk about you and your problems
All that I seem to do is spend the night
Just talking 'bout you and your problems
No matter what I say I can't get it right."

"Let's talk about me", The Alan Parson Project

3.1. Frequency Response

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}.

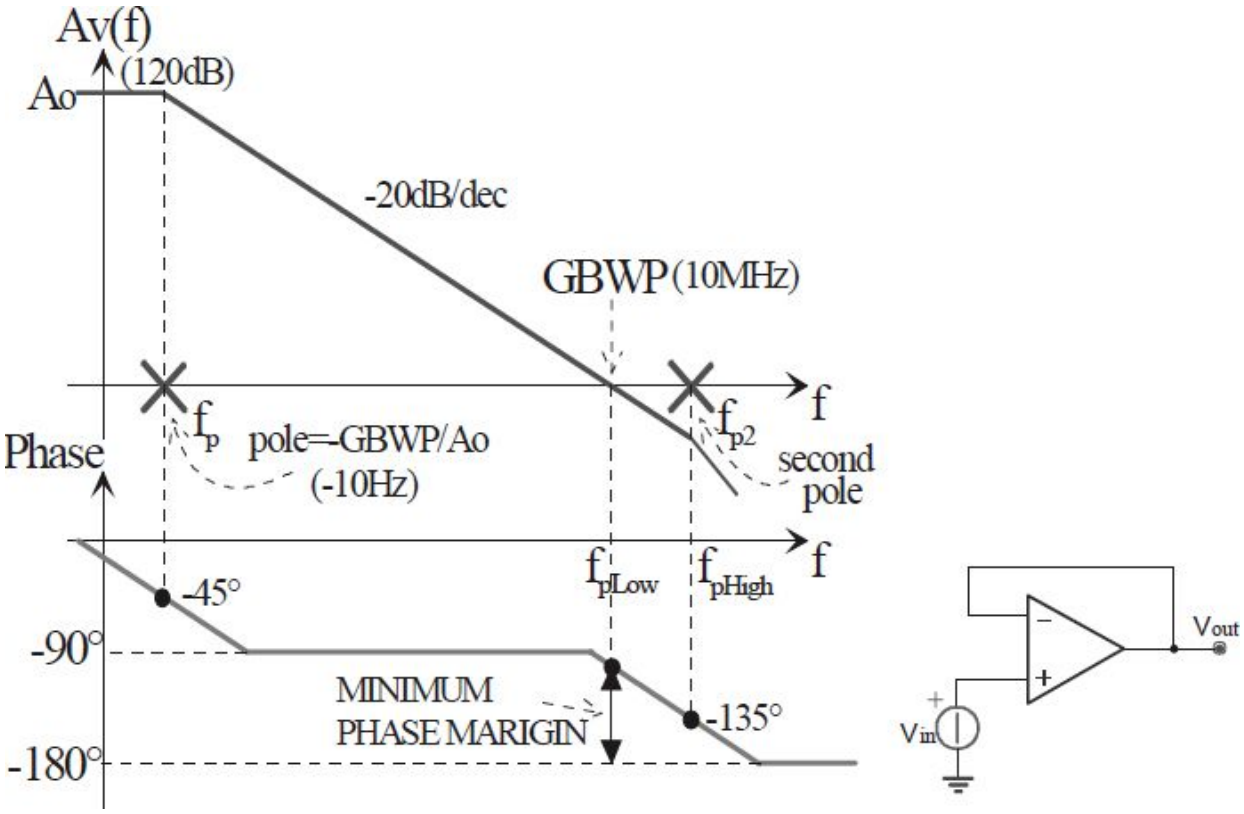


Fig. 3.1: Open-loop Bode plots of a compensated OpAmp and its buffer connection.

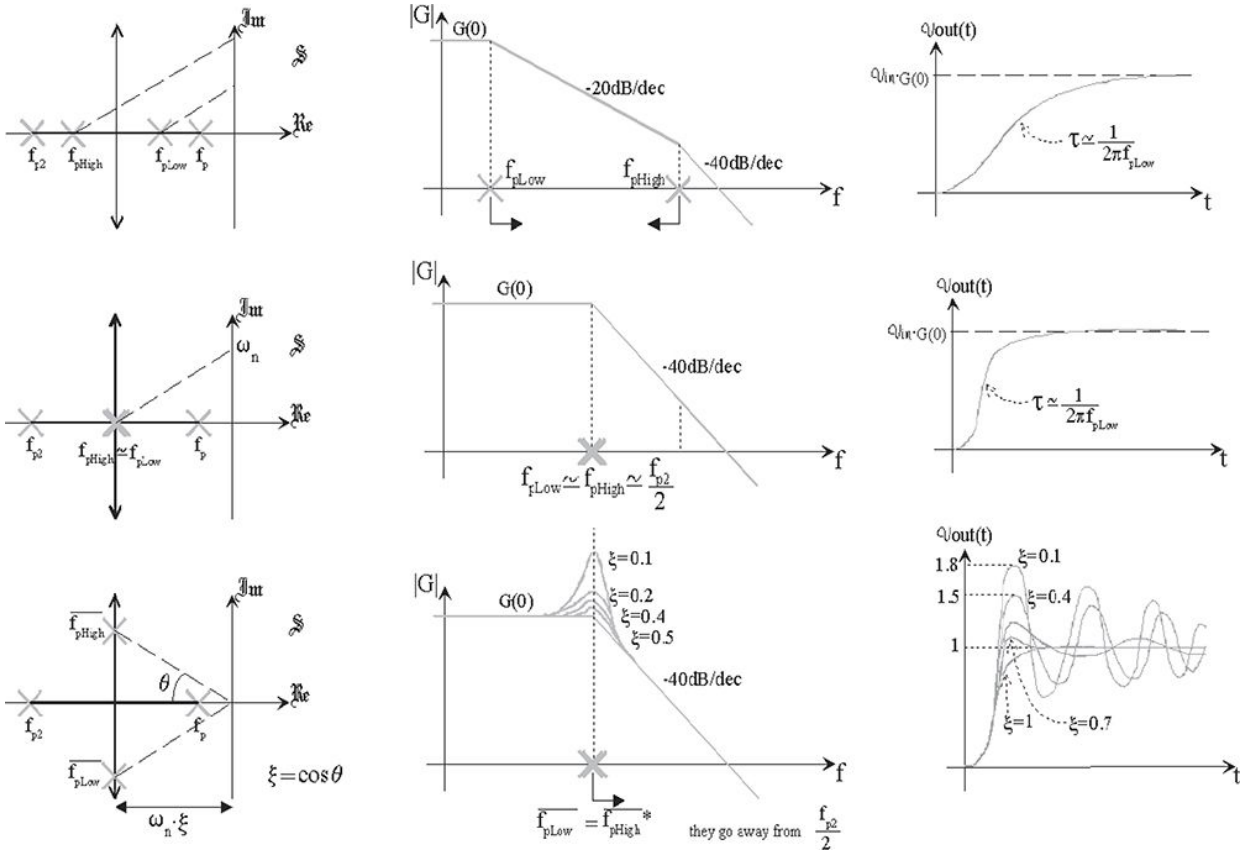


Fig. 3.2: Closed-loop pole migration as G_{loop} changes (at the bottom), corresponding closed-loop gain G (in the middle), and the unit step response (at the top).

3.2. "INTERNAL" OPAMP COMPENSATION

Depending on the internal compensating capacitor C_c of the OpAmp, a "voltage mode" operational amplifier can be compensated (if the second pole in the Bode plot of its frequency response occurs when |A(s)|<1) or not. Usually, in this kind of OpAmp, the internal compensation (increasing C_c) reduces the SR.

3.2.1 DOMINANT-POLE COMPENSATION (DIRECT AND MILLER COMPENSATION) Frequency compensation is a process that is applied to circuits which are not stable enough: the purpose is to change the loop gain of the circuit so that the phase margin increases to an acceptable level. The procedure is not unique,

but depends on the type of the circuit you want to compensate: each method has its advantages and disadvantages. The motivation that leads to choose one rather than the other depends on the performance to be achieved.

Since $G_{loop}(s)=-A(s)\cdot B(s)$, to perform frequency compensation, it is necessary to modify A or B or both through an appropriate electric network placed between the two nodes of the circuit. Now, we will analyze some compensation methods that modify only the forward gain of the system: next paragraphs will describe different methods that modify the ideal gain leaving the forward block gain unchanged. Let us start with the dominant-pole method. To better understand the reasoning that follows, we refer to Fig. 3.3. In addition to high frequency poles in which f_2 is the most dominant and which are limiting bandwidth and cannot be eliminated (a good design of an integrated circuit will try to move them to higher frequencies to increase the bandwidth, but, consequently, also the noise), it is possible to introduce C_c in order to place a pole at the frequency f_1. To compensate the Op-Amp, C_c shall be increased to move the dominant pole to $f_{d,comp}$. On the contrary, to leave the Op-Amp not compensated, it will be enough to move the pole back to $f_{d,not\ comp}$ if we would like, for instance, a value that meets the requirements of the SR and the bandwidth ($f_c=f_2$) based on the Op-Amp feedback ($|1/\beta|$).

To assist the user in the operational compensation, many manufacturers make some of the non-compensated Op-Amp pins available in order that a capacitor can be externally connected. This is the case of the LM108. As shown in Fig. 3.4, it is possible to insert a capacitor between two specifically-made terminals or between the ground and one Op-Amp terminal that helps improve the supply noise rejection.

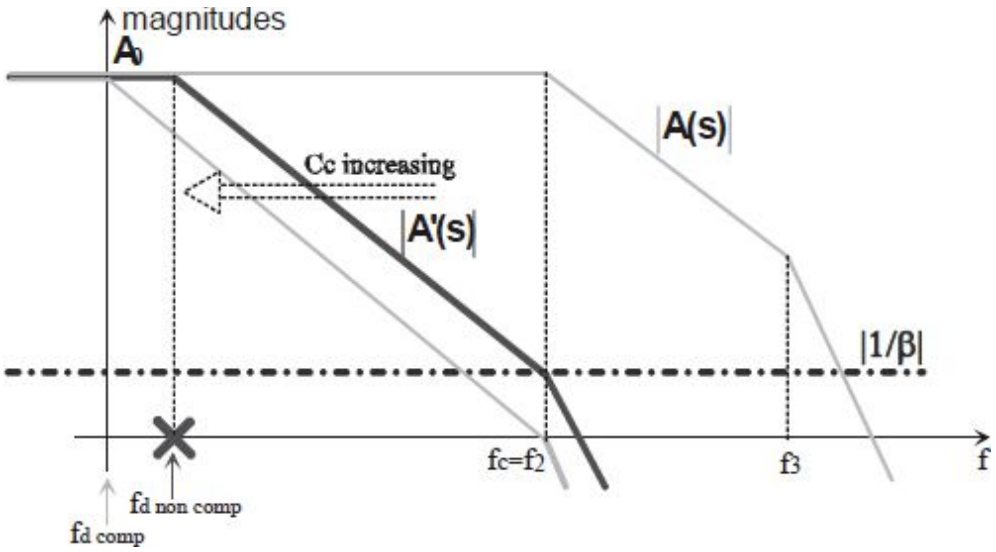


Fig. 3.3: Dominant-pole compensation.

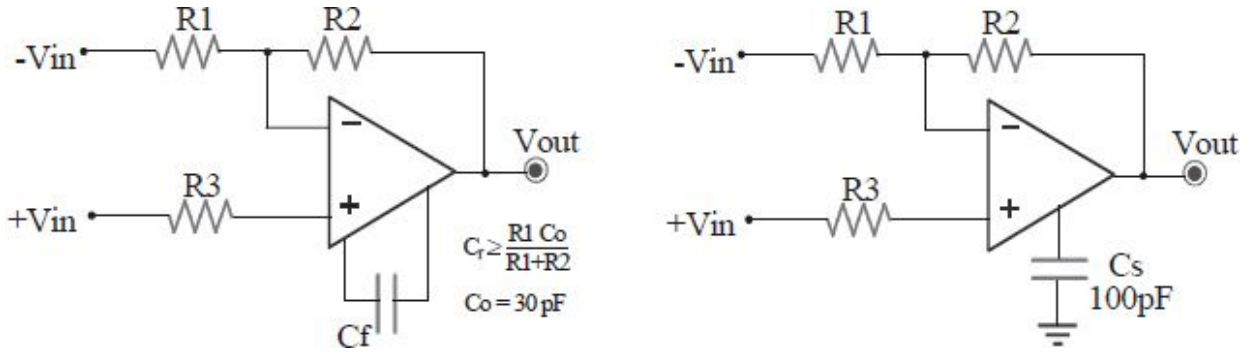


Fig. 3.4: Dominant-pole compensation by introducing a capacitor externally connected to the Op-Amp.

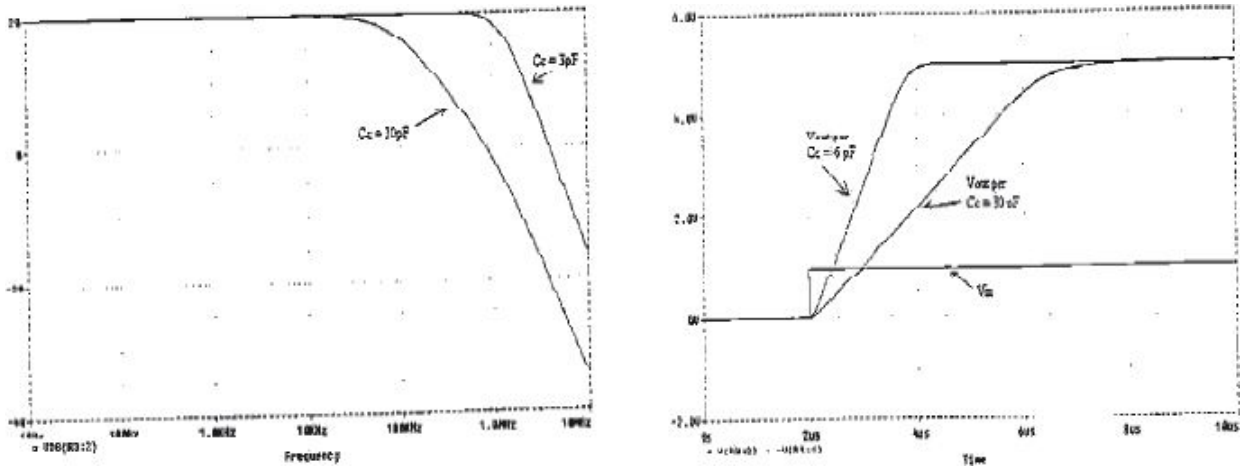


Fig. 3.5: SPICE simulation of the LM108.

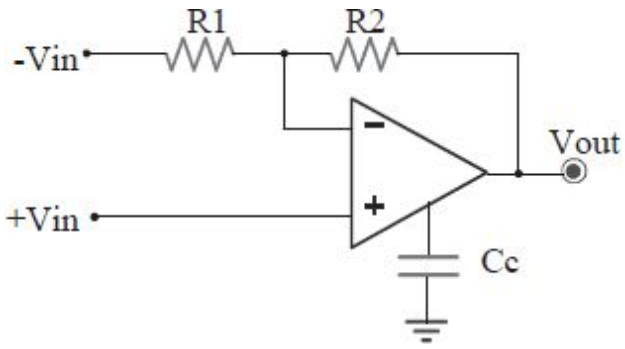


Fig. 3.6: Compensation through the insertion of an external C_c.

We observe that over-compensation (that is $f_d << f_{d\ comp}$) reduces the bandwidth too much and therefore the output switching speed, as demonstrated by the graphs of SPICE simulations carried out for the LM108, which are shown in Fig. 3.5.

On the data sheet, manufacturers give the value of the capacitance to compensate the OpAmp in the worst case (in terms of the closed-loop circuit stability) of buffer configuration ($|1/\beta|=1$). This capacitance is usually denoted by C_o and gives: $f_c = \frac{A_o}{2\pi \cdot R_{ic} \cdot C_o}$

In general, to have $\beta=R_1/(R_1+R_2)$ and a phase margin of 45°, it is enough to choose: $C_c = C_o \cdot \frac{R_1}{R_1 + R_2}$

and insert the capacitor as shown in Fig. 3.6.

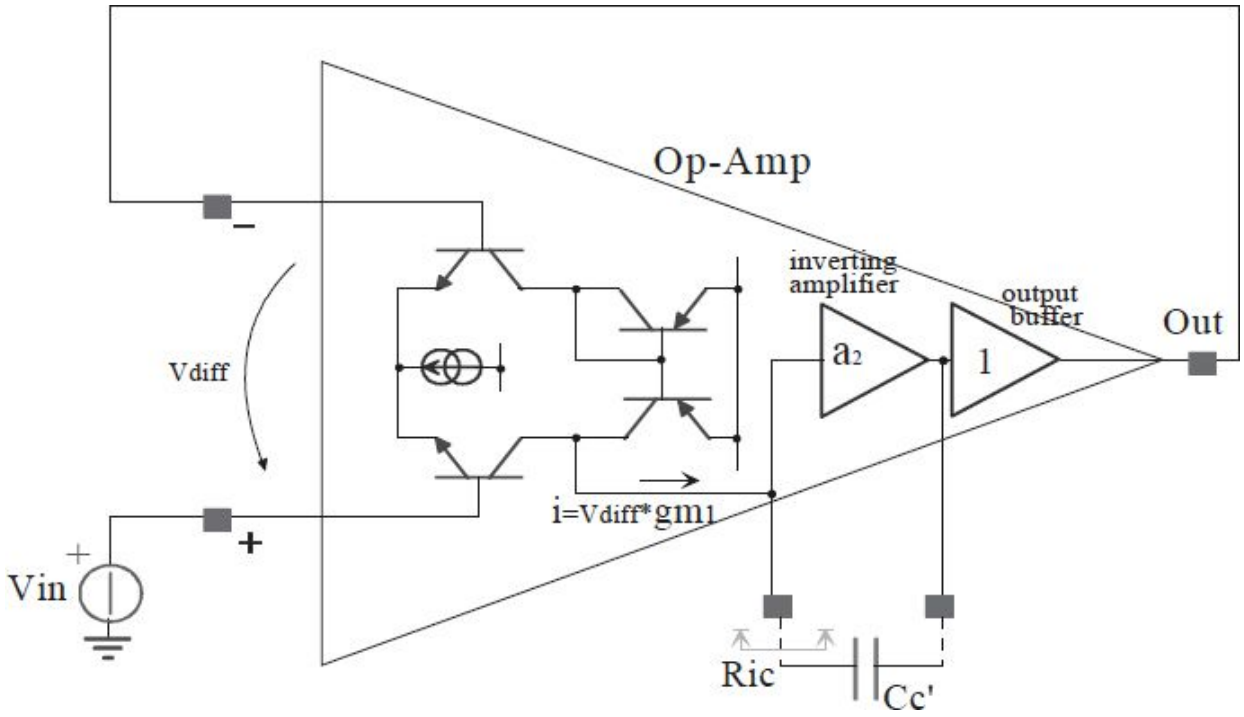


Fig. 3.7: Op-Amp internal structure: note where the compensating capacitor is placed.

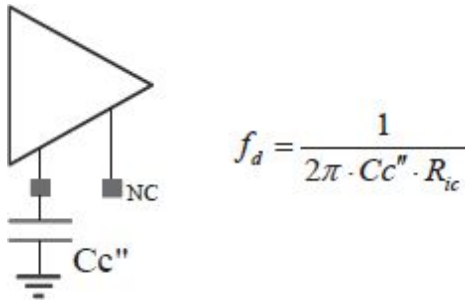


Fig. 3.8: Connection of the capacitance that does not exploit the Miller effect.

The internal structure of the Op-Amp can be represented as in Fig. 3.7. In this case, thanks to the Miller effect, the capacitance seen at the input node of the gain block a_2 ($|a_2|>>1$) is amplified by $1-a_2$, then we have:

$$f_d = \frac{1}{2\pi \cdot C_c' \cdot R_{ic} \cdot (1-a_2)}$$

On the other hand, had we used the capacitor in the way shown above, we would have had what is shown in Fig. 3.8. At the same desired f_d, a much larger and cumbersome C_c should have been used. For instance, with

R_{ic}=10MΩand a desired f_d=10Hz, C_c"=1μF is required whereas using the Miller effect, C_c'=10nF is enough if a_2=100.

3.2.2 Pole-Zero Compensation

This method is a variation of the dominant-pole compensation. It consists in "erasing" the first pole (f_1) of the Op-Amp through a zero, leaving the second pole (f_2) unchanged, and of placing an additional pole at low frequencies (f_p). Proceeding with this way, if you want to compensate the circuit with a 45° phase margin, you can assign to the corner frequency f'_c the value of the second pole of the Op-Amp.

The practical way to create a pole and a zero in the Op-Amp open-loop gain consists in connecting a series RC network either between the second stage input terminal and the ground or between the input and the output of the second stage. In the latter case, the Miller effect of magnification of the capacitance that gives the pole is exploited. The result of such compensation is shown in Fig. 3.9. This will ensure stability with an open-loop bandwidth (f_p) higher than that obtainable from using the dominant-pole compensation method (f_d) (see Fig. 3.3). As already mentioned, it is possible to perform pole-zero compensation by using an RC network in two different ways that are shown in Fig. 3.10 and Fig. 3.11, where the obtained results are highlighted.

To ensure pole-zero cancellation (actually impossible to realize with precision due to components' degree of tolerance and thermal drifts), $f_z \equiv f_1$ will be chosen. The dominant pole will become: $f_p = \frac{f_1 \cdot R_{cm}}{R_{cm} + R_{ic} \cdot (1 - a_2)}$

and will be chosen based on the required feedback:

$$f_p = \frac{f_2}{A_0 \cdot \beta}$$

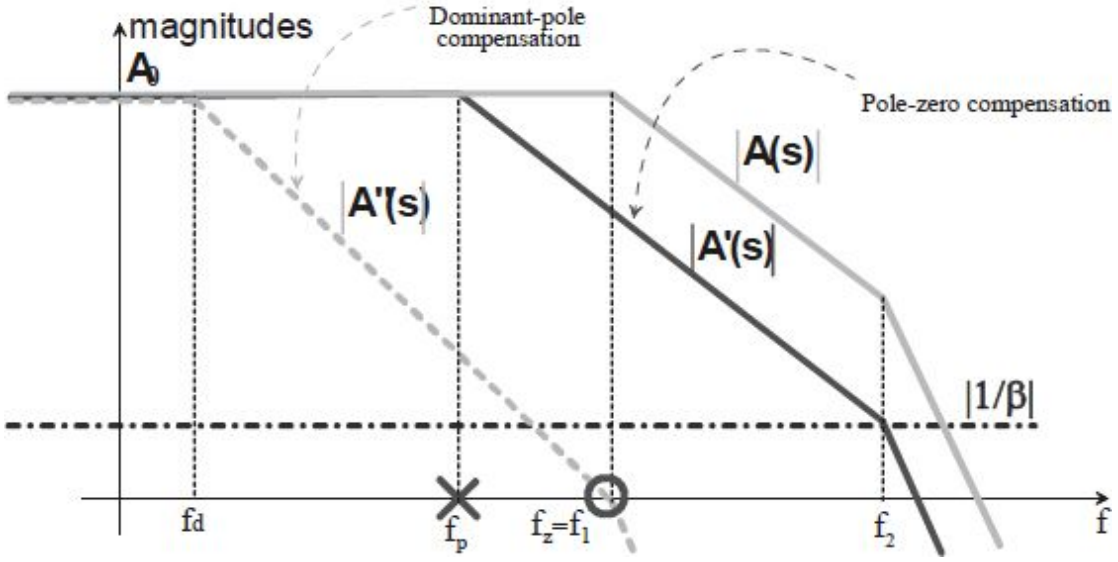
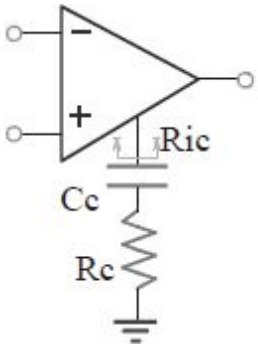


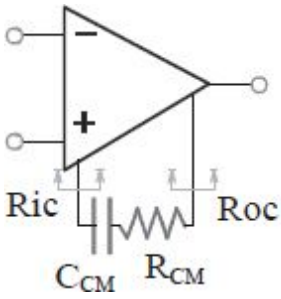
Fig. 3.9: Bode plots of the Op-Amp compensated with the pole-zero method.



$$f_z = \frac{1}{2\pi \cdot Rc \cdot Cc}$$

$$f_p = \frac{1}{2\pi \cdot (Ric + Rc) \cdot Cc} = f_z \cdot \frac{Rc}{Ric + Rc}$$

Fig. 3.10: Pole-zero compensation not exploiting the Miller effect.



$$f_z = \frac{1}{2\pi \cdot Rcm \cdot Ccm}$$

$$f_p = \frac{1}{2\pi \cdot (Rcm + Ric \cdot (1 - a_2)) \cdot Ccm} = f_z \cdot \frac{Rcm}{Rcm + Ric \cdot (1 - a_2)}$$

Fig. 3.11: Pole-zero compensation exploiting the Miller effect.

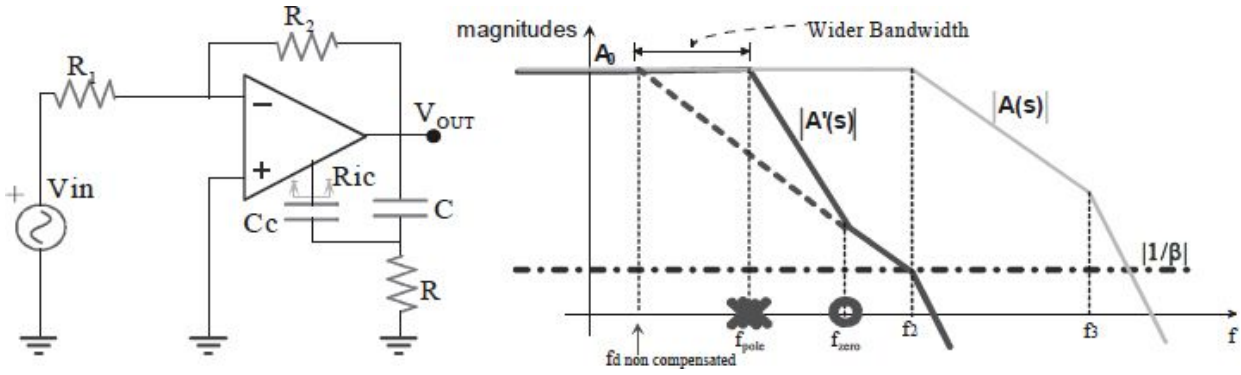


Fig. 3.12: Two-pole compensation.

3.2.3 Two-Pole Compensation

The aim is to make the Miller compensation only near the zero-crossing frequency so that the open-loop bandwidth (first pole f_d) obtained is much higher than $f_{dnotcomp}$ achievable with the pole-zero method. The circuit that realizes this compensation and its corresponding Bode plot are shown in Fig. 3.12. In addition to the two mentioned poles, the network also introduces a zero, and, if $C>>C_c$, we have: $f_{zero} \cong \frac{1}{2\pi \cdot R \cdot C}$

The capacitor C is directly connected to the output to exploit the higher current availability with respect to the internal node which “shows”, from the outside, an impedance value of Ric.

To estimate component values, it is possible to proceed, using the following three steps:

- Knowing C_0, R_1, and R_2, calculate $C_c = C_o \cdot \frac{R_1}{R_1 + R_2}$
- Choose $C=10 \cdot C_c$;
- Place the zero at least a decade before f_2, that is, choose $R = \frac{10}{2\pi \cdot f_2 \cdot C}$

Fig. 3.13 shows a comparison, made through a SPICE simulation, between the LM101A in inverting configuration (A_V=-10) Miller compensated by using 3pF and 30pF and the two-pole method with C_c=3pF, C=47μF, and R=3Ω. It is possible to note that the latter method retains higher open-loop bandwidth.

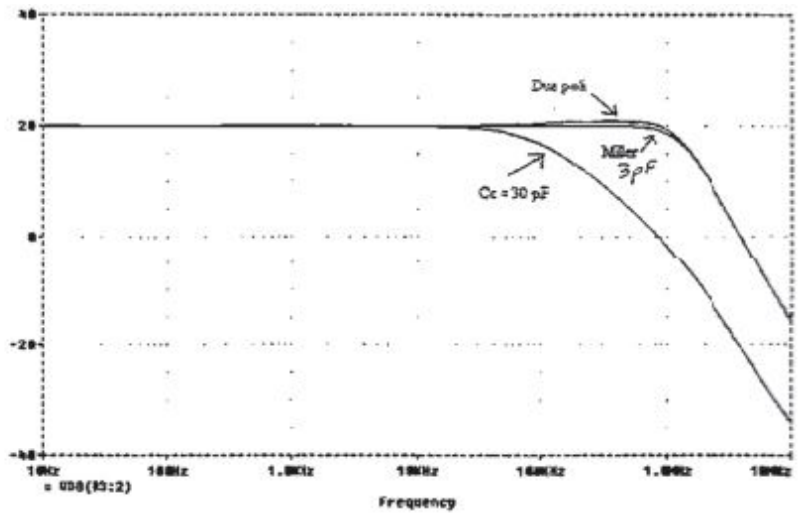


Fig. 3.13: SPICE comparison between the single pole and the two-pole compensation.

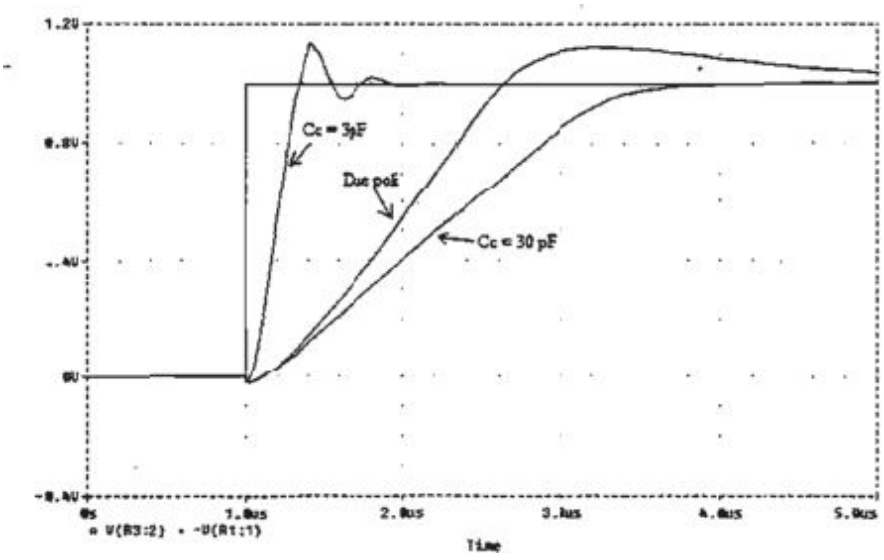


Fig. 3.14: SPICE transient simulation of the step response of a compensated buffer.

In the case of non-inverting buffer configuration (A_v=1), Miller compensation with 3pF and 30pF and the two-pole method with C_c=30pF, C=70pF, and R=3,9kΩhave been used. Fig. 3.14 demonstrates the transient response obtained with a SPICE simulation of the buffer mentioned above. You can see that, although C_c is equal to 30pF, the transient speed is greater than that of the simple two-pole configuration.

3.2.4 Feedforward Compensation

Usually in “voltage-mode” Op-Amps, there is a “slow” stage that limits the bandwidth of the whole integrated circuit. To eliminate the effect of this stage, it is possible to introduce a parallel “fast” path. Suppose that we can model the OpAmp as depicted in Fig. 3.15.

The block a_1 is the “slow” block (with the low frequency pole). The transfer function of the system modeled in such a way can be written as: $A' = (a_1 + h) \cdot a_2 = A + h \cdot a_2$

We choose h(f) to be high-pass, for instance h(0)=0 and h(∞)=1, as shown in Fig. 3.15. In this way, only the inverting terminal benefits from such a change.

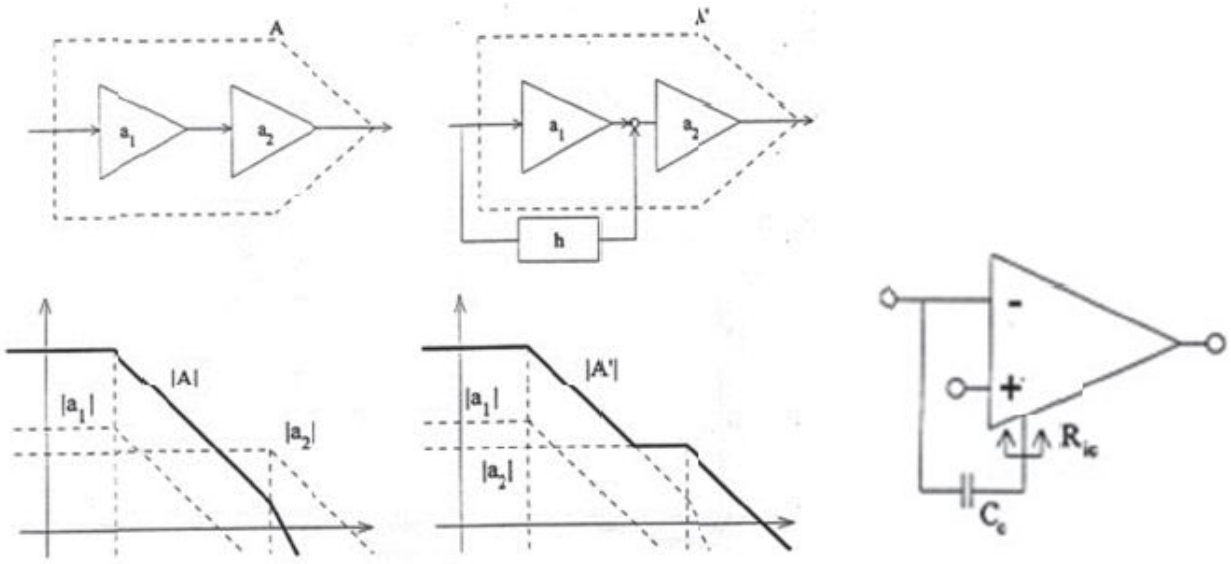


Fig. 3.15: Modeling of the Op Amp through two blocks (left) and Feedforward compensation (right).

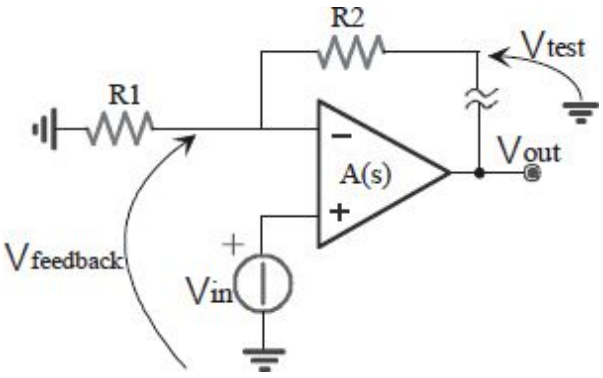


Fig. 3.16: Operational amplifier in non-inverting configuration.

3.3. "External" Compensation of OpAmps

3.3.1 Non-Inverting Configuration

Now let us study the Op-Amp in non-inverting configuration shown in Fig. 3.16. Let us proceed with the estimate of the quantities useful for studying the circuit stability. To calculate β, we can make a cut on the terminal of R_2 close to the output, as shown in Fig. 3.16. Then, we apply a signal V_{test} at this point and read the signal shown as $V_{feedback}$ in the figure.

At this point, βis given by: $$\beta = \frac{v_{feedback}}{v_{test}} = \frac{R_1}{R_1 + R_2}$$

Now, the loop gain is immediately computable and equal to:

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

The closed-loop gain is computed by applying feedback theory rules:

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

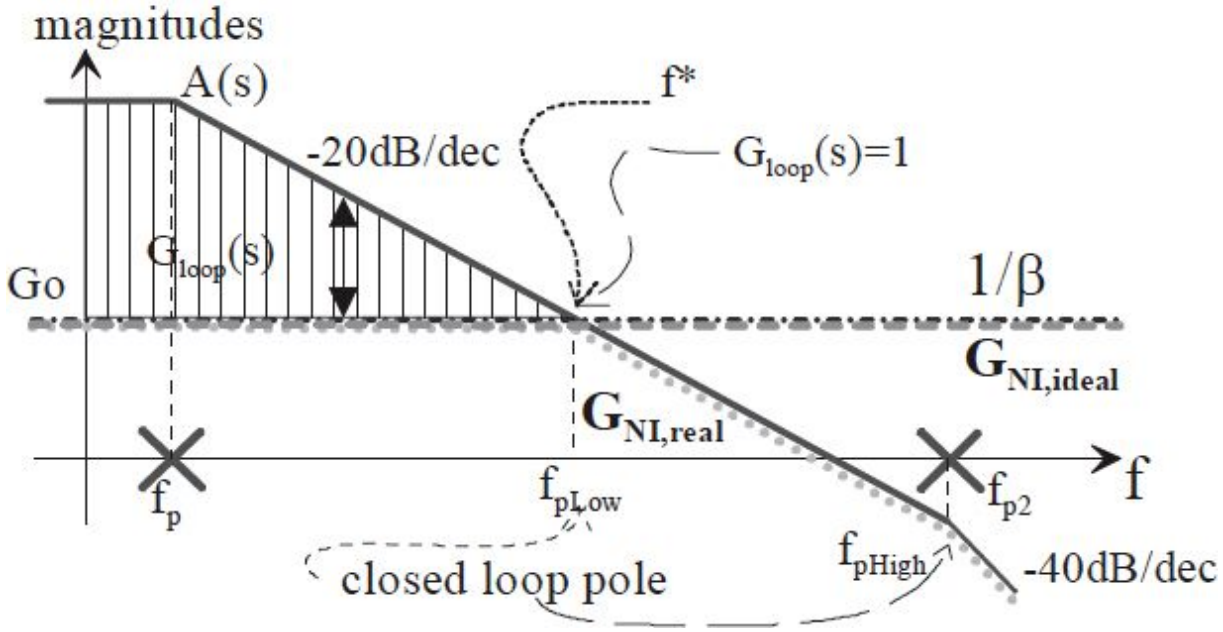


Fig. 3.17: Bode plot of the non-inverting configuration.

Fig. 3.17 depicts the frequency responses just calculated. For this configuration (but only for this one), the **closed-loop gain** of the stage is equal to the lower of A(s) and 1/β(s). Not being a general case, it s not convenient to use this method to graphically obtain the closed-loop gain. It is better to "forget" the two graphs of A(s) and 1/β(s) and plot the **ideal gain** as expected by the theory. In the case of the circuit shown in Fig. 3.16, the ideal gain would be equal to $1+R_2/R_1$. Once plotted on the same graph of Fig. 3.17, we realize that the ideal gain coincides (but only for this configuration) with the 1/β graph.

Well, the **real gain** is virtually identical to the ideal one as long as the stage is "ideal", that is, as long as $G_{loop}>>1$ or until the frequency f^*. From then on, the feedback collapses (because A(s) or β(s) becomes too small). It can be shown that from then on, the real gain continues as the ideal one, but also experiences the same "flexions" of A(s) or β(s), that is, -20dB/dec

corresponding to their poles. Although β is constant in this case, hereafter we will also deal with more general cases with β(s) having poles and zeros.

Finally, as soon as the feedback does not work anymore, that is, at the frequency f^* at which $|G_{loop}(f^*)|=1$, the closed-loop gain suffers a decline because it has a pole. This is why in the case of Fig. 3.17, the pole f_{pLow} coincides with f^*. The next pole f_{pHigh} will be the first pole in OpAmp A(s) to appear beyond f^*, that is, the second pole f_{p2}.

3.3.2 Stability Criteria

We prefer to plot 1/β instead of β because, in this way, we can better appreciate the intensity of G_{loop}, equal to the distance (i.e. ratio) between A(s) and 1/β(s), and extract G_{NI} (coincident with 1/β in this case) as long as there is a significant G_{loop} (>10). In addition, through this chart, it is possible to analyze the stability of the circuit as different quantities change. For the same purpose, you can use the root locus reproduced in Fig. 3.18. It is important to note that the axes are in linear scale whereas in the previous Bode plot, the scale was logarithmic. This fact implies that in the root locus, it is possible to see the movement of both poles as β changes whereas in the Bode plot, the retreat of the high frequency pole is not perceptible.

As already mentioned, the loop has intensity equal to the difference between A(s) and 1/β(s). Concerning its phase, it is possible to draw the corresponding Bode plot, recalling that each pole of G_{loop} (that is, of A(s) or β(s)) introduces a phase shift of -90° while each zero introduces a phase shift of +90°. The **Bode stability criterion** states that the configuration is stable if, at the frequency f^* where $|G_{loop}(f^*)|=1$, the phase of $Gloop(f^*)$ is far from -180°. In fact, if the phase shift is exactly -180°, it would mean that the loop reapplies the signal from the OpAmp output to its inverting input with a phase shift that would result in net positive feedback, thus causing instability and oscillations. An idea about the distance from this condition is provided by the **phase margin** that is what is lacking in the $G_{loop}(f^*)$ phase to reach -180°.

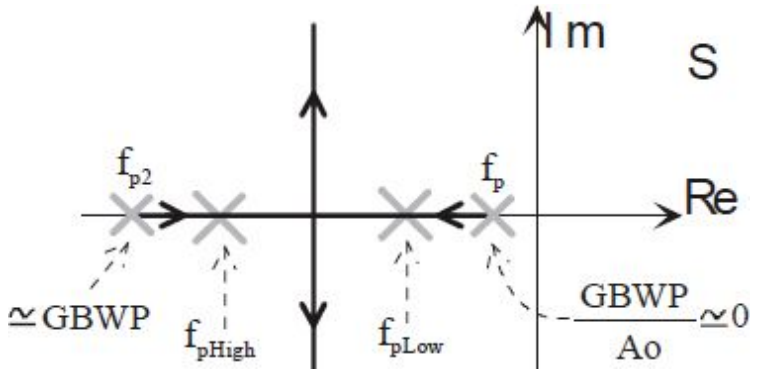


Fig. 3.18: Root locus of the non-inverting configuration.

In order not to be forced to draw the Bode plot of the phase of Gloop(f), it is possible to proceed with the other way. It should be noted that to reach -180°, Gloop must have suffered from the intervention of a number of poles N_{poles} and of zeros N_{zeros} such that N_{poles}-N_{zeros}=2 holds. This implies that at the frequency f^* where $|G_{loop}(f^*)|=1$, the **closure angle** with which A(s) and 1/β(s) intersect each other is equal to 40dB/dec. If, instead, the closure angle is 20dB/dec, the configuration is certainly stable (this is true if the network is minimum phase, i.e. without poles and zeros with positive real part, a condition observed in almost all circuits). It may also happen that the closure angle changes exactly in correspondence of f^* because, by sheer coincidence, at that frequency, there is also a pole or a zero of A(s) or β(s). Finally, it is possible to predict whether the stage is stable or not simply by observing the closure angle at the point on the Bode plot of the magnitude of |A(s)| and |1/β(s)| where $|G_{loop}(f)|=1$. Depending on the closure angle, we will have the following cases:

- 20dB/dec before, 20dB/dec after — stable configuration;
- 20dB/dec before, 40dB/dec after — stable configuration with PM=45°;
- 40dB/dec before, 20dB/dec after — stable configuration with PM=45°;
- 40dB/dec before, 40dB/dec after — unstable configuration with PM≈0°;
- Closure angles greater than 40dB/dec — certainly divergent configuration.

3.3.3 Boundary Condition of PM=45°

It is possible to consider the condition of PM=45° as the boundary for reliable stability. In this case, the system response to a step will be the typical fast response of a two-pole system, but without an excessive overshoot. As will be seen in the demonstration that follows, a phase margin equal to 45° corresponds to ξ=0.5, which corresponds to the response shown at the top right of Fig. 3.2.

The configuration that has the maximum intensity of feedback (i.e. the signal that comes back from the OpAmp output to its input) is the buffer, which has β=1 (Fig. 3.19). Note that if the second pole is coincident with the GBWP of the OpAmp (i.e. the frequency for which |A(f)|=0dB), the corresponding closure angle is equal to 20dB/dec before and 40dB/dec after, that is, the phase margin is 45°. The poles in the S plane are not coincident,

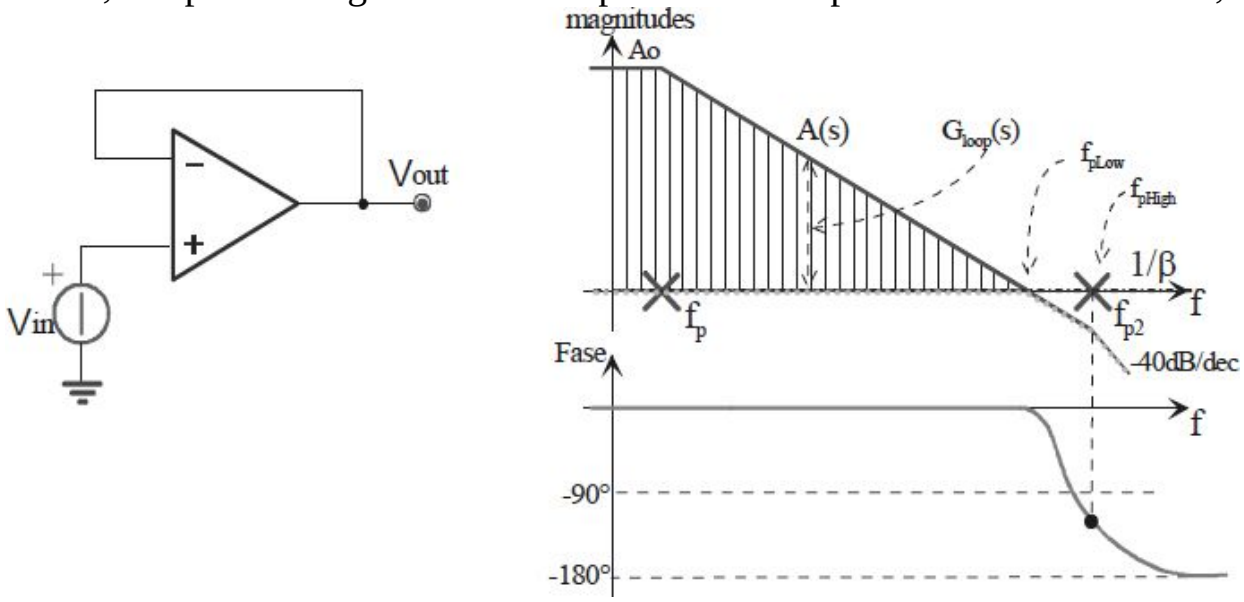


Fig. 3.19: Bode plots of the frequency response of a buffer.

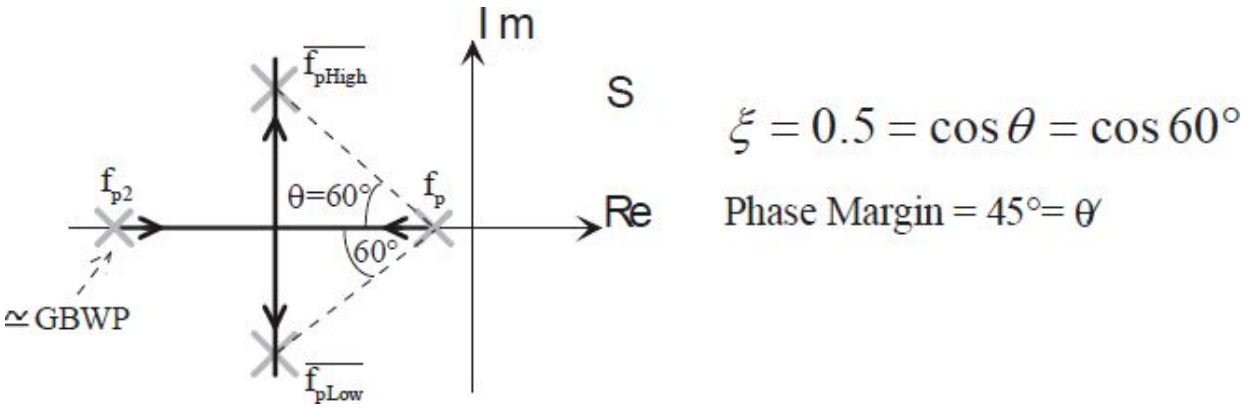


Fig. 3.20: Root locus and closed-loop pole positions in the case of phase margin equal to 45°.

but are complex conjugated at about 60°. To verify what has just been said, it is enough to solve the following quadratic equation:

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

which yields, for β=1:

$$s^2 + s\cdot\left(\frac{1+A_0}{A_0}\right)\cdot GBWP + \left(\frac{1+A_0}{A_0}\right)\cdot GBWP^2 = 0$$

Since $1+A_0 \cong A_0$, it can be simplified as in the following: $s^2 + s \cdot GBWP + \cdot GBWP^2 = 0$
whose solutions are:

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

Fig. 3.20 represents the root locus graph where it is possible to locate the positions of the just calculated poles. Recall the expression for the phase margin:

$$PM = 180° - \sum arctg\frac{f^*}{f_{pi}} = 180° - arctg\frac{f^*}{f_{p1}} - arctg\frac{f^*}{f_{p2}} = 90 - arctg\frac{f^*}{f_{p2}}$$

where f^* represents the frequency where $G_{loop}(s)$ falls to 0dB, i.e. A(s) intersects 1/β(s). In this case, f_{plow} (the first closed-loop pole) coincides with f^* and also with the second OpAmp pole f_{p2}. Since:

$$\left|f_{plow}\right| = \frac{\frac{f_{p2}}{2}}{\cos\theta} = \frac{\frac{GBWP}{2}}{\cos 60°} = \frac{\frac{f_{p2}}{2}}{\frac{1}{2}} = f_{p2}$$

we obtain:

$$PM = 90° - arctg\frac{|f_{p2}|}{|f_{p2}|} = 90° - arctg1 = 90° - 45° = 45°$$

If we assume that f_{p1} is sufficiently lower than f_{p2} so that it is possible to consider its contribution to the phase margin equal to 90°, the PM can be

$$PM = 90° - arctg\frac{|f_{plow}|}{|f_{p2}|}$$

calculated by the following simple expression:

Note that, although the poles of G_{loop} are f_{p1} and f_{p2}, the closed-loop pole is f_{plow}.

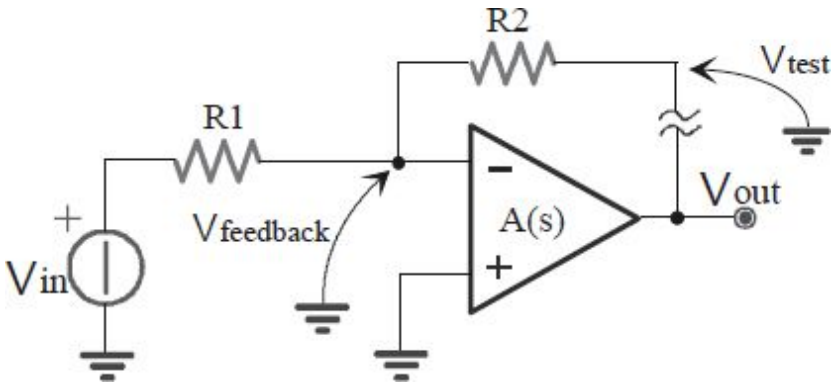


Fig. 3.21: Inverting configuration.

3.3.4 Inverting Configuration

We will examine the frequency response of the OpAmp in inverting configuration (Fig. 3.21). Proceeding as what has been done in the previous section, we calculate the quantities needed for the frequency response study of the circuit. It is found that:

$$\beta = \frac{v_{feedback}}{v_{test}} = \frac{R_1}{R_1 + R_2}$$ (the same result as the N.I. configuration)

$$G_{loop} = -A(s)\cdot\frac{R_1}{R_1 + R_2} = -A\cdot\beta$$ (the same result as the N.I. configuration)

Note that the forward gain $\tilde{A}$ is different from that of the case analyzed in the previous paragraph, where it coincided with the A(s) of the OpAmp. In fact, the following expression is found for the ideal gain:

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

What has been said above implies that, in the Bode plot, the real gain will not be consistent with the lower of A(s) and 1/β(s) since in this case, the closed-loop gain is reduced by a factor $R_2/(R_1+R_2)$. This is clearly visible in Fig. 3.22. Hence, here it is shown that this rule is not always effective and therefore should be avoided.

Let us see, in general, how to answer the following three questions:

- In what frequency range does the stage have well feedback?
- The stage is stable or unstable? And what about its PM?
- What is the Bode plot of the real gain of the stage?

As mentioned in §3.3.2, to answer the first two questions, only the feedback loop is taken into account. Hence, we plot A(s) of the OpAmp and 1/β(s) of the feedback block, regardless of what the forward gain $\tilde{A}$ is. The frequency range where |A(s)| is much greater (let us say by at least a factor of 10) than |1/β(s)| will be where the stage behaves ideally, i.e. it has well feedback. The frequency f^* where the intersection of |A(s)| and |1/β(s)| occurs is shown with a vertical wavy line. This is the frequency at which the stage ceases to have well feedback, and, then, all the considerations concerning the ideal gain and all feedback benefits (the increase in the impedance along the branches of the loop, the reduction in the impedance values at the nodes of the loop, the increase in the accuracy of the gain, etc.) do not apply anymore.

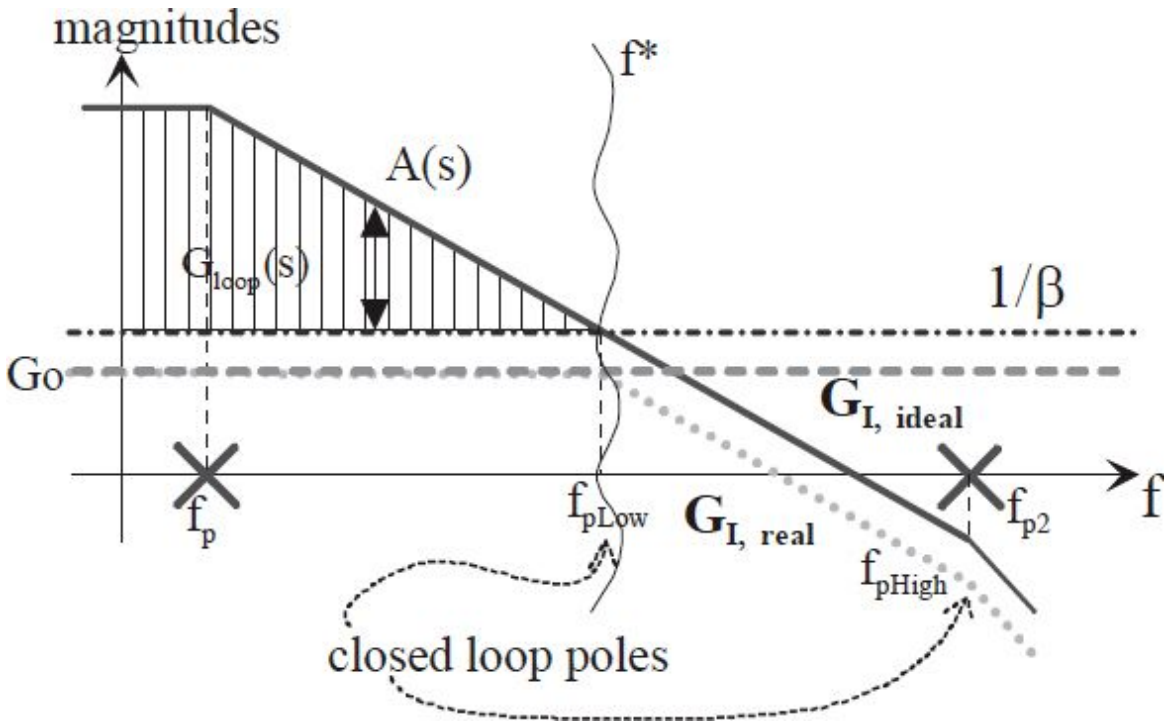


Fig. 3.22: Magnitude Bode plots of the inverting configuration.

To answer the question about stability, just look at the two closure angles between A(s) and 1/β(s) before and after f^*. Observations made in §3.3.2 apply.

Finally, to answer the last question, namely plotting the frequency response of the closed-loop real gain, just proceed in this way. "Forget" A(s) and 1/β(s) graphs as if they had never been plotted and plot the ideal closed-loop gain of the stage instead or what would be expected by the theory if the OpAmp were ideal (infinite A_0 and perfect virtual ground). In the case of the circuit depicted in Fig. 3.21, it is $G_{I,ideale}=|-R_2/R_1|$ (see Fig. 3.22).

Well, the real gain will coincide with the ideal one only when there is enough G_{loop}, that is, until the frequency f^*. Beyond this point, the real gain, instead of continuing to follow the ideal gain (constant in this case), "will collapse as A(s) and β(s) collapse". To be more precise, if A(s) is falling with a -20dB/dec slope beyond f*, the real gain will start to decline with an additional -20dB/dec compared to its previous slope. If A(s) were falling or rising with different slopes, the real gain would do the same, **compared to the trend that it previously had** (in this case, the trend that it had was constant with a slope equal to 0dB/dec). That is why in Fig. 3.22, the real gain G_I of the inverting configuration beyond f^* starts to drop first by -20dB/dec and then, in correspondence of the second pole f_{p2} of the OpAmp, by -40dB/dec. This means that the first pole of the closed-loop gain (real) does not coincide with any actual pole of the OpAmp, in fact it is due to the fact that the feedback is not effective anymore. On the other hand, the second pole coincides almost perfectly with the second pole of the OpAmp. The word "almost" comes from the fact that the coincidence is almost perfect in the bilogarithmic scale of the Bode plot while in the linear scale of the root locus of Fig. 3.18, the slight discrepancy is more detectable.

Of course, if also β(s) decreases or increases with a certain slope (which is not the case in this example), the real gain would be influenced and, compared to the trend that it previously had, its slope will change of the same angle of the β(s).

Desiring to be more precise, we can improve the plot of the real gain around f^*. In the case of stable configuration (closure angles equal to

20dB/dec before and after f*), in correspondence of f*, the curve should be -3dB lower than the value plotted with the dashed line. In the case of configuration with PM=45° (closure angles are 20dB/dec).

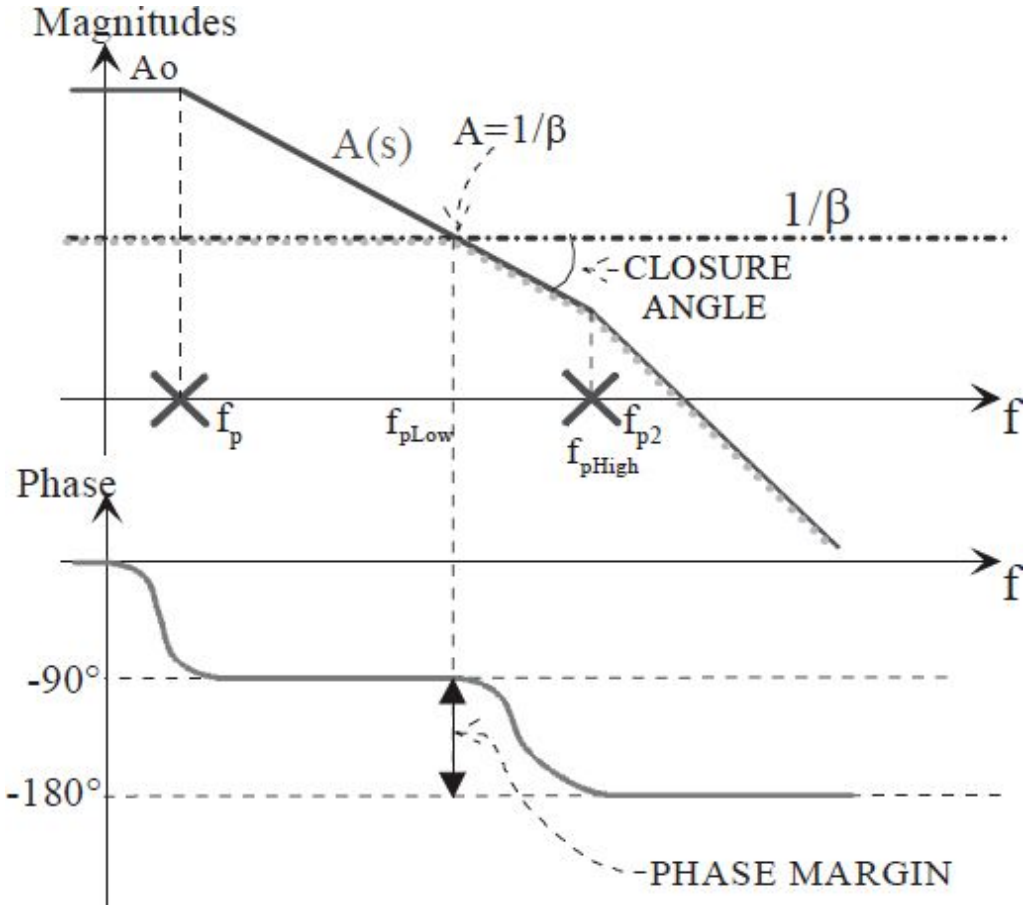


Fig. 3.23: Phase margin and closure angle on Bode plots.

and 40dB/dec before and after f*, or vice versa), a slight peaking in correspondence of f^* should be drawn. On the contrary, in the case of the markedly unstable configuration (closure angles are equal to or greater than 40dB/dec before and after f^*), that peaking should be much more pronounced, as shown in Fig. 3.2.

It may happen that the OpAmp is not compensated, and then it is important to check the stability and the phase margin. Fig. 3.23 shows how to identify these parameters from the plots. If the intersection is right at the slope change between –20dB/dec and –40dB/dec, the PM is equal to 45°. In this particular case, it happens that $f_{plow} \equiv f_{phigh} \equiv f^*$ are coincident with f_{p2}. In

the case of a greater closure angle, as in Fig. 3.24, the first closed-loop pole f_{plow} is at a frequency higher than that of the second OpAmp pole f_{p2}.

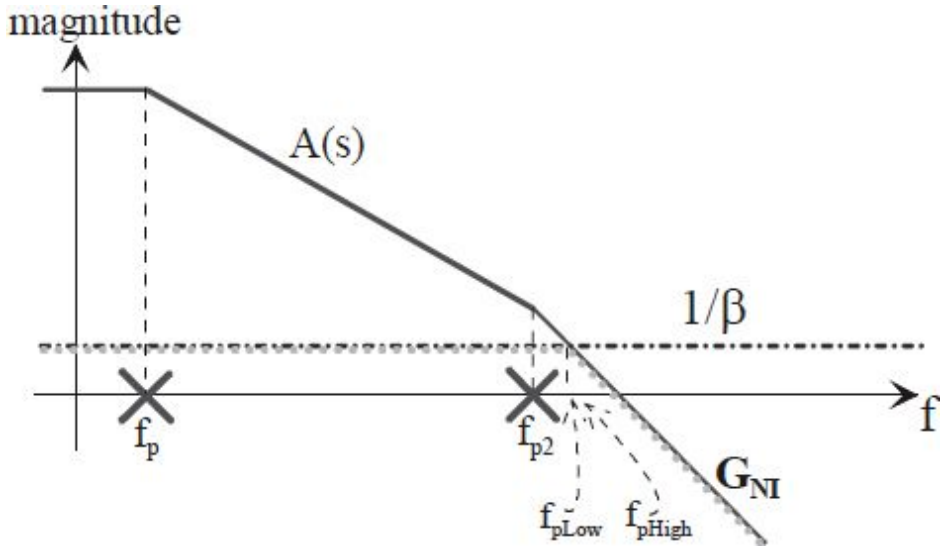


Fig. 3.24: Non-compensated OpAmp where $f_{plow}>f_{p2}$.

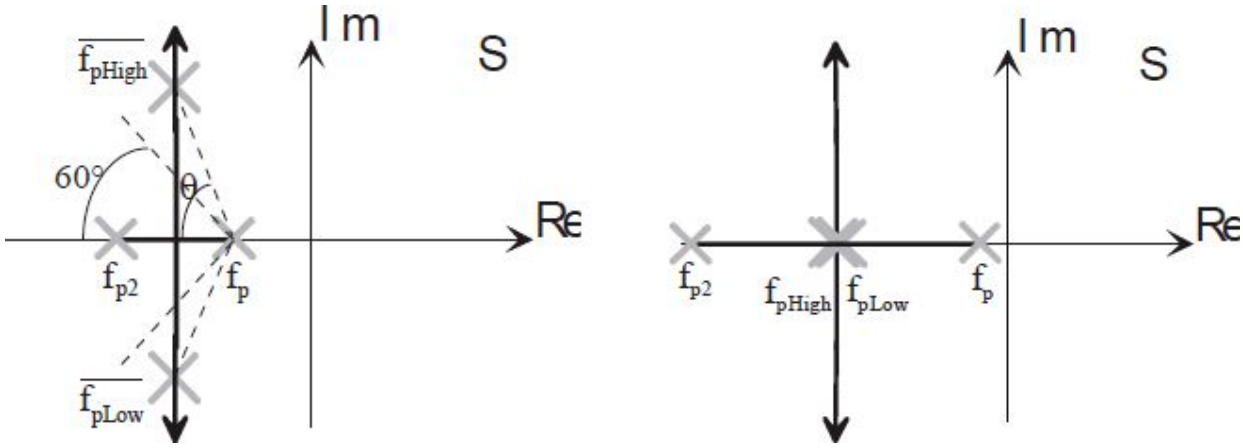


Fig. 3.25: Calibration (i.e. localization of the closed-loop poles) using the root locus for θ>60° (on the left) and for θ=0° (on the right).

It is immediately verified that the phase margin is less than 45°:

$$PM = 90 - arctg\left(\frac{f_{plow}}{f_{p2}}\right) < 45°$$

In fact, the closed-loop poles are complex conjugates (they seem almost coincident on the Bode plot) and are very far from the real axis (i.e. very close to the imaginary axis). As derived earlier, from Fig. 3.25, note that for a PM less than 45°, the angle θ is greater than 60°. To gain more sensitivity on the link between the root locus and the Bode plots, we can derive the PM corresponding to the situation where the closed-loop poles are real and coincident, or rather the condition θ=0°. By a simple calculation, it is found:

$$PM = 180° - 90° - arctg\frac{|f_{plow}|}{|f_{p2}|} \cong 90° - arctg\frac{|\frac{f_{p2}}{2}|}{|f_{p2}|} = 90° - arctg\frac{1}{2} = 90° - 26° = 64°$$

3.4. Effect of the Feedb Capacitance

3.4.1 Non-Inverting Configuration

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:

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

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 $1/zero = -C\cdot R_2$.

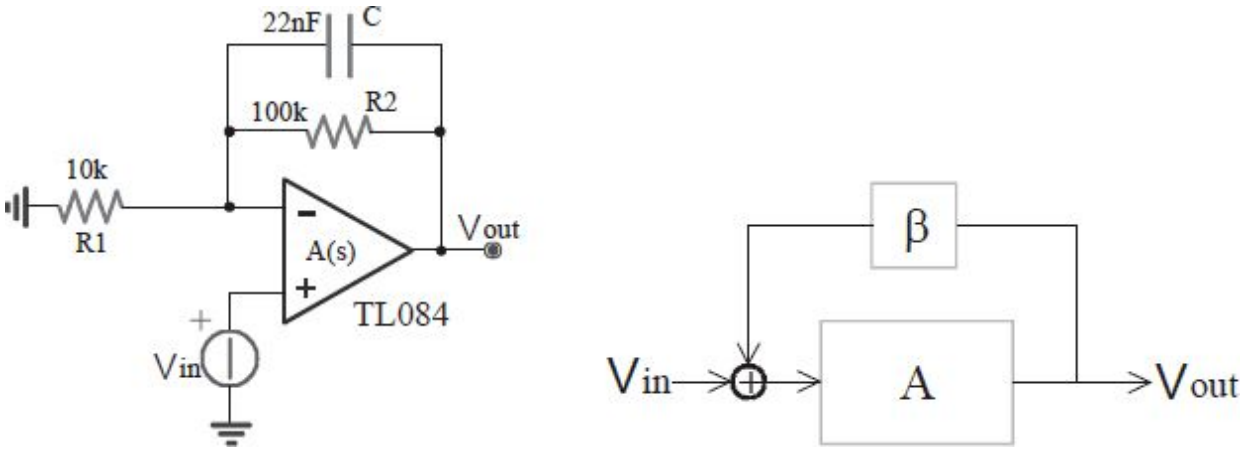


Fig. 3.26: Non-inverting circuit with capacitance on the feedback branch and the system model.

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).

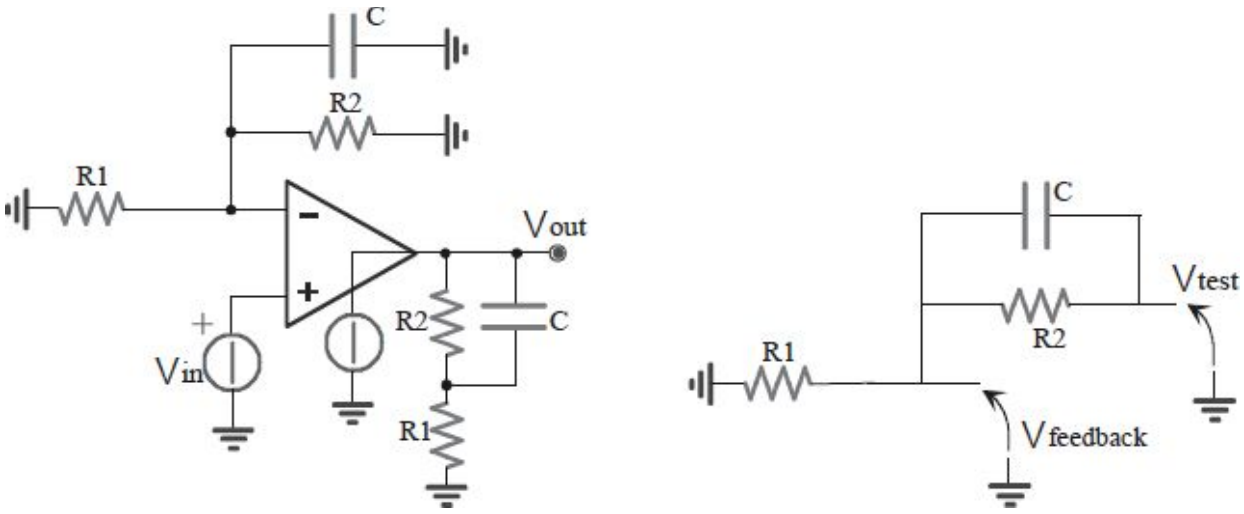


Fig. 3.27: Circuits for the calculation of forward (left) and feedback (right) blocks.

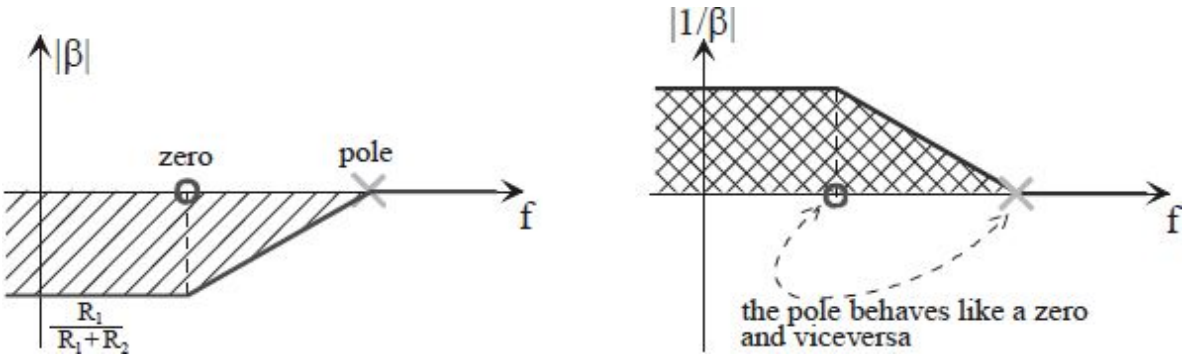


Fig. 3.28: Bode plots of β(s) and 1/β(s).

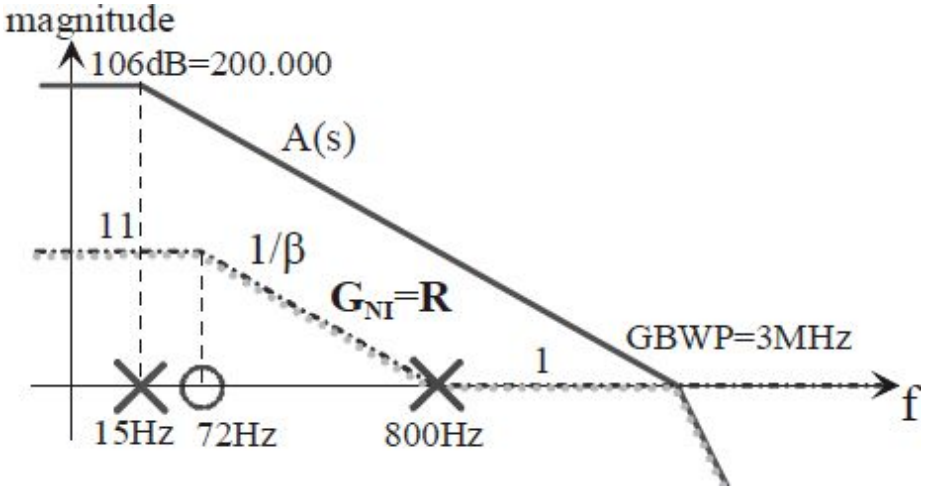


Fig. 3.29: Bode plots of the different transfer functions.

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 11·72=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 $G_{loop}>>1$, on the root locus, the closed-loop pole migrates towards the open-loop zero (Fig. 3.30). It is interesting to note that, for $f\rightarrow\infty$ (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.

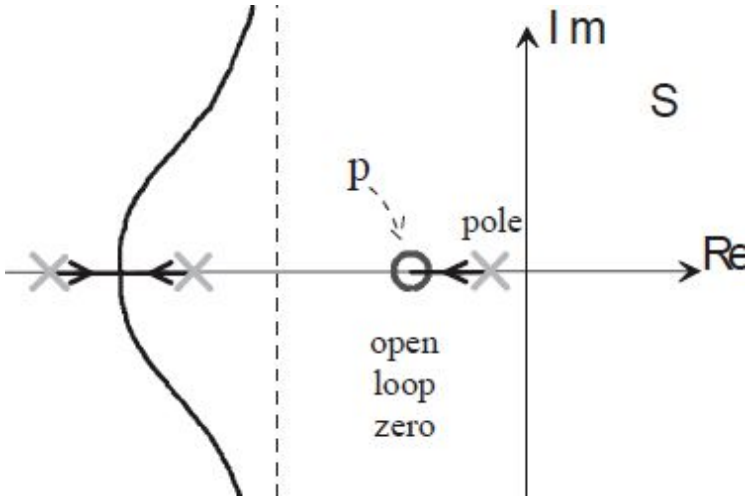


Fig. 3.30: Root locus: the pole tends to the zero.

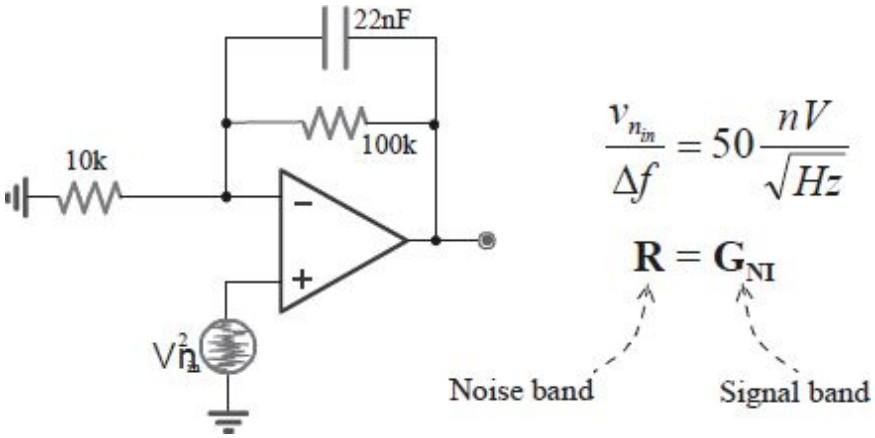


Fig. 3.31: Equivalent circuit for noise performance study.

Let us analyze the circuit in terms of **noise**. For our study, we can think of a noise generator applied to the positive terminal of the OpAmp (Fig. 3.31). Thus, even for the noise, the circuit is still the non-inverting configuration.

So, in this configuration (NI for the signal), the bandwidth of the signal coincides with the bandwidth of the noise. It is easy to verify the following expressions:

area with high gain

$$v_{n\,out}^2 = \frac{v_{n_{in}}^2}{\Delta f} \cdot (11)^2 \cdot 1{,}57 \cdot 72Hz = (5{,}8\mu V)^2$$

area with unity gain

$$v_{n\,out}^2 = \frac{v_{n_{in}}^2}{\Delta f} \cdot 1^2 \cdot 1{,}22 \cdot 3MHz = (100\mu V)^2$$

Numbers 1.57 and 1.22 that appear in the expressions just computed represent the multiplicative factors to derive the equivalent noise bandwidth

in the case of a system with one and two poles, respectively. We note that the component of high frequency noise is higher due to the wider bandwidth.

3.4.2 Inverteting Configuration

Having examined the non-inverting configuration, let us study the frequency response of the operational amplifier in inverting configuration. In what follows, we refer to the circuit depicted in Fig. 3.32. Let us proceed again in two ways: the laborious method that uses $\tilde{A}$ and the other one proposed above, which instead looks only at A(s) and 1/β(s). We begin to calculate the transfer function of the forward and feedback blocks, $\tilde{A}$ and β:

$$\beta = \frac{R_1}{R_1 + \left(R_2 \left\| \frac{1}{sC}\right.\right)}$$ (same equation of N.I. case)

$$\tilde{A} = -\frac{R_2}{R_2 + R_1} \cdot \frac{1}{1 + sC\left(R_1 \| R_2\right)} \cdot A(s)$$ (different equation from N.I. case)

The expression of the real closed-loop gain is:

$$G_{I,real}(s) = \frac{\tilde{A}(s)}{1 - \beta \cdot A(s)} \quad \text{con} \quad A(s) \neq \tilde{A}(s$$

Desiring to make calculations, one finds:

$$G_{I,real}(s) = \frac{G_{op}}{1 - Gloop} = \frac{\tilde{A}(s)}{1 + \beta \cdot A(s)} = \frac{-\frac{R_2}{R_2 + R_1} \cdot \frac{1 \cdot A(s)}{1 + sC\left(R_1 \| R_2\right)}}{1 + \frac{A(s) \cdot R_1}{R_1 + R_2 \left\| \frac{1}{sC}\right.}} =$$

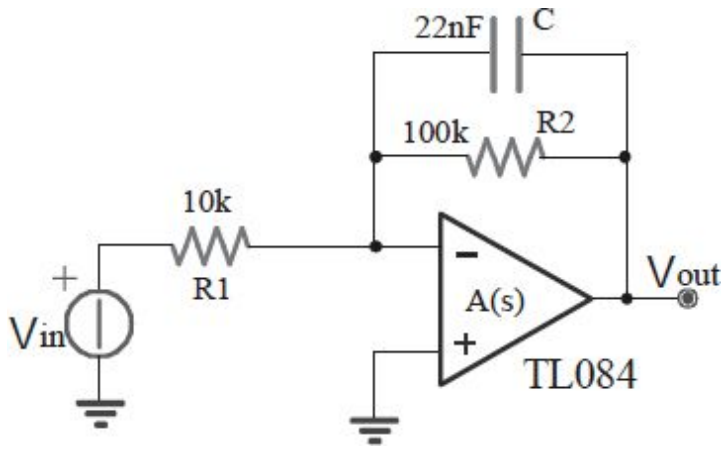


Fig. 3.32: Operational amplifier in inverting configuration.

$$= \begin{cases} -\dfrac{R_2}{R_1}\cdot\dfrac{1}{1+sCR_2} & \text{for low frequencies, i.e. Gloop>>1} \\ \dfrac{R_2}{R_2+R_1}\cdot\dfrac{A(s)}{\left(1+sC\left(R_1\|R_2\right)\right)} & \text{for high frequencies, i.e. Gloop<<1} \end{cases}$$

To proceed in a graphical way, we should also plot:

$$\tilde{A}(s) \qquad \text{(high frequency response)}$$

$$\frac{\tilde{A}(s)}{\beta\cdot A(s)} \qquad \text{(low frequency response)}$$

In this case, we will have to plot the following transfer functions that represent the trend of the ideal gain at various frequencies:

$$\tilde{A}(s)=-\frac{A(s)\cdot R_2}{R_2+R_1}\cdot\frac{1}{\left(1+sC\left(R_1\|R_2\right)\right)} \qquad \text{(high frequency)}$$

$$\beta=\frac{R_1}{R_1+\dfrac{R_2}{sCR_2+1}}=\frac{R_1}{R_2+R_1}\cdot\frac{\left(1+sCR_2\right)}{\left(1+sC\left(R_1\|R_2\right)\right)}$$

$$\frac{\tilde{A}(s)}{\beta\cdot A(s)}=\frac{R_2}{R_1}\cdot\frac{1}{1+sCR_2} \qquad \text{(low frequency)}$$

Fig. 3.33 shows these plots together with that of G_I.

The same conclusion can be reached easier by using the proposed method, without appealing to the concept, moreover not so evident, of $\tilde{A}$. It is enough to plot A(s) and 1/β(s) to find the frequency f^* at which $|G_{loop}=1|$. We can note that this is equal to the OpAmp GBWP and that the configuration is stable with a phase margin equal to 45°. Then, we plot the ideal gain: at low frequencies (capacitor open circuit), it is equal to $|-R_2/R_1|=10$; at the infinite

frequency (C short circuit), it is equal to 0, i.e. -∞dB. The pole will coincide with the zero of β at 72Hz. The real gain will be equal to the ideal gain only up to f^*=3MHz. Beyond this frequency, it will collapse, compared to the trend that it previously had (-20dB/dec), as A(s) collapses (-40dB/dec), leading to -60dB/dec (see Fig. 3.33).

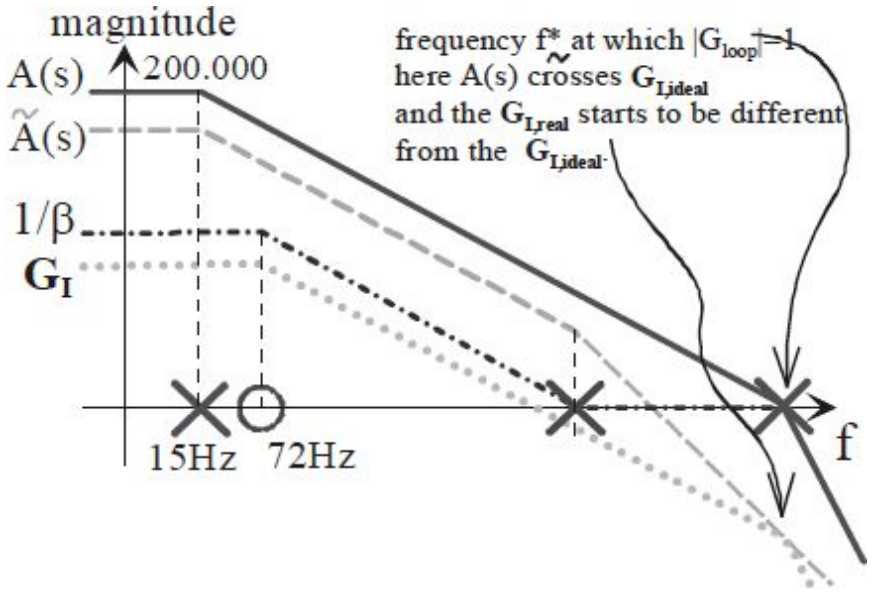


Fig. 3.33: Magnitude Bode plot of the different transfer functions.

Let us see what happens now in terms of noise. At low frequencies, it is as if:

$$|G_I| = \left|-\frac{R_2}{R_1}\right| = \left(1+\frac{R_2}{R_1}\right)-1 \cong \left|\frac{1}{\beta}\right|-1$$

At high frequencies, on the other hand, the gain G_I tends to become small due to the increasing feedback until the maximum (β=1) that imposes the signal in the positive OpAmp terminal to its output. Hence, any potential noise affecting this terminal can reach the output with unity transfer. This is the case of the OpAmp noise, which is modeled in this way. Fig. 3.34 shows the plot of noise transfer function. Note that the bandwidth values for the signal and the noise are not the same.

Although the bandwidth has this trend, since 72Hz is only 24ppm compared to the overall bandwidth equal to 3MHz, the noise at the output node is not very large. In fact, assuming an input noise of 50nV/√Hz, the output noise would be equal to only:

$$v_{out_{n\_source}} = \sqrt{\frac{v_s^2}{\Delta f} \cdot 72 \cdot 1{,}57 \cdot (10)^2} = 5\mu V_{rms}$$

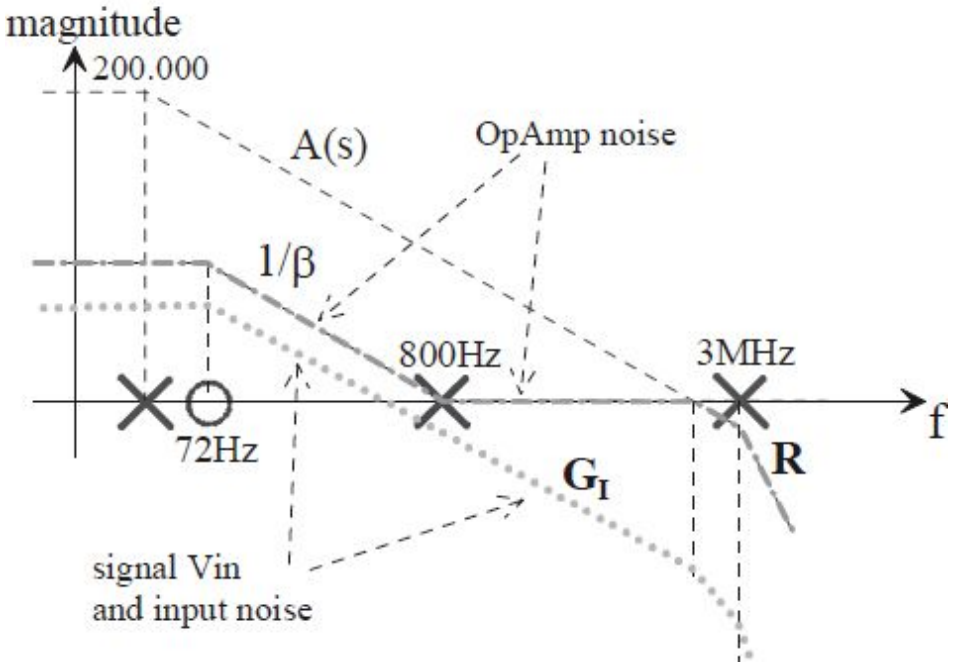


Fig. 3.34: Input source Vin and OpAmp noise transfers towards the output.

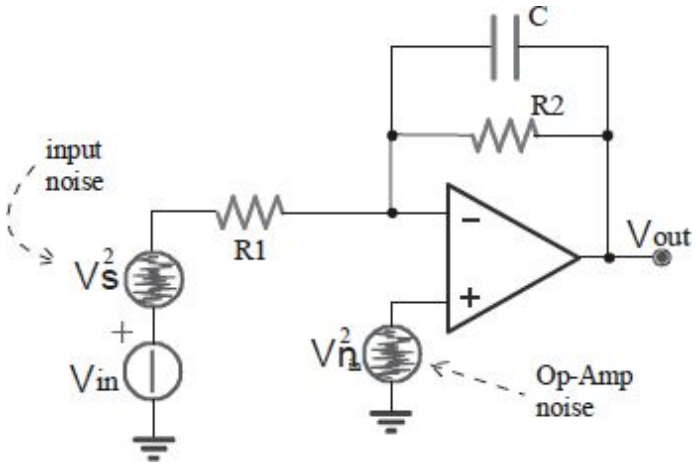


Fig. 3.35: Equivalent noise generators.

Instead, the OpAmp noise at the output would be about:

$$v_{out_{n,op-amp}} = \sqrt{\frac{v_{n,in}^2}{\Delta f} \cdot GBWP 1{,}57 \cdot (1)^2} = 109 \mu V_{rms}$$

In Fig. 3.35, equivalent noise generators discussed above are shown. We must, therefore, pay attention to input noise filtering. To filter this input noise, it is possible to use the same filter used for the signal v_{in}. However, to reduce the OpAmp noise bandwidth, it is worth acting on the internal phase compensation even if also the external compensation is used for the stability and signal filtering. Fig. 3.36 shows how noise bandwidth changes by acting as what has just been said.

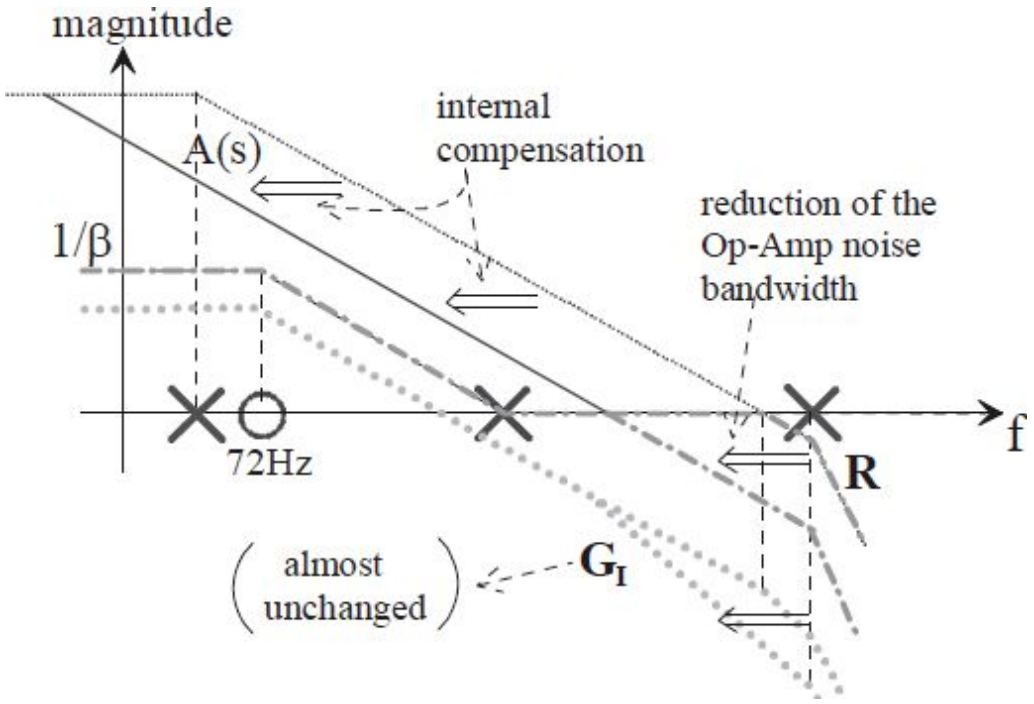


Fig. 3.36: Noise bandwidth reduction.

3.4.3 Integrator

The integrator is a circuit shown in Fig. 3.37. To study the frequency response, we operate as what has been told in the previous paragraphs. We cut the feedback loop and calculate the transfer functions by using the equivalent circuit of Fig. 3.38, on the left. We calculate the forward gain again, more for exercise than for real use, since we prefer the proposed method.

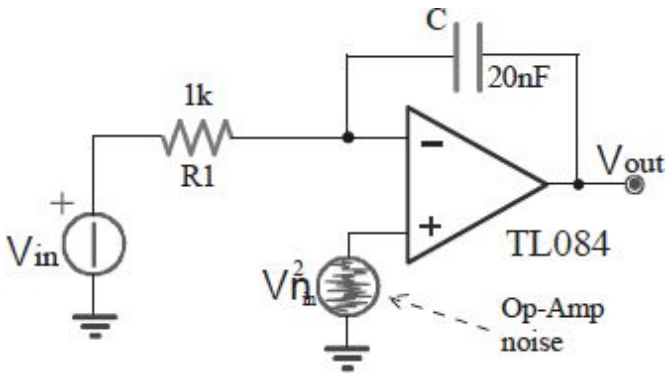


Fig. 3.37: Integrator circuit.

It is found:

$$\tilde{A} = \frac{\frac{1}{sC}}{\frac{1}{sC} + R} \cdot A(s) = \frac{A(s)}{1 + sRC}$$

We use the circuit on the right of Fig. 3.38 to calculate the feedback block:

$$\beta = \frac{R}{R + \frac{1}{sC}} = \frac{sRC}{1 + sRC}$$

Notice that β has a zero at the origin (Fig. 3.39), which implies that the feedback does not act at low frequencies. To find the closed-loop gain, we must proceed as what was done for the inverting configuration. Performing

calculations, we find:

$$G_{INT,\ reale} = \begin{cases} \dfrac{\tilde{A}}{A(s)\cdot\beta} = \ldots = \dfrac{1}{sCR} \\ \tilde{A} = \dfrac{A(s)}{1+sCR} \end{cases}$$

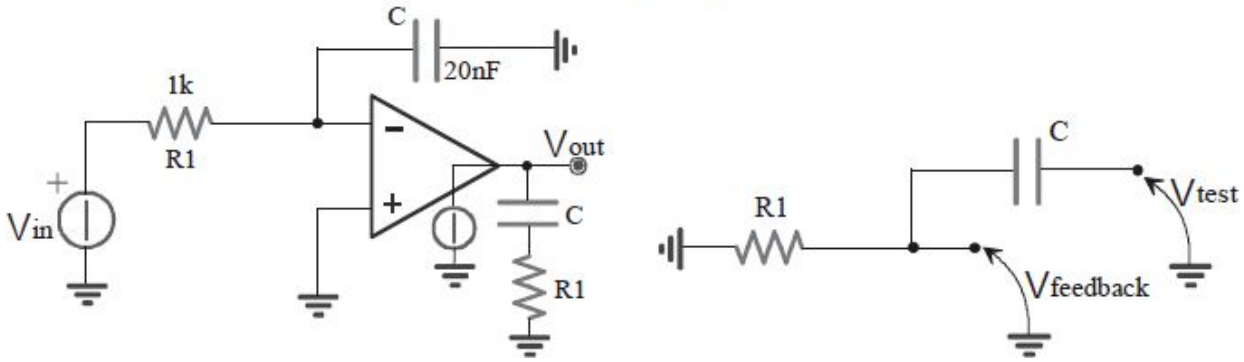


Fig. 3.38: Calculation of forward block gain (on the left) and β (on the right).

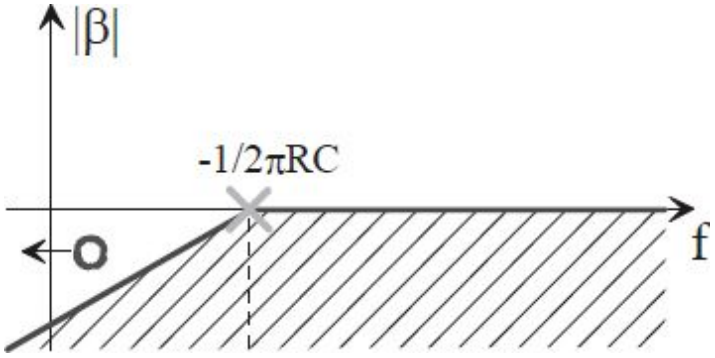


Fig. 3.39: Magnitude Bode plot ofβ.

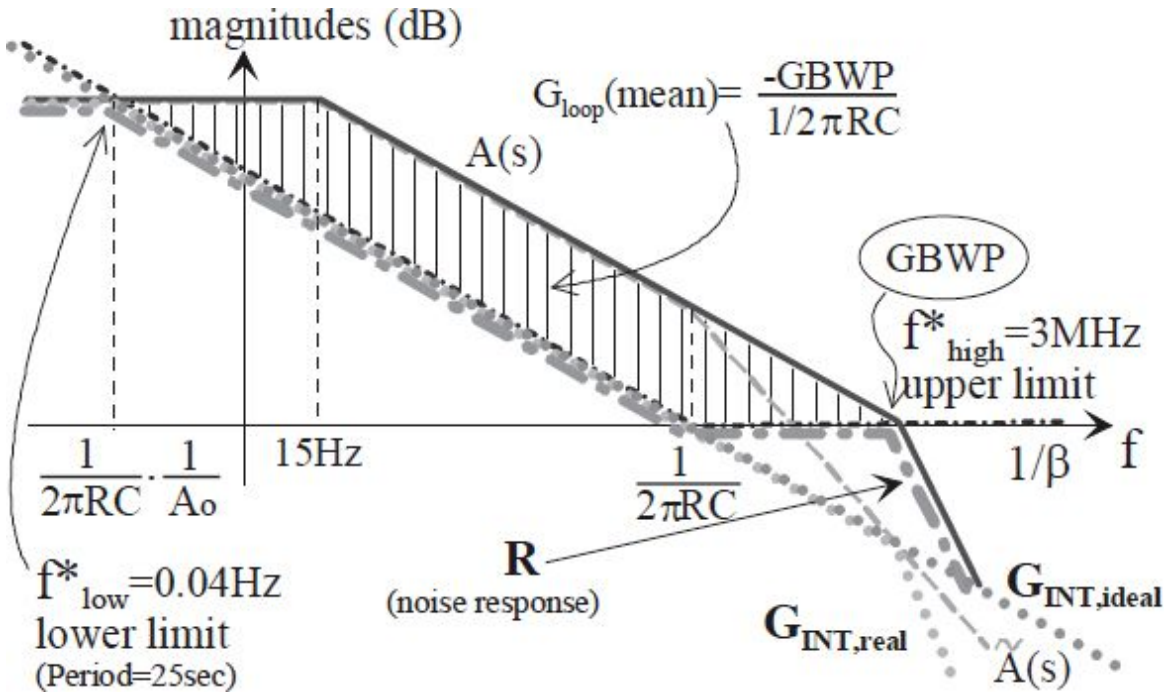


Fig. 3.40: Bode plots for the analysis of the integrator.

The trend of $G_{INT,real}$ is immediately obtainable from the plots of Fig. 3.40. Desiring to proceed with the proposed fast method, simply plot A(s) and 1/β(s). Without resorting to the Laplace transform of β(s), just see that at low frequencies (C open), β is zero (zero at the origin), then 1/β tends to infinity (pole at the origin). On the other hand, at high frequencies (C closed), β is 1, and then 1/β=1. The pole of β (which boosts 1/β) will be 1/RC. For the first time so far, we note that A(s) and 1/β(s) intersect each other in two points, f^*_{low} and f^*_{high}. Only between them has the stage well feedback. At too low frequencies (at worst at the dc frequency), the stage does not have feedback because C is not sufficiently conductive. At too high frequencies, the stage does not have feedback anymore because the OpAmp has a poor gain. At f^*_{low}, the closure angle is good (20dB/dec) and likewise at f^*_{high} (PM=45°): the stage is stable.

Finally, to plot the real gain, we proceed first with the ideal one $G_{INT,ideal}$, which is $-Z_2/Z_1=-1/sRC$; which is ∞ at the dc frequency and 0 at the infinite frequency and intersects the 0dB axis in correspondence of the pulsation ω=1/RC. The real gain $G_{INT,real}$ will coincide with the ideal one in the

medium frequency range. Below f^*_{low}, instead, $G_{INT,real}$ will suffer the rise (heading to the left, i.e. heading towards lower frequency gradually) of 1/β and then will acquire a slope of +20dB/dec, compared to the ideal of -20dB/dec, $G_{INT,real}$ will therefore become constant. The value will be precisely equal to A_0 of the OpAmp because at low frequencies, it is as if the block β were not present (C open), and then the OpAmp works with an open-loop. Above f^*_{high}, instead, $G_{INT,real}$ will undergo the collapse of A(s) with a -40dB/dec slope and, then, will collapse with a -60dB/dec slope (since the ideal gain had already a -20dB/dec slope). The noise bandwidth is greater than the signal bandwidth (**R≠G**), but the effect is negligible due to high gain at low frequencies.

In conclusion, the frequency range in which this circuit behaves as an integrator is between f^*_{low} and f^*_{high}, which is where there is a significant G_{loop}. The frequency dynamics is therefore equal to the distance between $1/A_0RC$ and GBWP, i.e.:

$$\text{DYNAMICS} \cong \log_{10}\left[A_0 \cdot \left(\frac{GBWP}{\frac{1}{2\pi RC}}\right)\right] \approx 8 decades$$

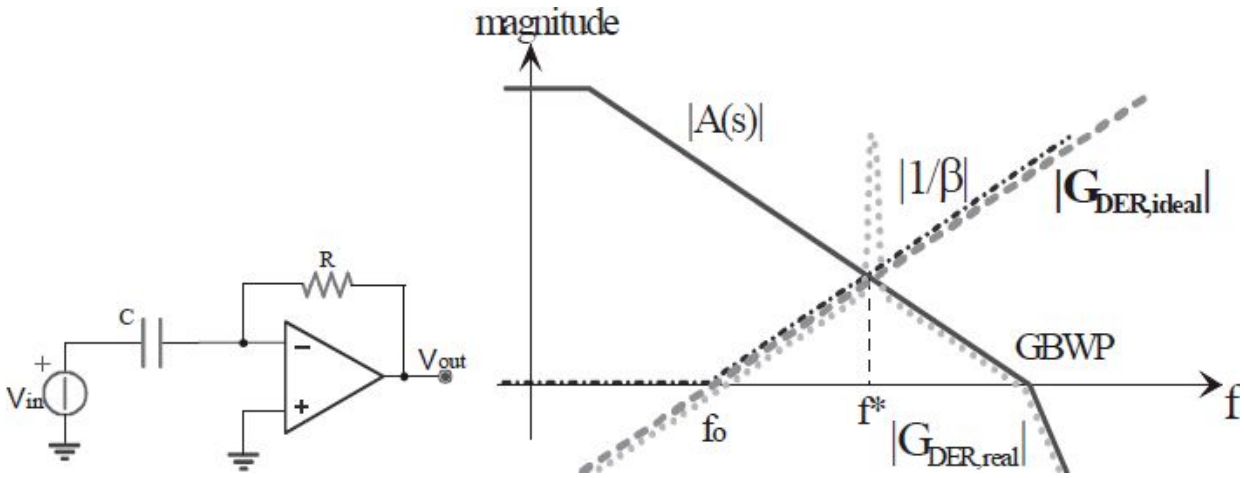


Fig. 3.41: Non-compensated derivator (on the left) and its Bode plot (on the right).

3.4.4 Derivator

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

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

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

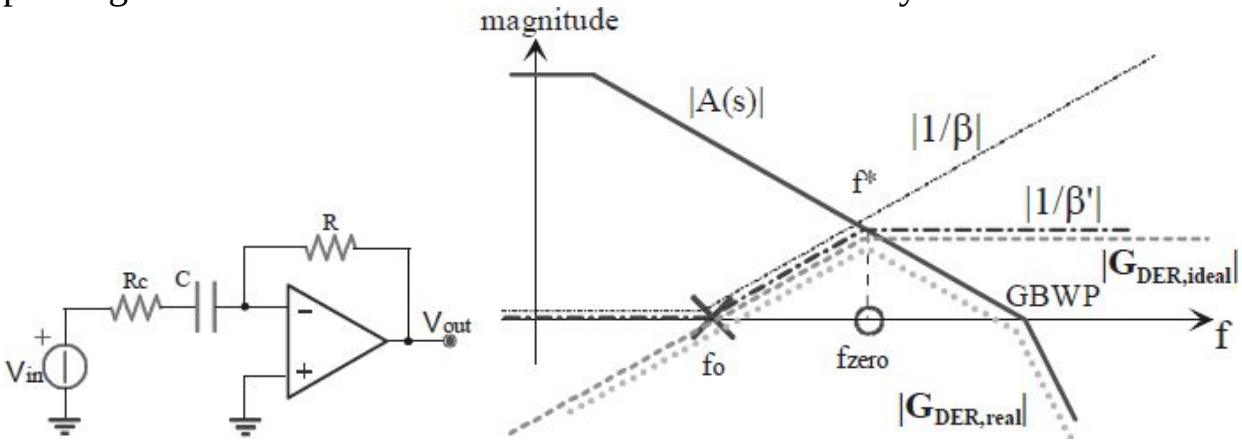


Fig. 3.42: Derivator compensated (on the left) thanks to the addition of R_C and its Bode plot (on the right).

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. $0 \div f_i$) and the frequency f_0 at which the gain should be 0dB: ▪ choose an OpAmp with GBWP$\geq 100 \cdot f_i$ and place $f^*=f_{zero} \geq 10 \cdot f_i$;

▪ if the source has $R_C>100\Omega$, choose $C=1/2\pi R_C f^*$ (making sure that $C>>C_{stray}$ of the OpAmp); otherwise choose $C>>C_{stray}$ and then $R_C=1/2\pi C f^*$; ▪ finally, obtain $R=1/2\pi C f_0$.

The following figures show the plots of some SPICE simulations performed on the derivator. For these simulations, an LM101A has been chosen with f_i=1kHz, C=16nF, and R=4.7kΩ. Fig. 3.43 demonstrates the Bode plot of the magnitude of G. Fig. 3.44 shows the step response of the non-compensated derivator while Fig. 3.45 is the same response, but this time for a compensated derivator. The compensation is the Miller one, with C_C=30pF.

3.5. Effect of the Input Capacitance

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 $G \neq 1/\beta$. To calculate it, we can proceed with the method that uses $\tilde{A}$,

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

$$G = \begin{cases} -\dfrac{R_2}{R_1} & \text{for } |Gloop| >> 1 \\ -\dfrac{A(s) \cdot R_2}{R_2 + R_1} \cdot \dfrac{1}{\left(1 + \dfrac{s}{pole}\right)} & \text{for } |Gloop| << 1 \end{cases}$$

obtaining:
and finally:

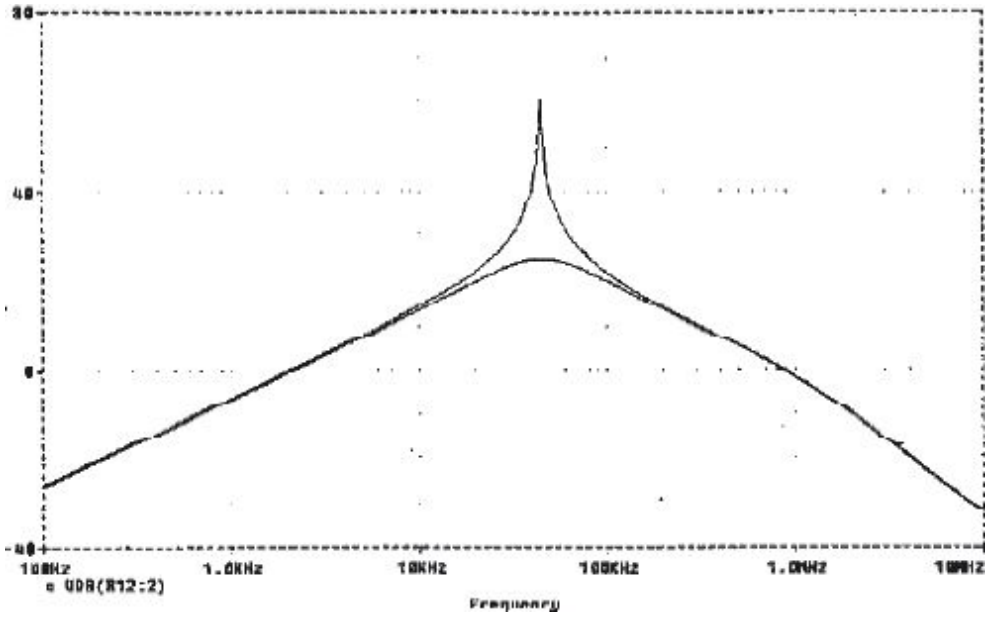


Fig. 3.43: SPICE simulation of the Bode plot of G.

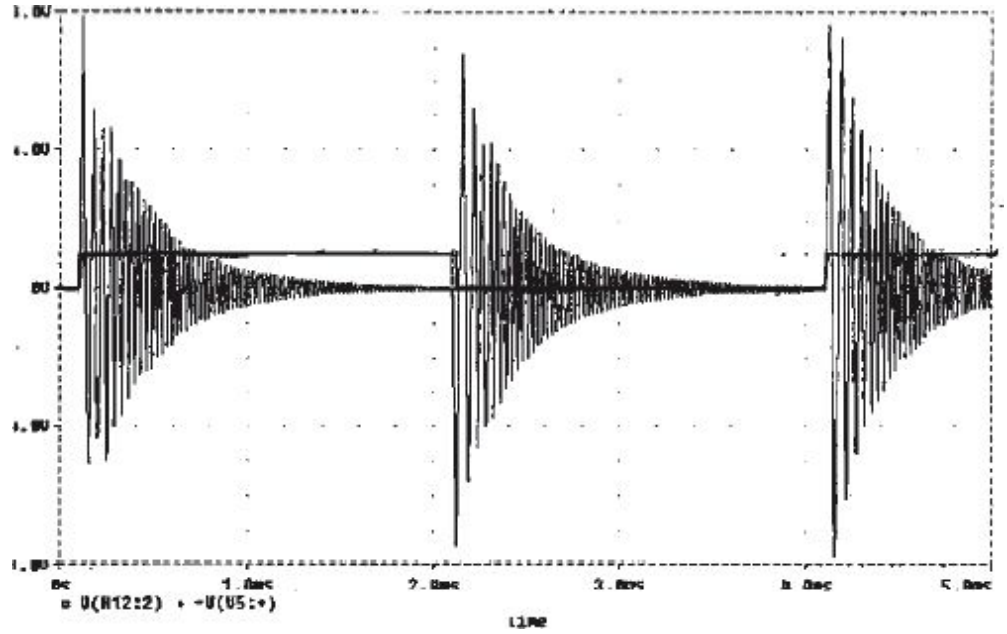


Fig. 3.44: SPICE simulation of the step response of the non-compensated derivator.

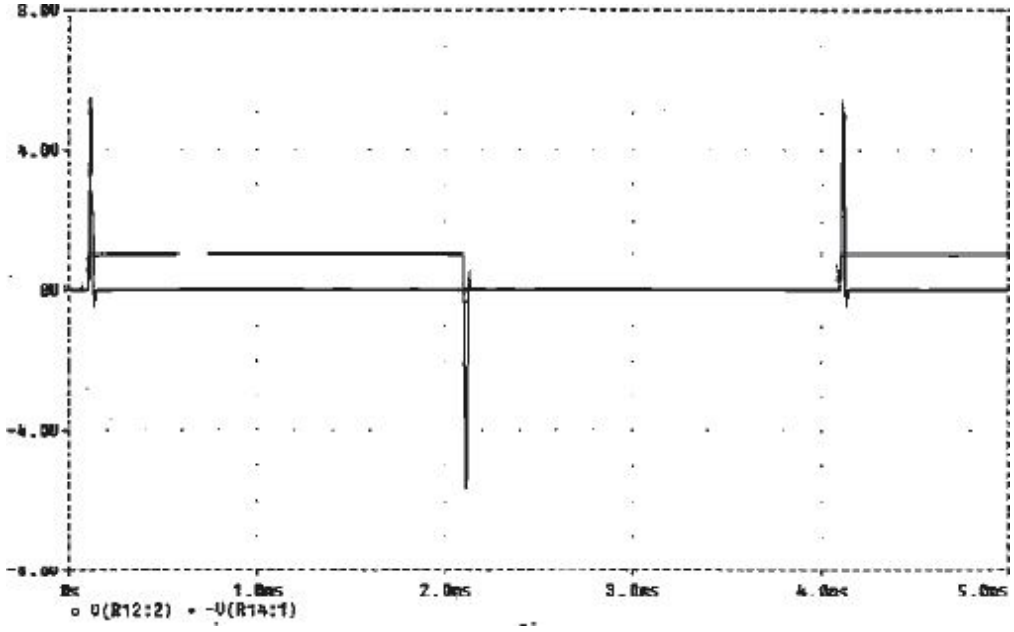


Fig. 3.45: SPICE simulation of the step response of the compensated derivator.

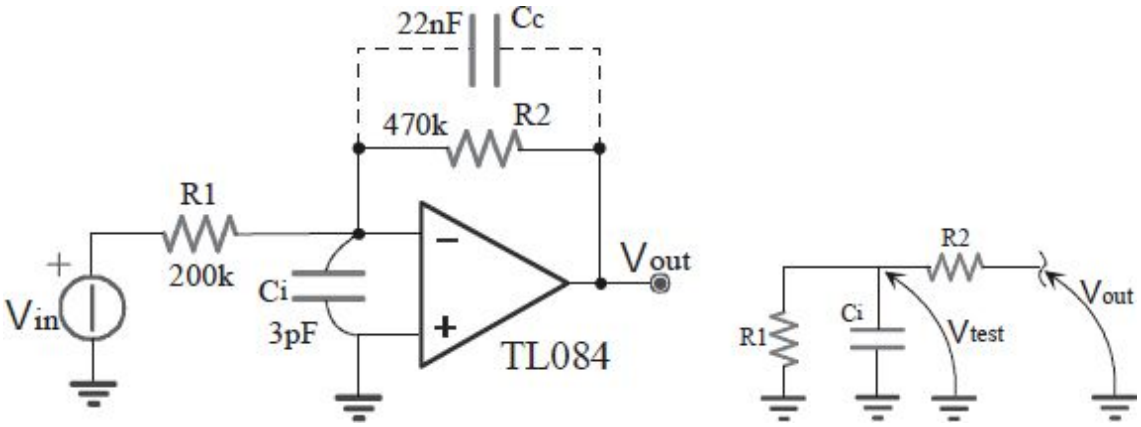


Fig. 3.46: Stage with input stray capacitance (left) and circuit for β(s) calculation.

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, $\beta=0$, then $1/\beta=\infty$. 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.

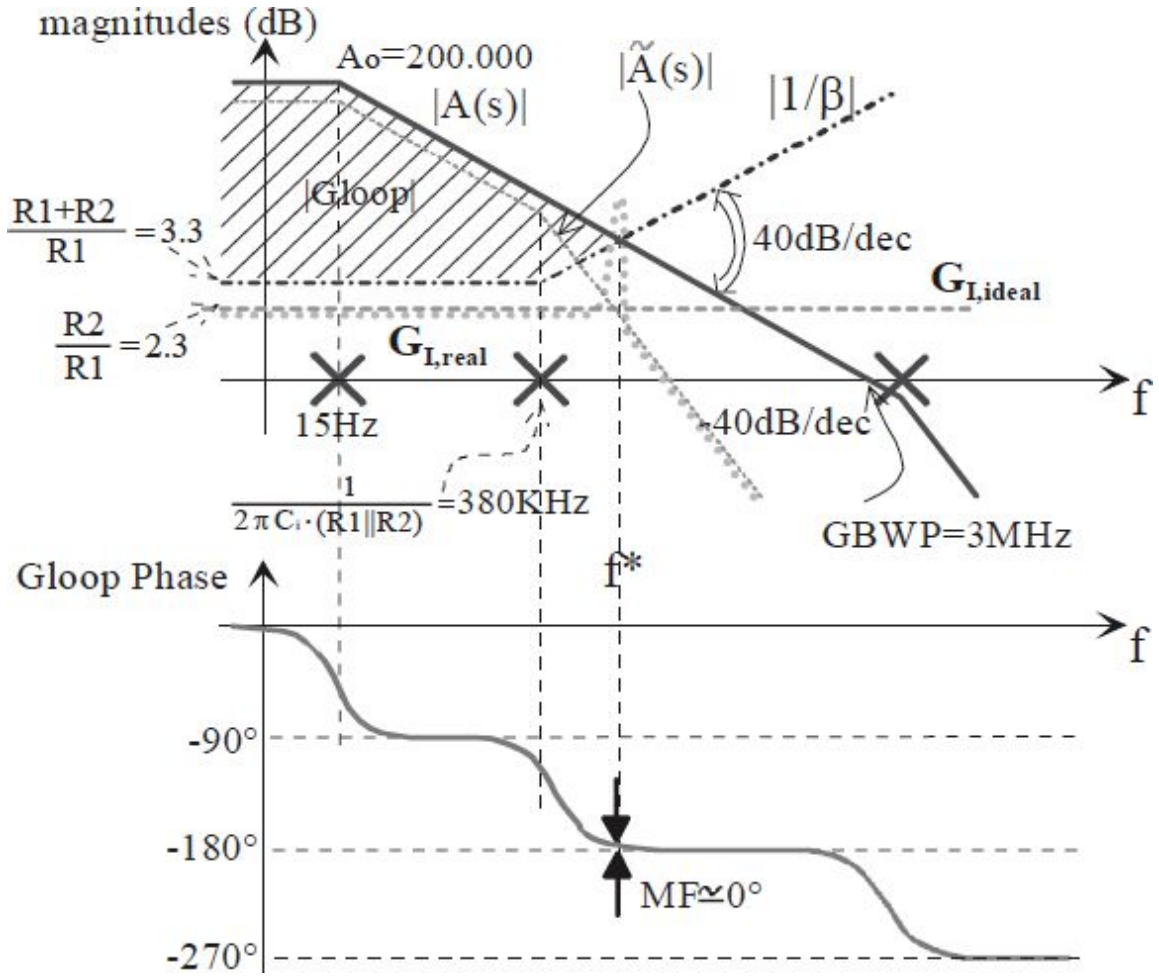


Fig. 3.47: Bode plots of the responses of interest and of the G_{loop} phase for the unstable circuit.

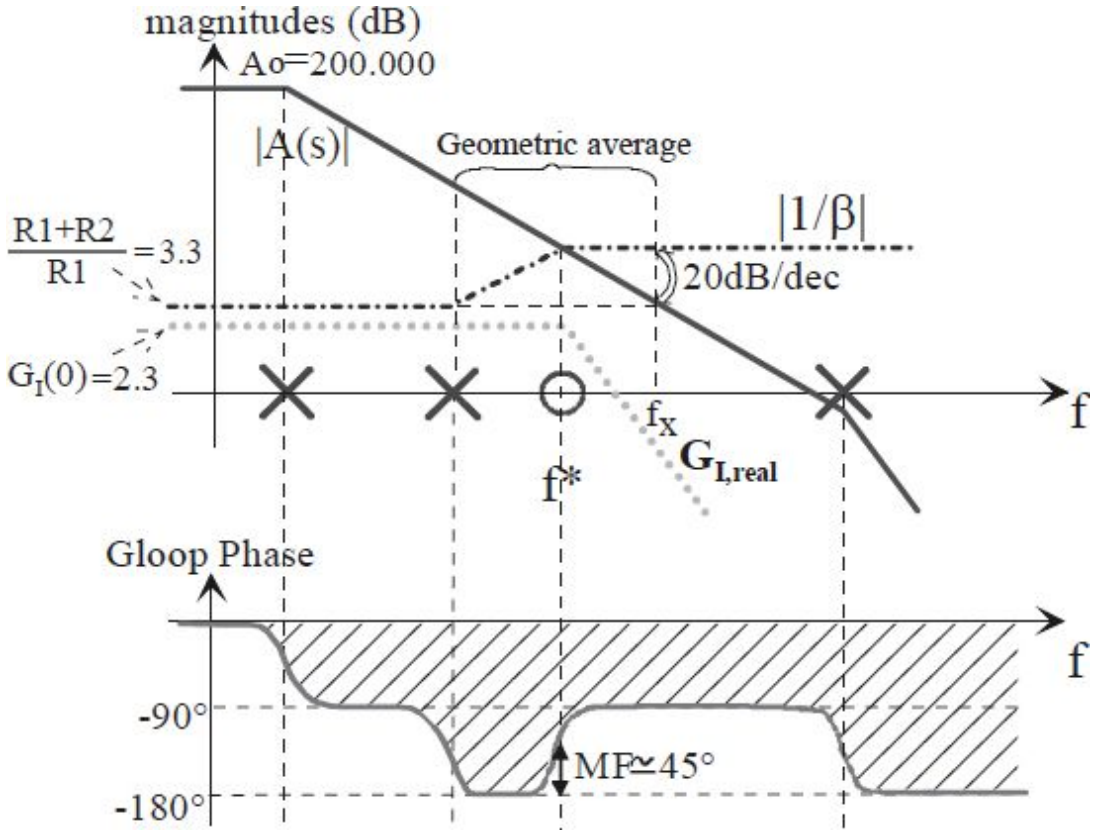


Fig. 3.48: Bode plots for the compensated circuit.

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=$\sqrt{(\text{pole}\cdot f_C)}$=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 $C_C << C_i$ 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 $|G_{loop}|>>1$. 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.

3.5.1 Example 1

Let us compute the critical value of C_i that makes the circuit not sufficiently stable (PM≤45°). With simple considerations, we find that this occurs when: $\frac{1}{C_i \cdot (R_1 \| R_2)} \equiv \frac{2\pi \cdot GBWP}{3.3}$

that is, for $C_i \geq 1.2pF$, the circuit is not sufficiently stable anymore.

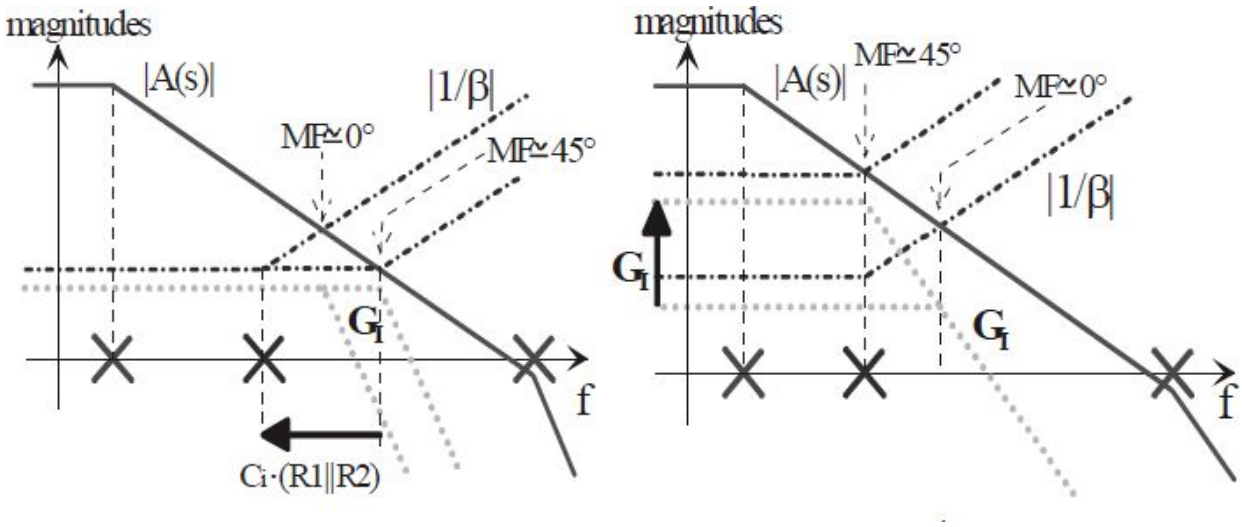


Fig. 3.49: Magnitude Bode plots as C_i (on the left) and R_1 (on the right) change.

3.5.2 Example 2

Given $C_i=3pF$, let us compute the values of the resistors R_1 and R_2 that ensure the complete stability (PM≥45°). The maximum feedback occurs for:

$$\frac{2\pi \cdot GBWP}{\frac{1}{C_i \cdot (R_1 \| R_2)}} \leq \frac{R_1 + R_2}{R_1}$$

With R_2=470k, we have $(R_1+R_2)/R_1 \geq 5.1$; that is, $G_I \geq 4.1$ and R_1=113kΩ. This means that the resistance must have low values. at the expense of increased consumption even if this causes more problems for the OpAmp in terms of the output current (which increases) and the input impedance (which decreases). Fig. 3.49 shows how Bode plots change if C_i and R_1 are modified.

The followings are the results of a SPICE simulation performed on the circuit depicted in Fig. 3.50, which uses an OPA627. The values of the components used in the simulation are as follows: R_1=47kΩ, R2=470kΩ, C_C=1,5pF, C_{stray}=15pF, and GBWP=16MHz.

The following figures demonstrate the SPICE simulations' results: note the significant difference between the step responses of compensated and non-compensated circuits.

Let us proceed with analyzing more examples of compensation circuits used in practice.

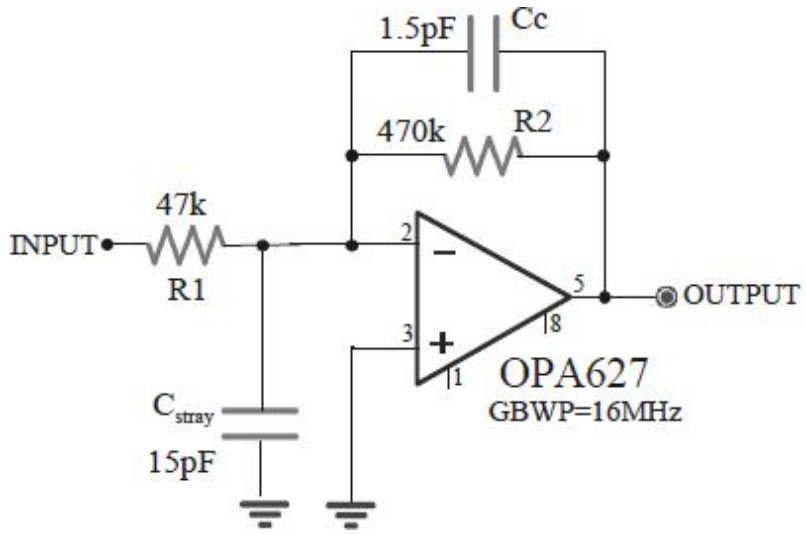


Fig. 3.50: Compensation for an input C_{stray} or a big feedback resistor.

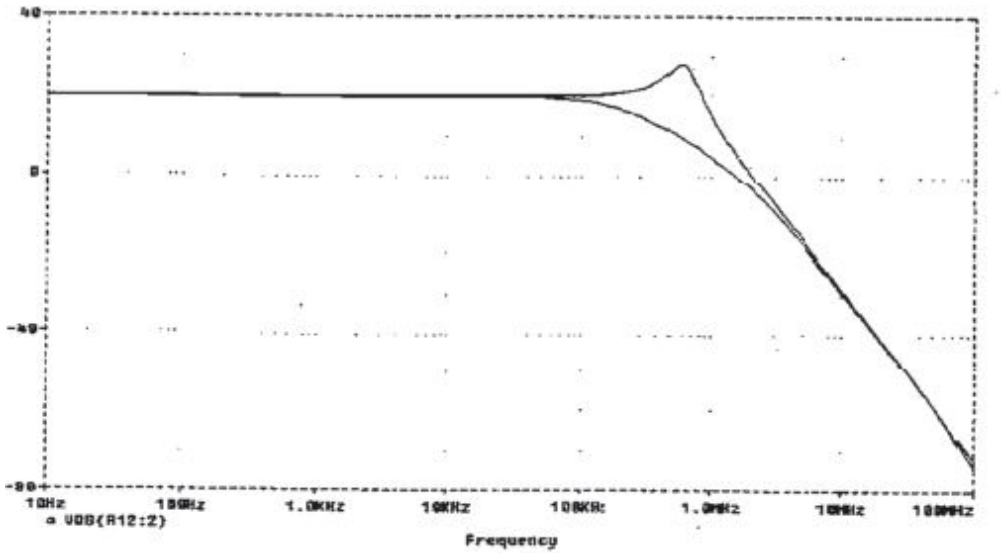


Fig. 3.51: Magnitude Bode plots: SPICE simulation.

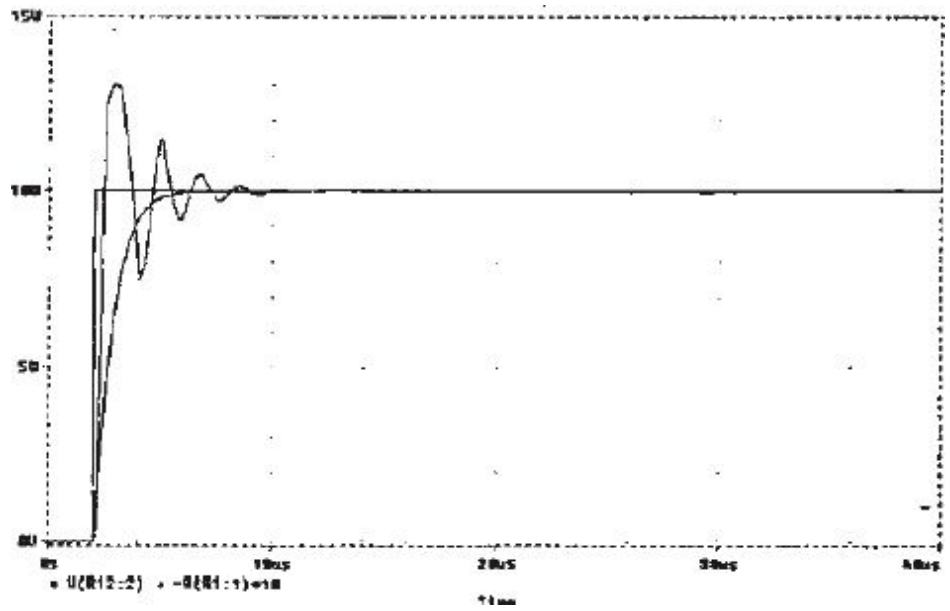


Fig. 3.52: Step response: SPICE simulation.

3.5.3 Compensation of an Inverting Buffer

Given C_i=3pF, we seek the values of R_1=R_2 to have an inverting buffer stable (note that this is the worst case); it is sufficient to impose the following condition: $\frac{2\pi \cdot GBWP}{\frac{1}{C_i \cdot (R_1 \| R_2)}} \leq \frac{R_1 + R_2}{R_1} = 2$

that is, $R_1 \| R_2 \leq 17k\Omega$; we can choose, for instance, $R_1=R_2=33k\Omega$. Fig. 3.53 shows one stable and one unstable circuit of this kind.

3.5.4 Compensation of a Photodiode Amplifier

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.

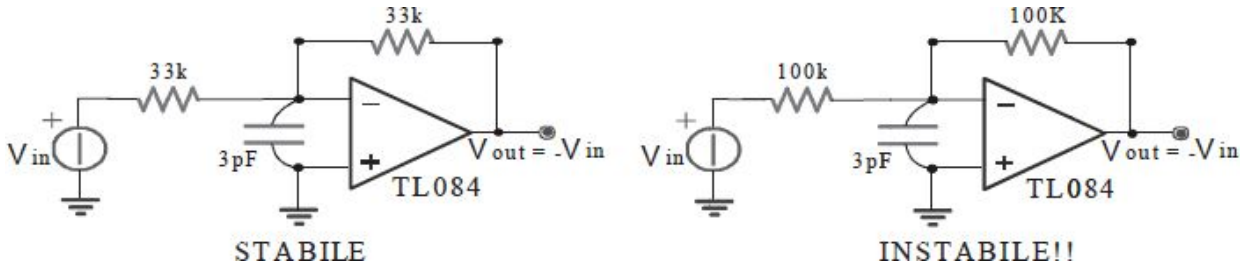


Fig. 3.53: Stable (on the left) and unstable (on the right) circuits, only different in resistance values.

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: $$G_{loop}(s) = -A_d(s) \cdot \frac{1+sRC_C}{1+sR(C_D+C_i+C_C)}$$
As can be seen, it is impossible to try to perfectly elide the pole with the zero, so the only thing that can be done is to choose C_C to reduce the closure angle between A_d and $1/\beta$. The zero must be placed at the frequency $f_{zero} = f_C = \sqrt{GBWP \cdot 580Hz} = 42kHz$ so that the closure angle is reduced, as shown in Fig. 3.55.

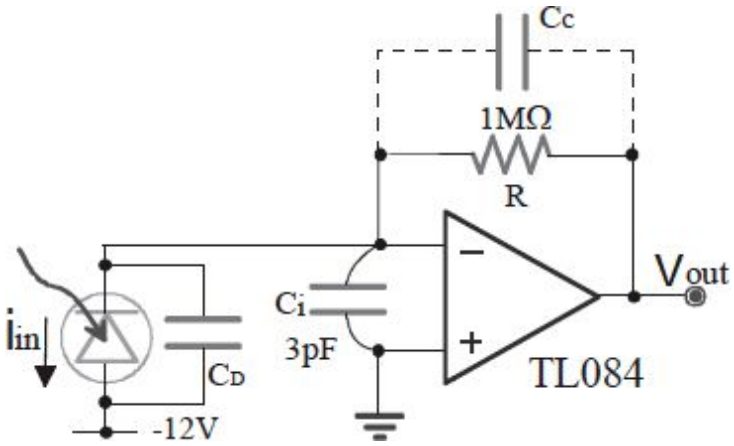


Fig. 3.54: Photodiode amplifier.

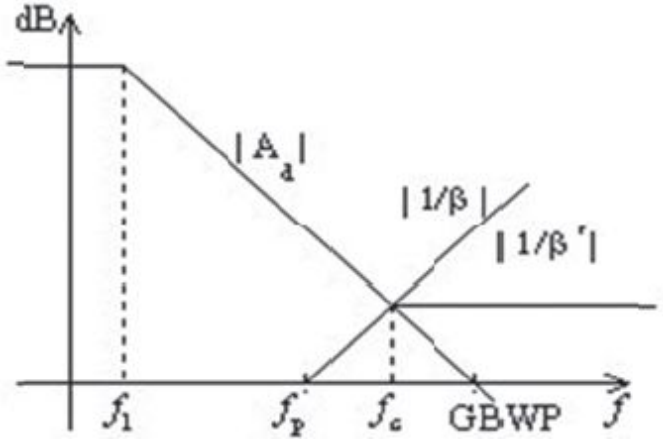


Fig. 3.55: Magnitude Bode plot.

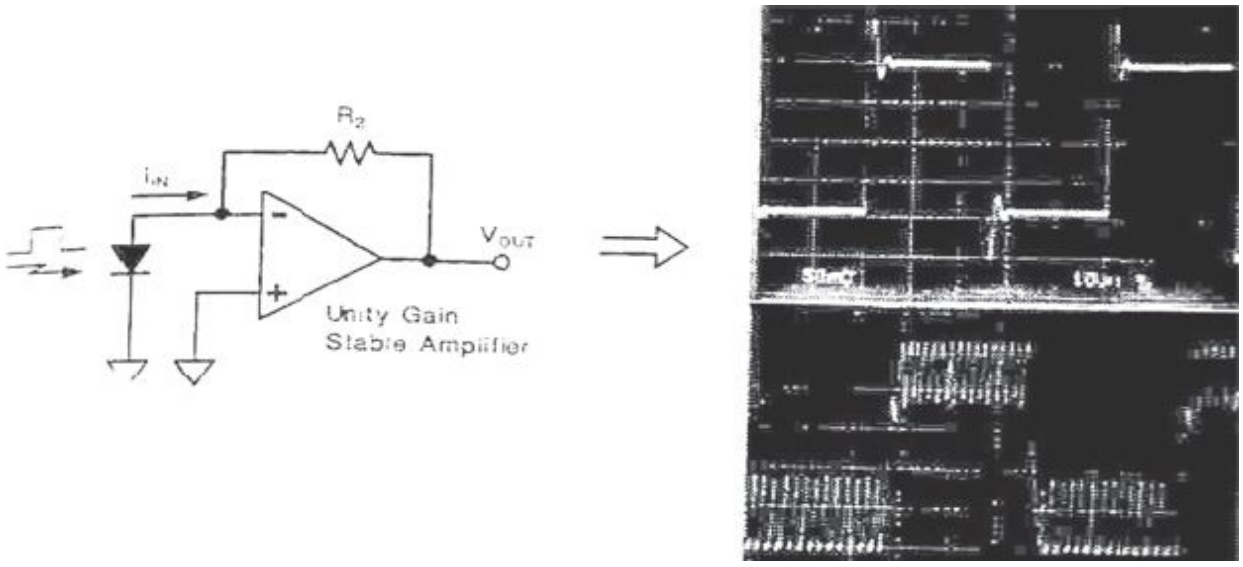


Fig. 3.56: Visible effects of not-perfect stability on a compensated OpAmp, which are shown on an oscilloscope.

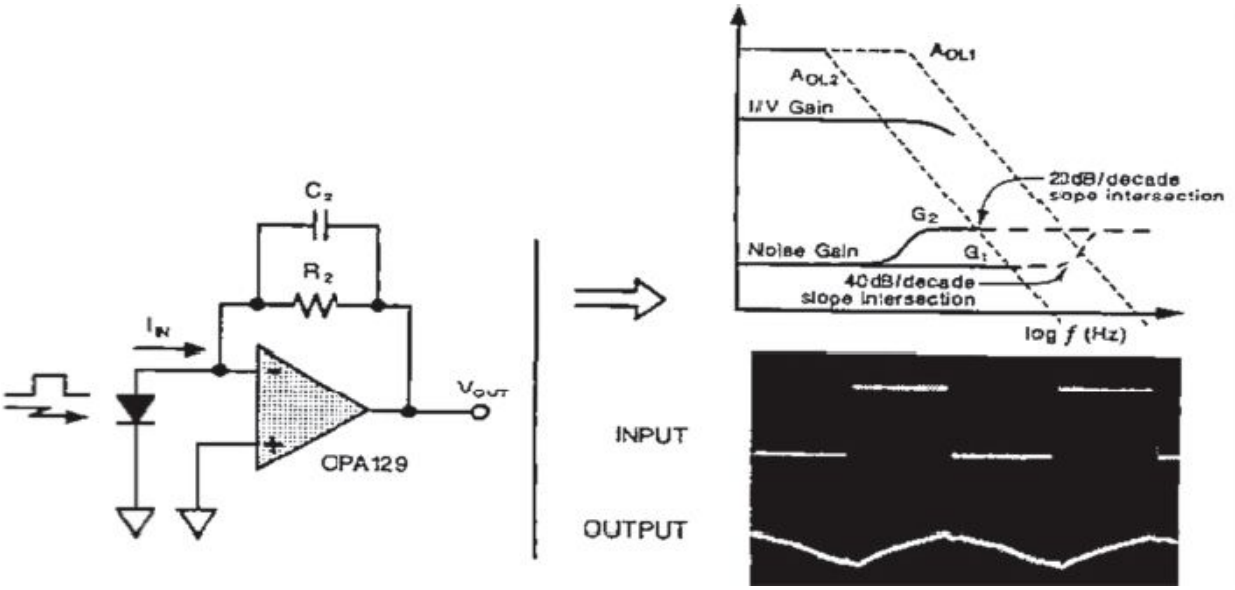


Fig. 3.57: Reduction in the output signal, which results from the bandwidth reduction caused by excessive compensation.

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.

There are some available dedicated integrated circuits which include both the photodetector and the amplifier OpAmp. They are called OEIC (Opto Electronic Integrated Circuits) (Fig. 3.58). Since these are monolithically integrated in silicon, the responsivity of the detector will be that typical of silicon, which is excellent for red wavelength, but unsuitable for working in

the second or third fiber optics communication windows (1.3 and 1.55μm), as shown in Fig. 3.59.

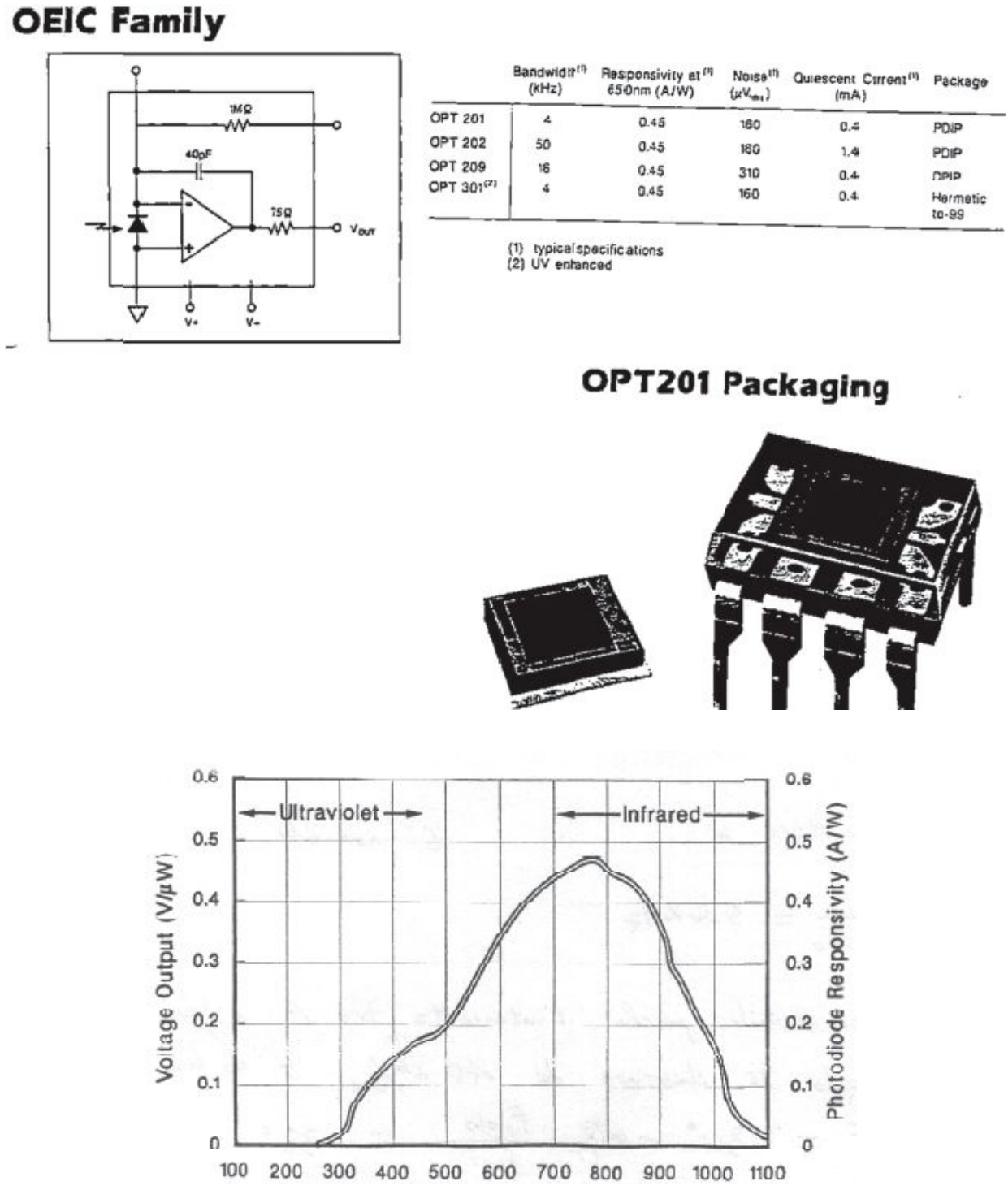


Fig. 3.58: Example of OEIC (Opto Electronic Integrated Circuit).

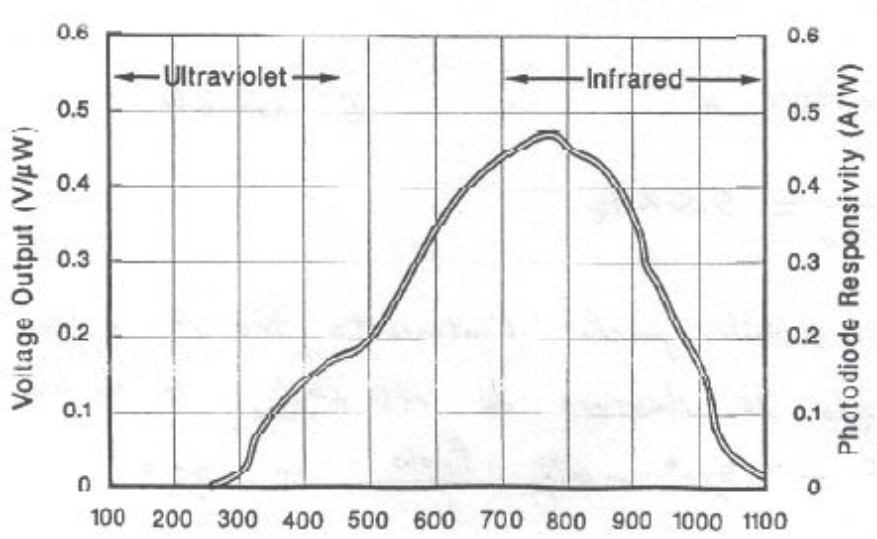


Fig. 3.59: Responsivity of the photodiode integrated in OPT201 (λ in nm).

3.6. Effect of Output Resistance and Load Capacitance

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.

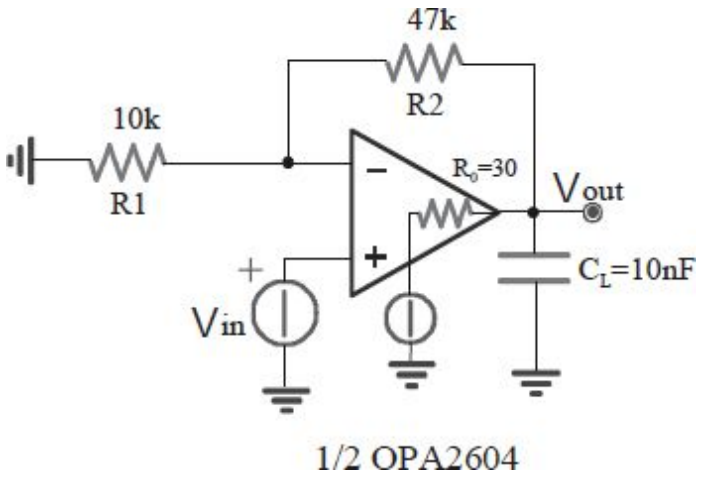


Fig. 3.60: Circuit with high load capacitance.

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 A*(s) as the cascade of A(s) and the pole:

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

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 A*(s) instead of A(s). We note that the system is unstable since the intersection between 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:

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

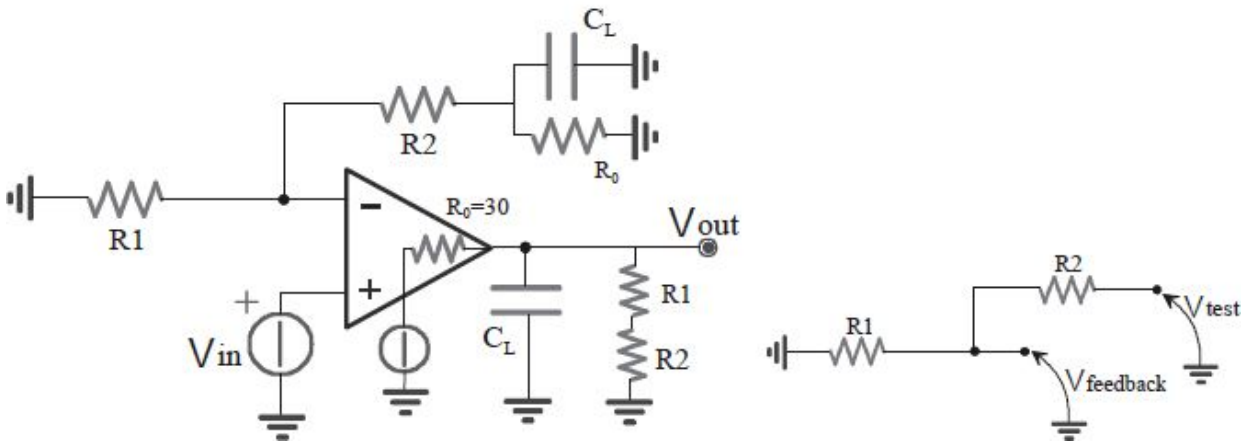


Fig. 3.61: Circuits for the calculation of the equivalent gain A(s) (on the left) and the feedback β(s) (on the right).*

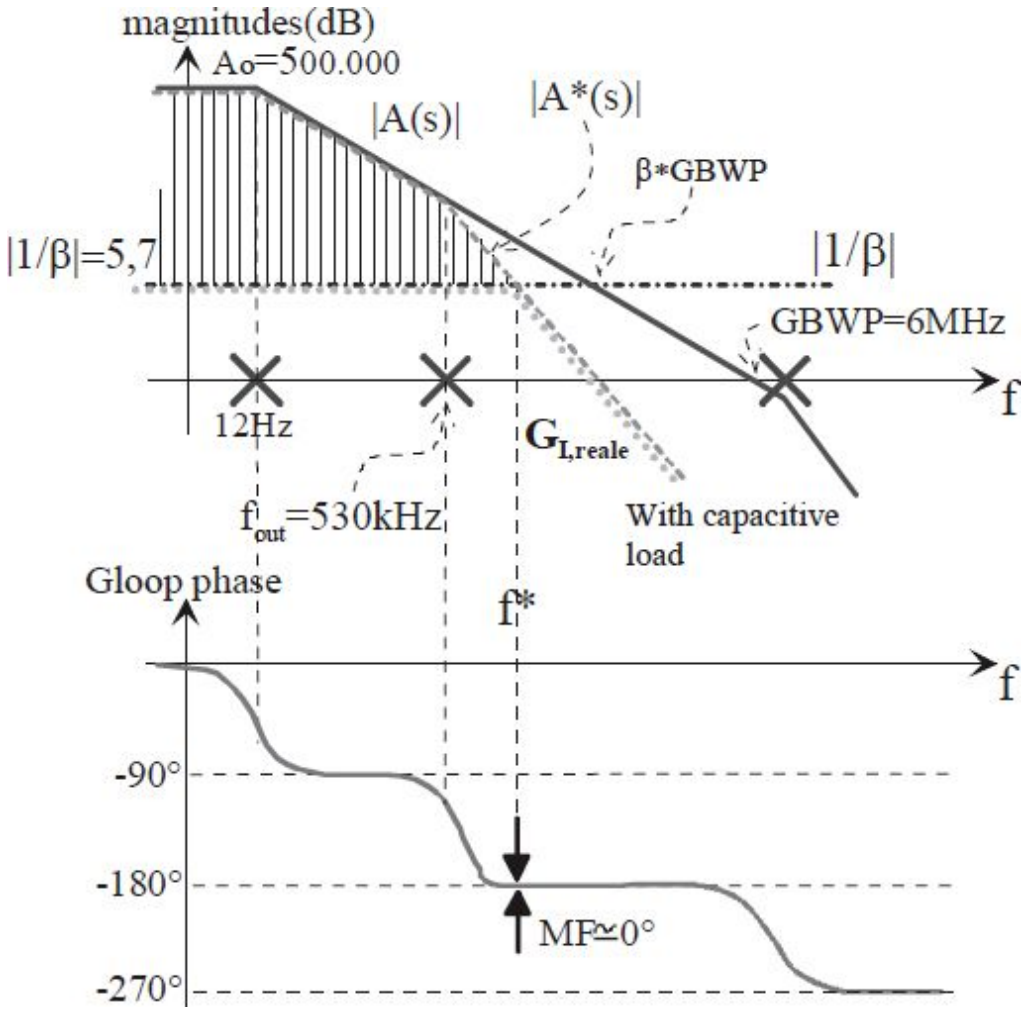
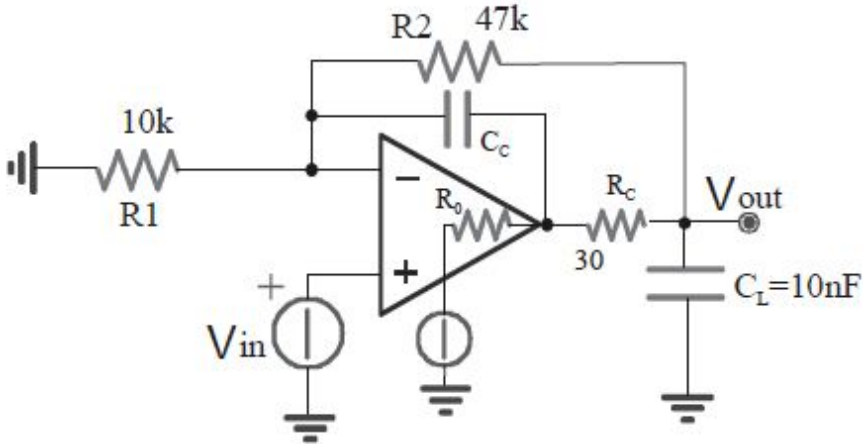


Fig. 3.62: Magnitude Bode plots of the transfer functions of different blocks.

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,

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



resulting in:

Fig. 3.63: Compensated circuit.

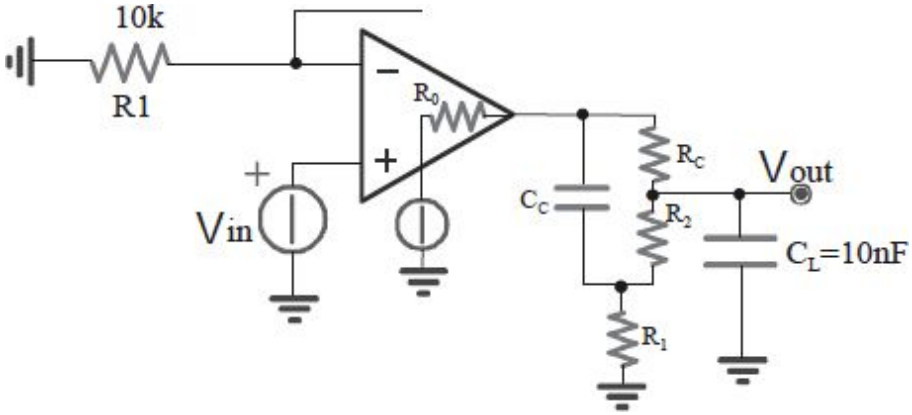


Fig. 3.64: Circuit for the calculation of the equivalent gain A(s).*

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.

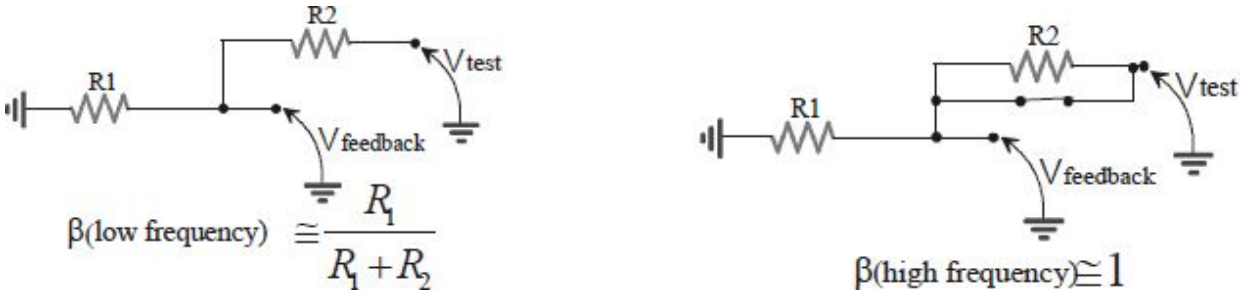


Fig. 3.65: Feedback block at low (on the left) and high frequencies (on the right).

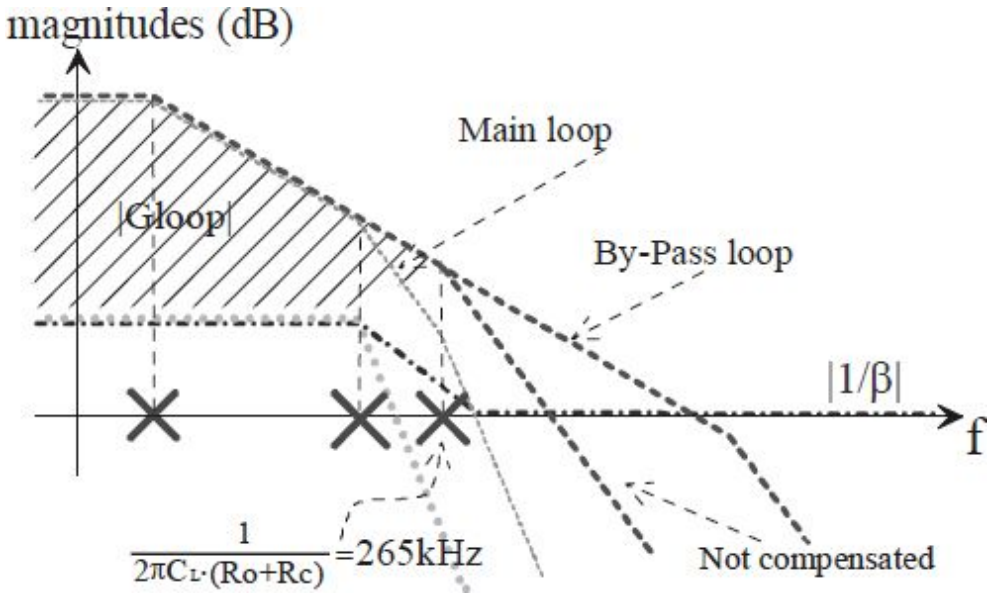


Fig. 3.66: Magnitude Bode plots.

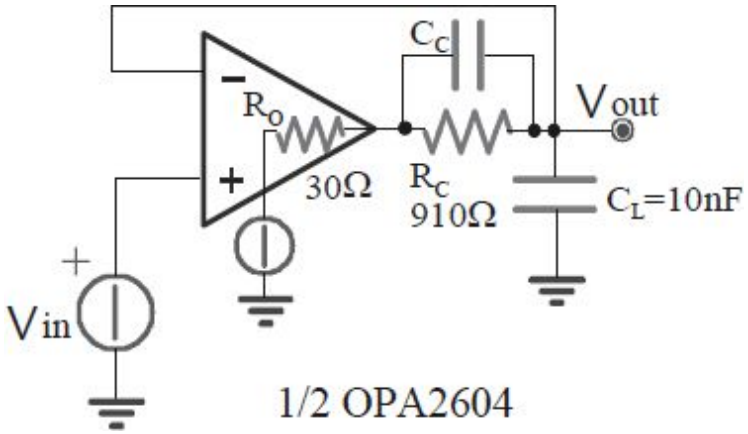


Fig. 3.67: Pole-zero compensation.

It is interesting to note that this circuit reduces the noise bandwidth **R,** although it is a non-inverting configuration. Indeed, for this reason, sometimes a capacitor C_L is purposely added at the output, obtaining the $R_C C_L$ filter that allows the abovementioned improvement.

Another method for compensating this circuit is what we will see now. This is a less empirical method, but it has the disadvantage of reducing the output voltage range, particularly if large output currents are required. It is pole-zero compensation, and the corresponding circuit is the one shown in Fig. 3.67.

The obtained Bode plot is shown in Fig. 3.68. To put discontinuities at a reciprocal distance of at least one decade, it is convenient to choose the following values:

$$R_C = 30R_0 \qquad C_C = \sqrt{\frac{0,018 \cdot C_L}{R_0 \cdot \beta \cdot GBWP}} << C_L$$

Note that, due to the large value of R_O, there will be a significant voltage drop.

Fig. 3.69 shows an example of a circuit to drive coaxial cables for video signals. Note the output impedance matching (source, cable, and load, all being 50Ω).

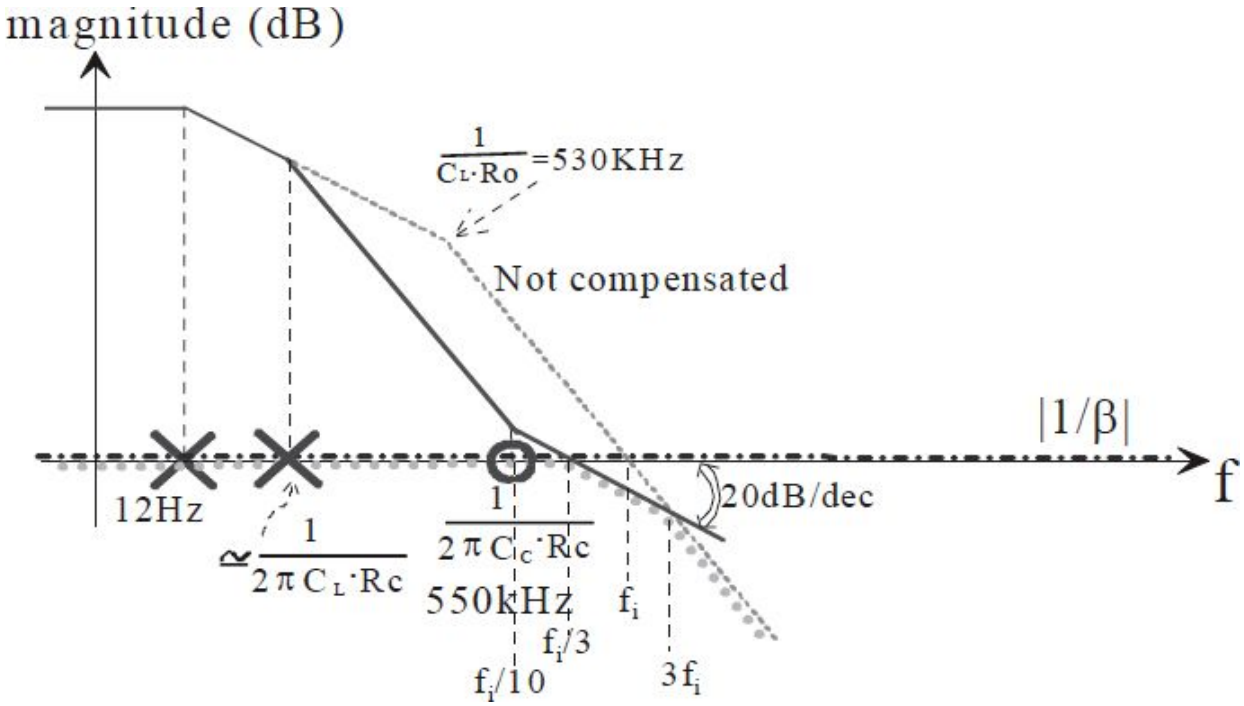


Fig. 3.68: Magnitude Bode plots.

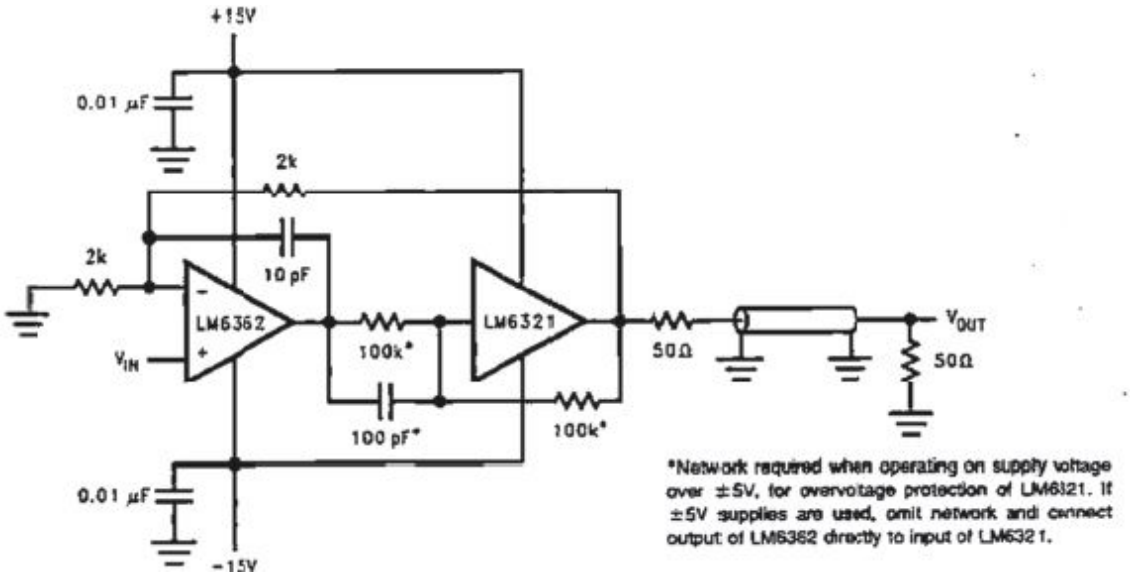


Fig. 3.69: Circuit to drive coaxial cables for video signals.

3.7. Negative Feedback Compensation

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 $\frac{1}{2\pi \cdot C_c \cdot \left(R_c + R_1 \| R_2\right)}$ 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: $\beta = \frac{R_1}{R_1 + R_2} \cdot \frac{1 + sR_c \cdot C_c}{1 + s\left(R_c + R_1 \| R_2\right) \cdot C_c}$

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.

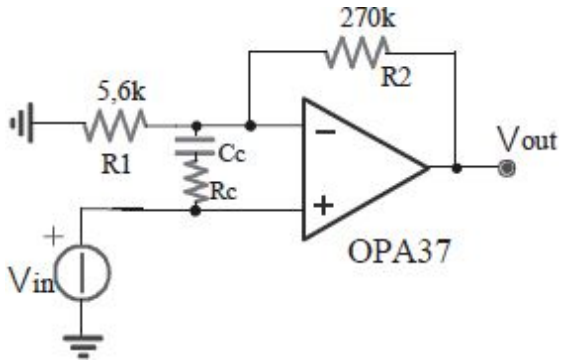


Fig. 3.70: Negative feedback compensation.

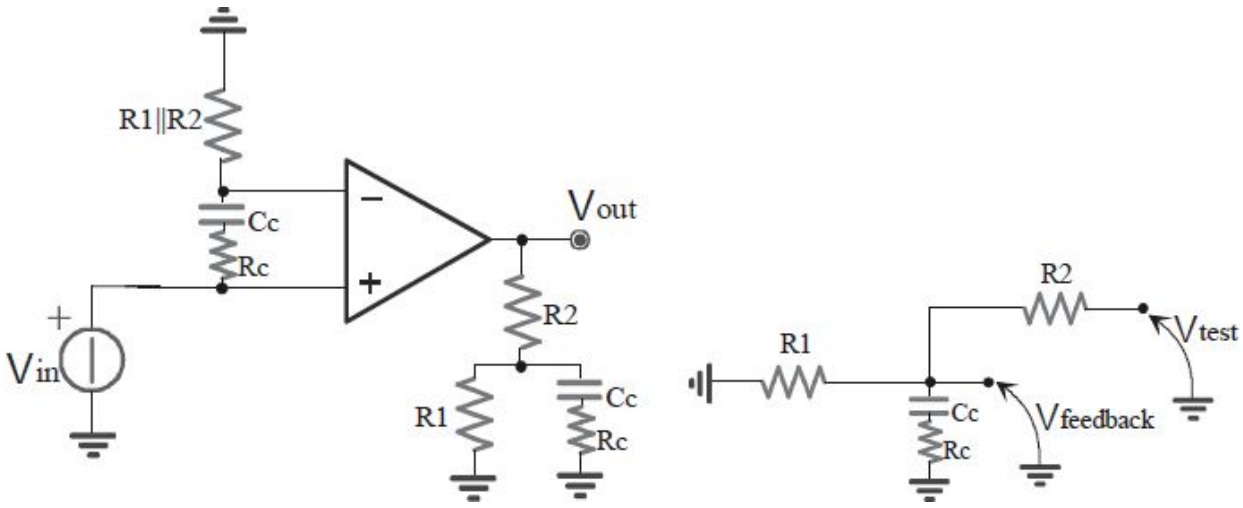


Fig. 3.71: Circuits for calculation of the forward (on the left) and the feedback blocks (on the right).

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.

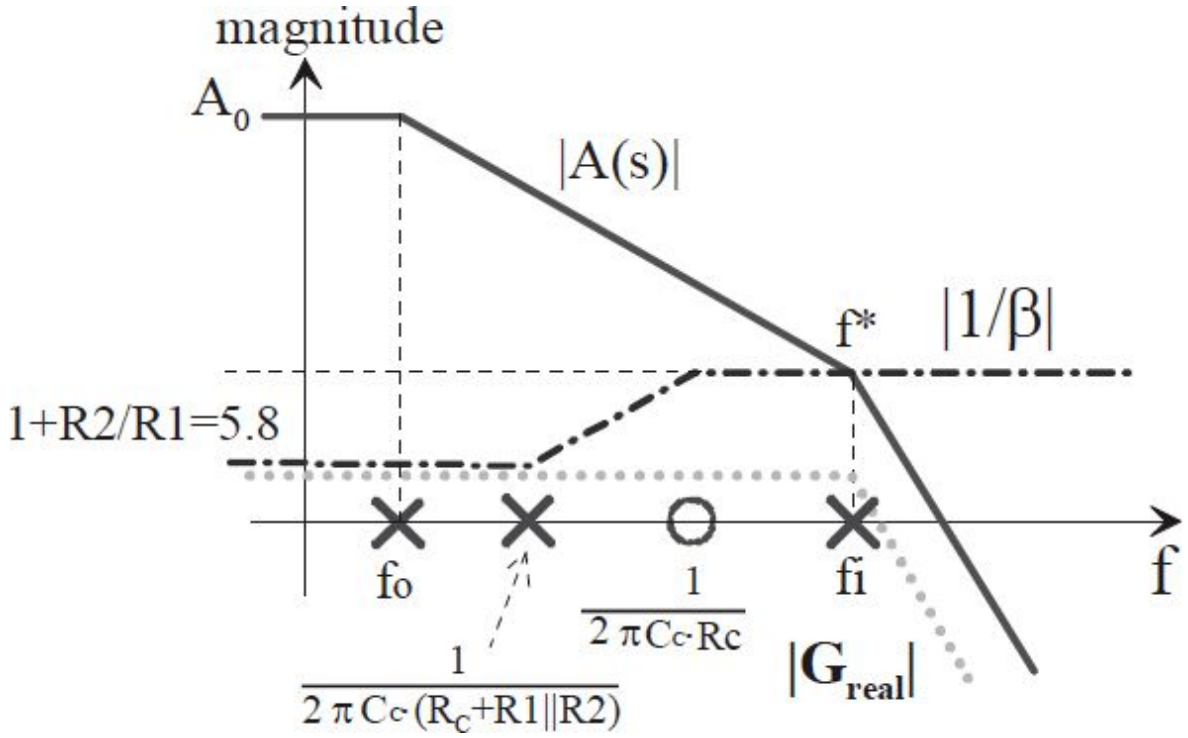


Fig. 3.72: Bode plots of the circuit in Fig. 3.70.

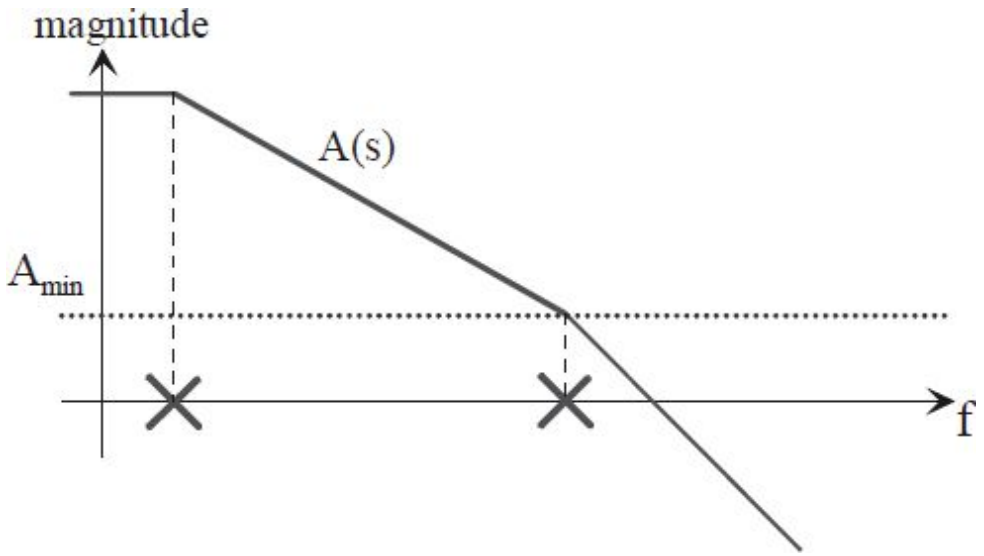


Fig. 3.73: A_{min} represents the minimum gain that can be used without needing any compensation.

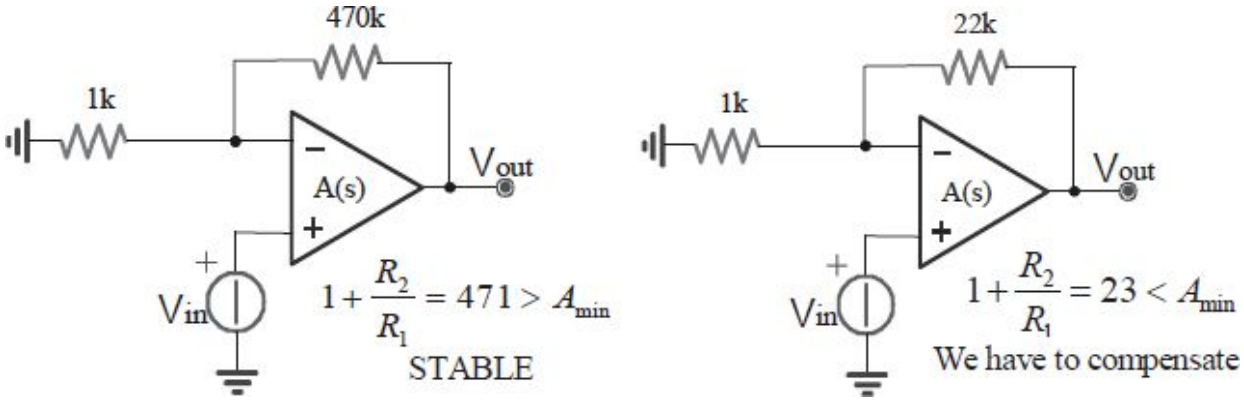


Fig. 3.74: Example of stable (on the left) and unstable (on the right) configurations.

3.7.1 EXAMPLE 1

Consider an operational amplifier with A_0=500.000 and A_{min}=100. In case of need of compensation, we choose: $\frac{1}{\beta}(highfrequency)=\frac{1}{\beta_0}+\frac{R_2}{R_C}=A_{min}$ and then we choose C_c so that the zero occurs at least a decade before $f_i=f_0\cdot A_0/A_{min}$.

3.7.2 EXAMPLE 2

Consider an operational amplifier with A_0=500.000, A_{min}=100 and f_0=1kHz. It is found:

$$R_C=\frac{R_2}{A_{min}-\frac{1}{\beta_0}}=\frac{270k}{100-5{,}8}=2866\Omega$$

the available resistance closest to this value is 2,7k.

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

and since $f_i=1k\cdot\frac{500.000}{100}=5MHz$, we obtain C_c=120pF.

Available OpAmps often offer the possibility of internal compensation (even for buffers) at the expense of the slew-rate. Others have the option of having a partial internal compensation which leaves A_{min}>1 to keep a higher slew-rate.

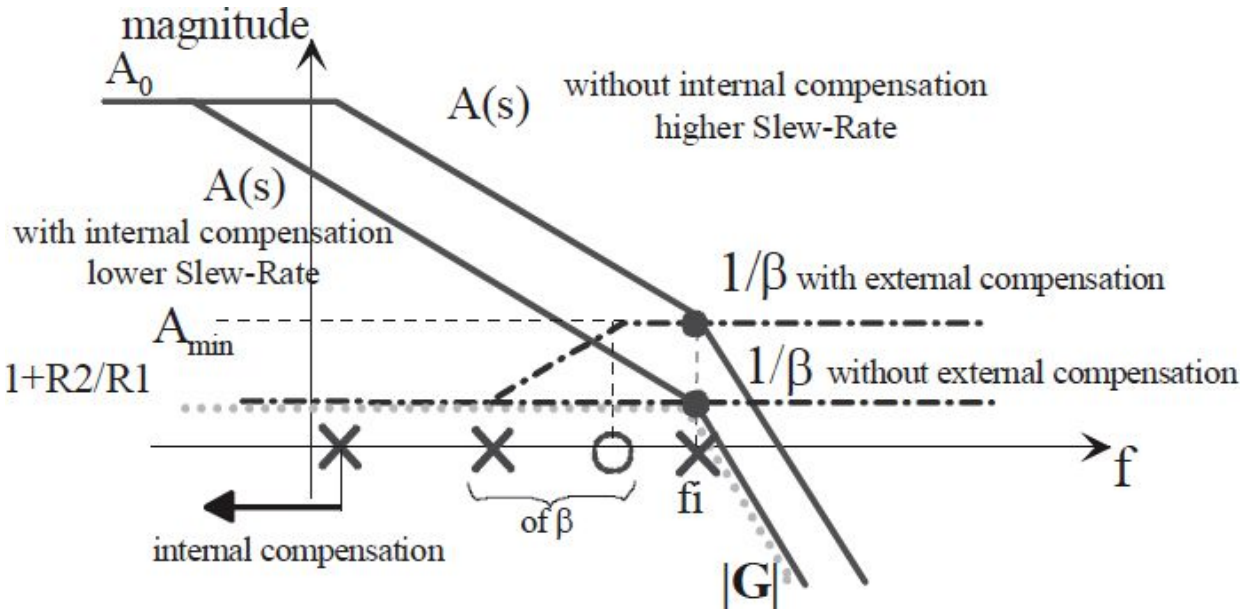


Fig. 3.75: Bode plot in the two cases described in the text.

In the latter case, we can use the higher slew-rate of the non-compensated OpAmp and then compensate it externally. Two products that match what has been said are the followings, for which we can note the difference in their slew-rates: OPA 37 SR=17V/μs OPA 27 SR=2V/μs

Let us now examine two possible options for compensation:

▪ A(s) with internal compensation, but 1/β without external compensation; ▪

A(s) without internal compensation, but 1/β with external compensation.

In both cases, we have stability (PM=45°) and the same closed-loop bandwidth, which is approximately:

$$f_i = f_0 \cdot \frac{A_0}{A_{min}} = 5MHz$$

Nonetheless, the second solution has a greater slew-rate, even though there is a corresponding increase in the settling-time because of the pole-zero pair.

To eliminate this "tail", simply remove C_C completely and leave only R_C, as shown in Fig. 3.76. In this case, 1/β coincides with the high frequency 1/β of Fig. 3.75, that is, with A_{min}.

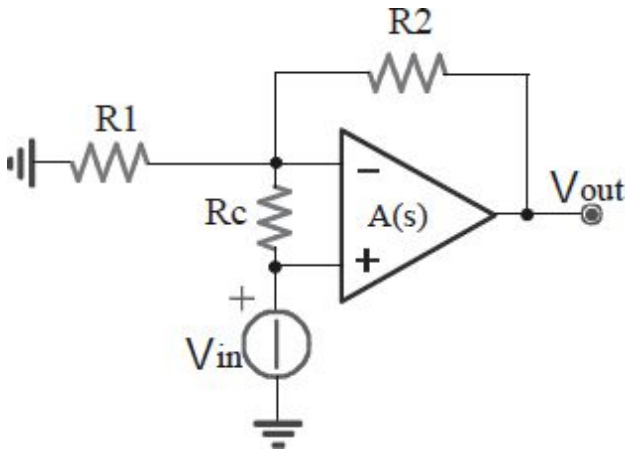


Fig. 3.76: Reduction of the tail in the settling-time.

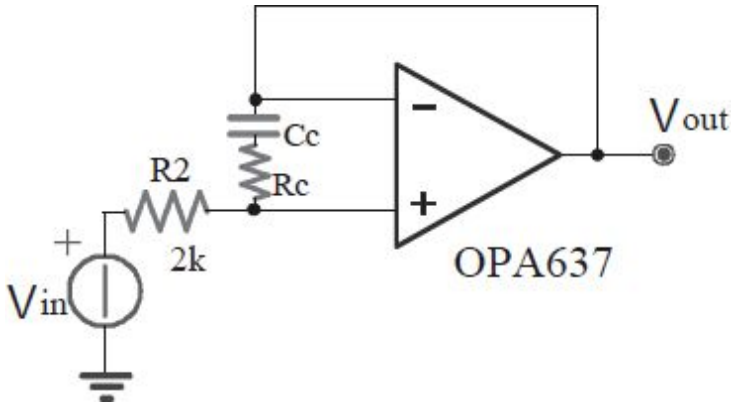


Fig. 3.77: Positive feedback compensation.

3.8. Positive Feedback Compensation

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:

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

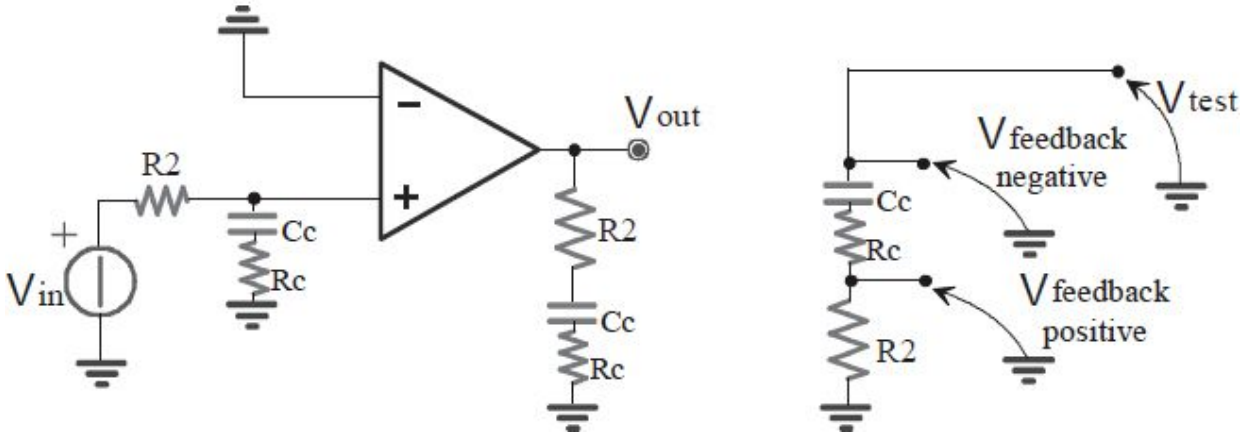


Fig. 3.78: Forward (on the left) and feedback (on the right) blocks

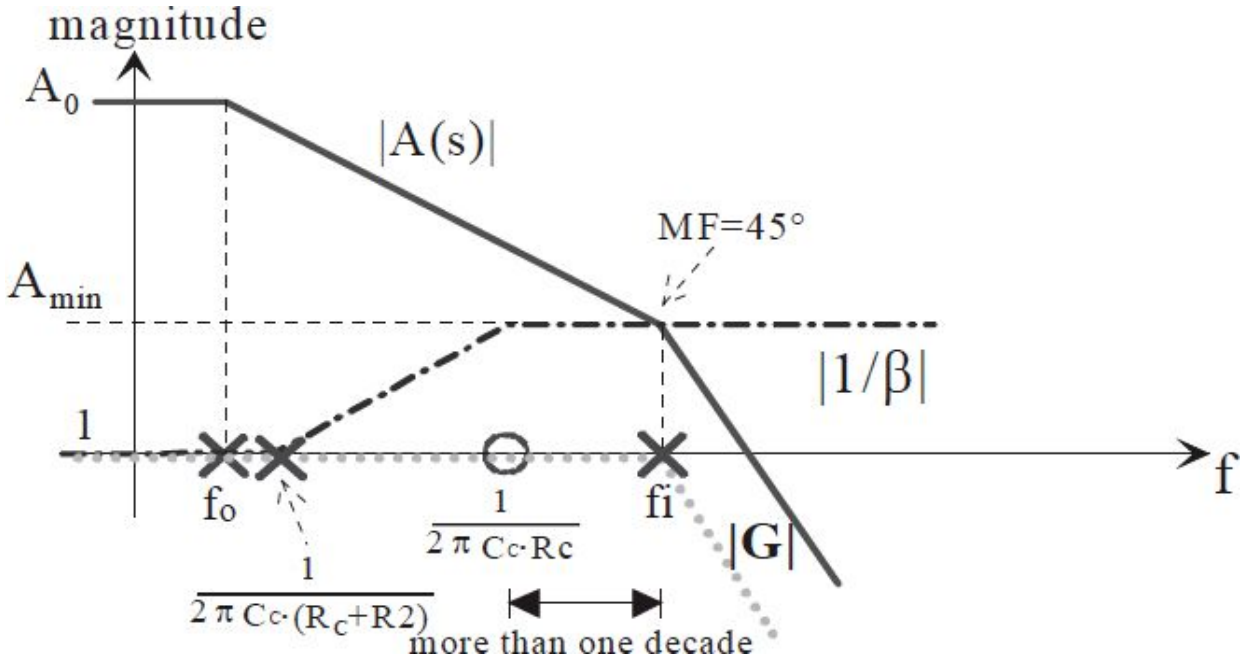


Fig. 3.79: Magnitude Bode plots.

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:

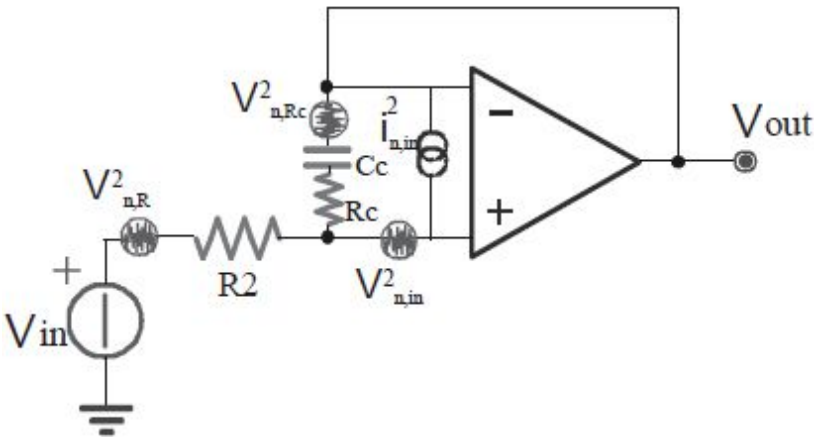
$$Z_{in} = \left(R_C + \frac{1}{sC_C} \right) \cdot G_{loop}(s)$$ where $G_{loop}(s)=A(s)\cdot\beta$ and $\beta<1$.

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.

Now R_2 is chosen according to the desired Z_{in}, such as by noise considerations, and to the equivalent scheme depicted in Fig. 3.80. Increasing R_2 increases also Z_{in}, but also its thermal noise $\sqrt{4kTR_2}$ at the amplifier input. A trade-off is needed. To make it weigh less than 1/10 of the

overall rms value, it will be sized to give a voltage noise equal to 1/√10,

$$R_2 = \frac{v^2_{n,in_{op-amp}}}{10 \cdot 4kT}$$



then:

Fig. 3.80: Equivalent noise generators.

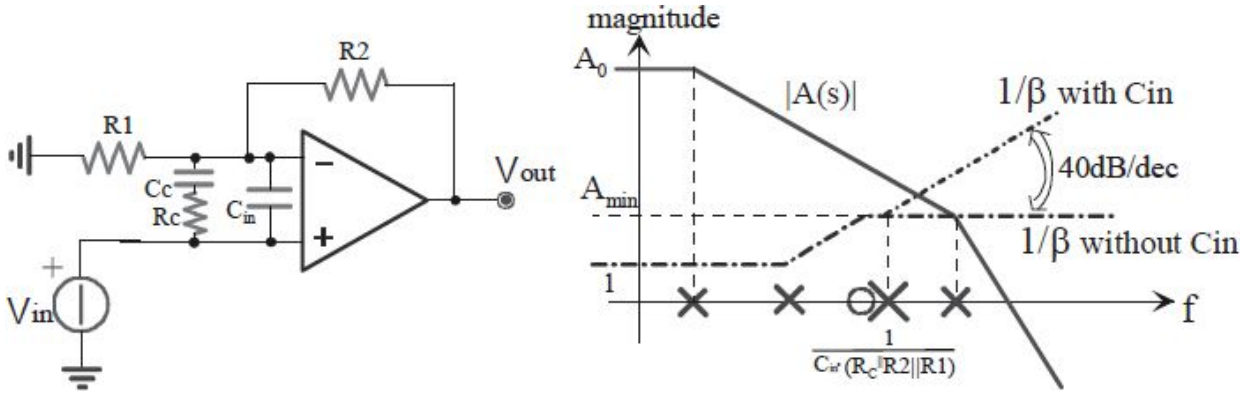


Fig. 3.81: Influence of capacitance C_{in} on negative feedback compensation.

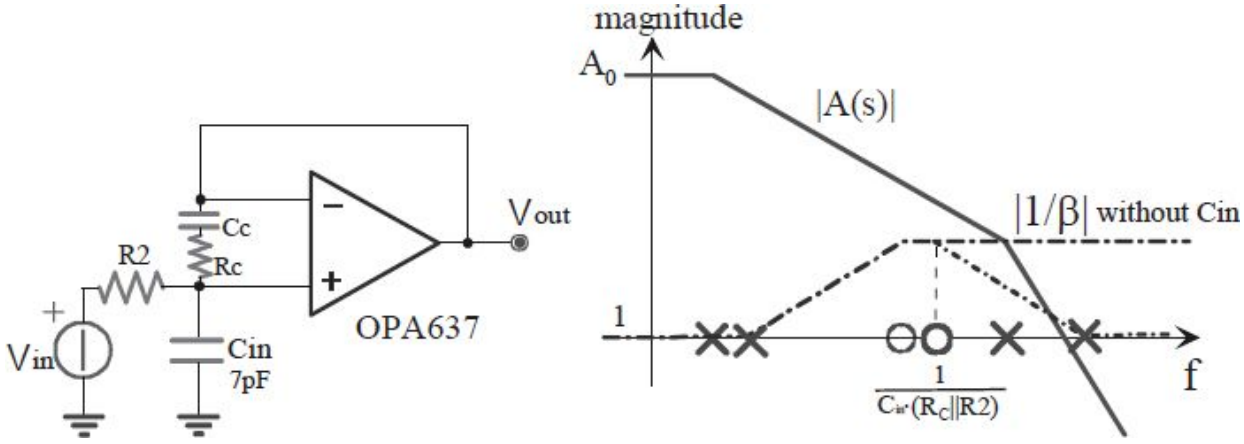


Fig. 3.82: Influence of capacitance C_{in} on positive feedback compensation.

Let us consider, for instance, the OPA637, which has 5nV/√Hz. This would entail the use of a resistor R_2=500Ω that would lead to the choice of R_C=125Ω and C_C=820pF. If we added a capacitor C_{in} in the case of negative feedback, this would worsen the stability since it would determine a rise of 1/β at high frequencies, as depicted in Fig. 3.81. On the contrary, in the case of positive feedback compensation, the presence of C_{in} improves stability, as Fig. 3.82 demonstrates.

3.9. Feedback Advance Compensation

This compensation allows to use non-compensated OpAmp to obtain gains lower than its A_{min}, by reducing the closure angle between A(s) and 1/β, as shown in Fig. 3.83.

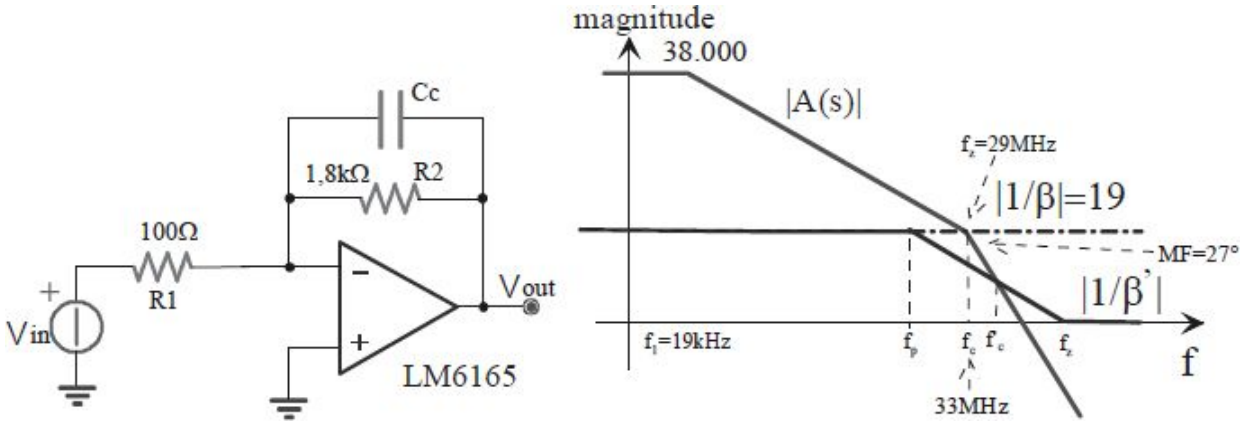


Fig. 3.83: Compensation of a non-compensated OpAmp.

4

Advanced OpAmps

"Did I disappoint you or let you down?
Should I be feeling guilty or let the judges frown?
'Cause I saw the end before we'd begun,
Yes I saw you were blinded and I knew I had won.
So I took what's mine by eternal right.
Took your soul out into the night.
It may be over but it won't stop there,
I am here for you if you'd only care.
You touched my heart you touched my soul.
You changed my life and all my goals.
And love is blind and that I knew when,
My heart was blinded by you.
I've kissed your lips and held your hand.
Shared your dreams and shared your bed.
I know you well, I know your smell.
I've been addicted to you."

"Goodbye my Lover", James Blunt

This chapter presents a number of special operational amplifiers, whose design differs from that of typical voltage-mode architectures for optimizing some features. For example, the instrumentation amplifier is characterized by high common mode noise rejection and high inputs impedances; the isolation amplifier is characterized by its ability to ensure galvanic isolation between the input signal and the output signal; the current feedback amplifier and the Norton amplifier are characterized by high bandwidth and high slew rate.

4.1. Instrumentation Amplifier (INA)

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.

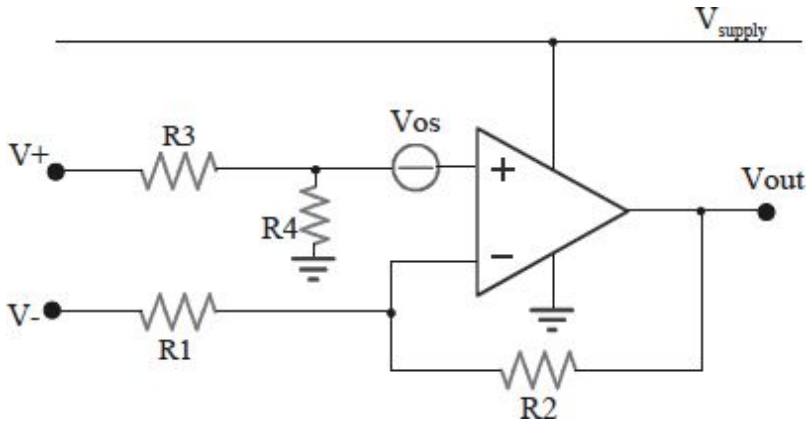


Fig. 4.1: Differential amplifier.

4.1.1 Differential amplifier

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: • 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.

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. 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*.

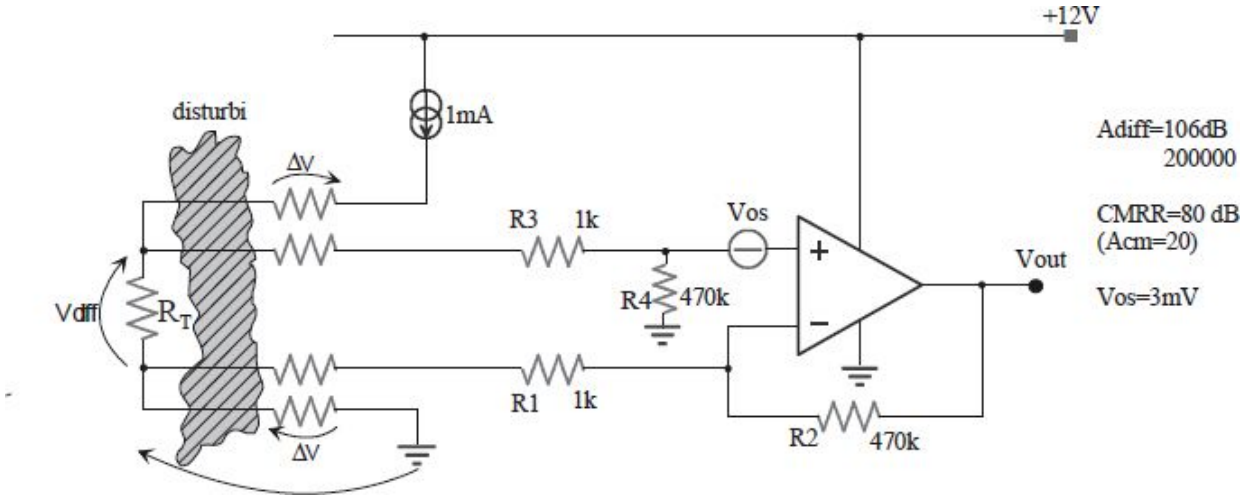


Fig. 4.2: An example of the use of a Differential amplifier.

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.

4.1.2 Instrumentation amplifier

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.

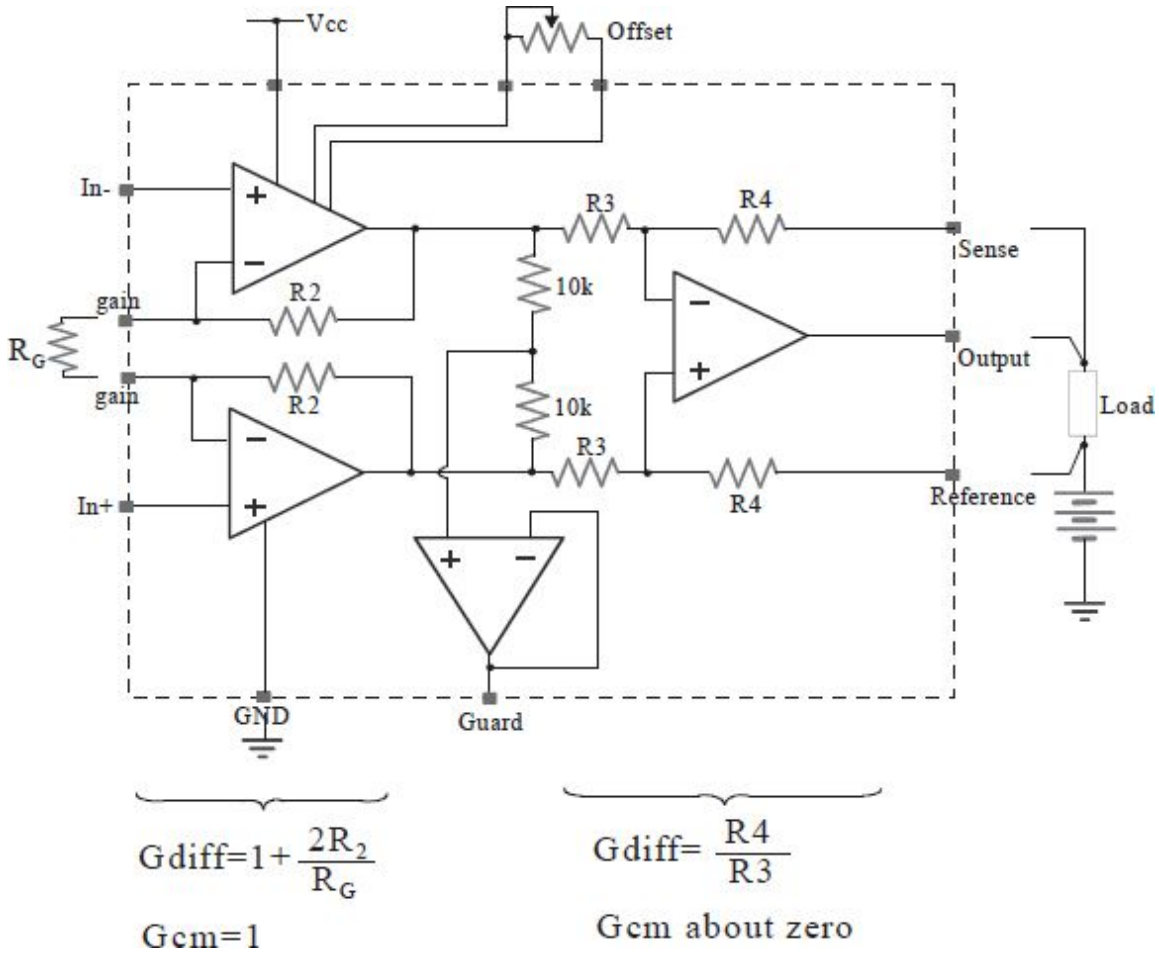


Fig. 4.3: Instrumentation Amplifier.

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.

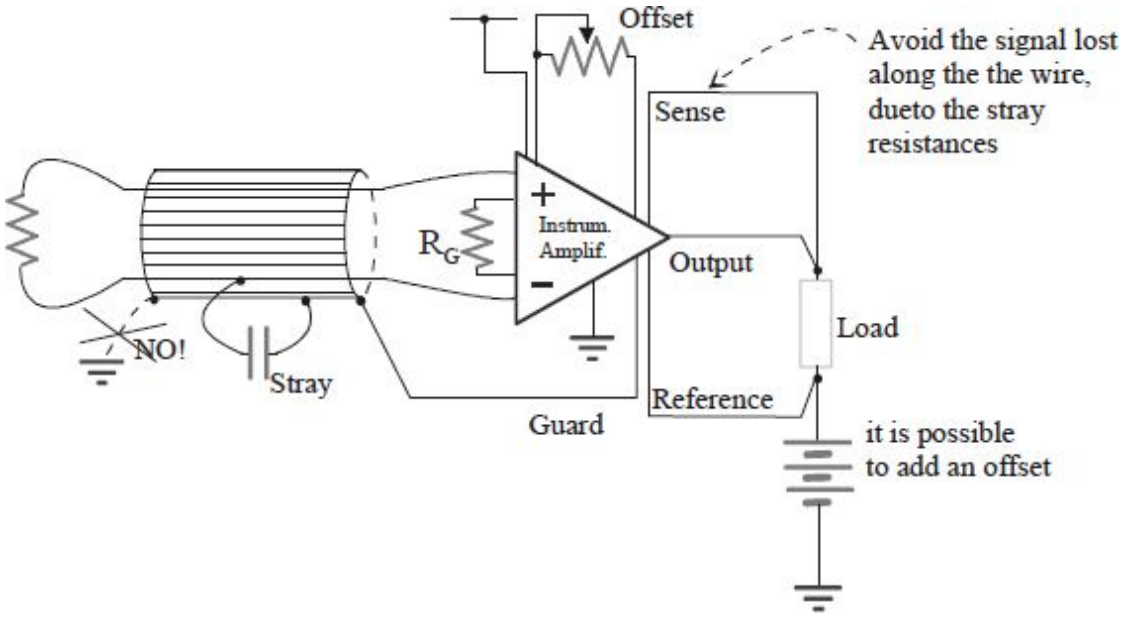


Fig. 4.4: An example of correct usage of the INA.

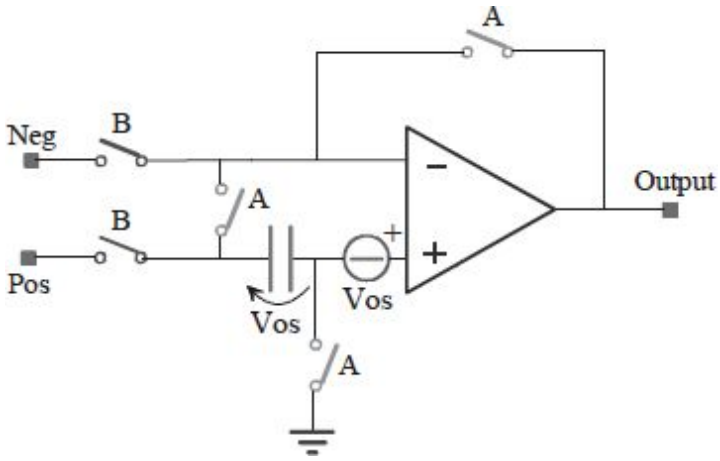


Fig. 4.5: Commutating Auto Zeroing (CAZ) Amplifier: phase A (Sampling); phase B (Zeroing).

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.

4.1.3 Auto zeroing

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.

4.1.4 INA: usage examples

The INAs are often used to acquire (with high CMRR) very weak signals from very high impedance sources. To process the signal well, it is fundamental to avoid the errors in the design of the electronics. Here are presented some examples of the usage of INAs.

The INA is often used for monitoring signals from patients (Fig. 4.6). The typical impedance of the electrodes is 100kΩ. The resistance R_B is used for guaranteeing the return of the bias current of the INA without which the INA could not work well.

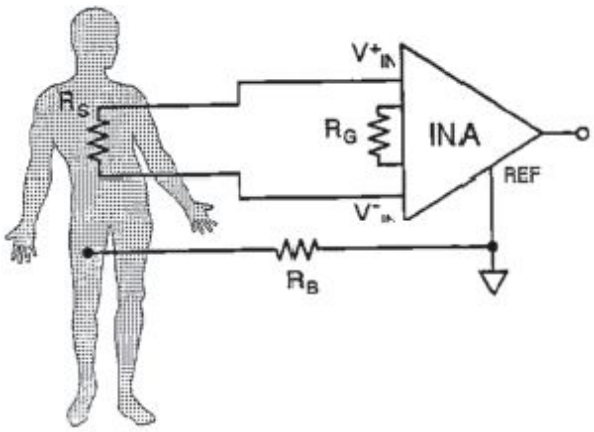


Fig. 4.6: INA for electro-medical applications.

The INA114 has an input offset of only 50μV and bias current of 2nA. Thus, with only 100kΩ it is possible to induce an additional offset error of 200μV. If the patient moves, he can cause an undesired differential signal that can be seen as low frequency noise. The input noise of the INA103 is equal to 1nV/√Hz (equivalent to that a 61Ω resistor), the lowest of the Burr Brown INAs. Unfortunately, because of the high source impedance, the current noise is very high, equal to 2pA/√Hz, and then the voltage noise is 100kΩ·2pA=200nV/√Hz. Instead, the INA111 with an FET input has a bias current of 20pA (and then there is an offset of 2μV) and, with a noise of 0.8fA/√Hz, contributes only for 80pV/√Hz to the input noise.

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 $1/(2\pi RC)$. 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.

We have the same problems with high-pass filtering. Let us study the solution illustrated in Fig. 4.9, in which the INA Reference terminal is used for making a feedback connection with a low-pass version of the actual output.

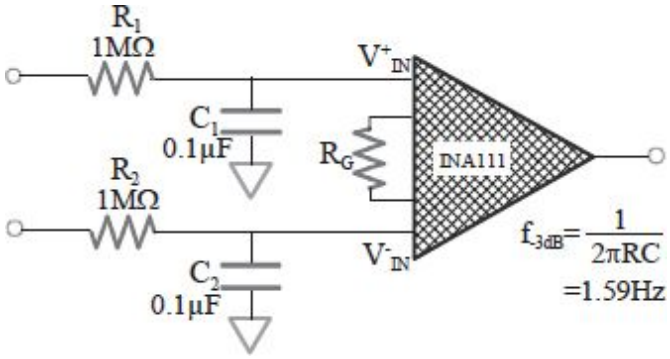


Fig. 4.7: Wrong low-pass filtering.

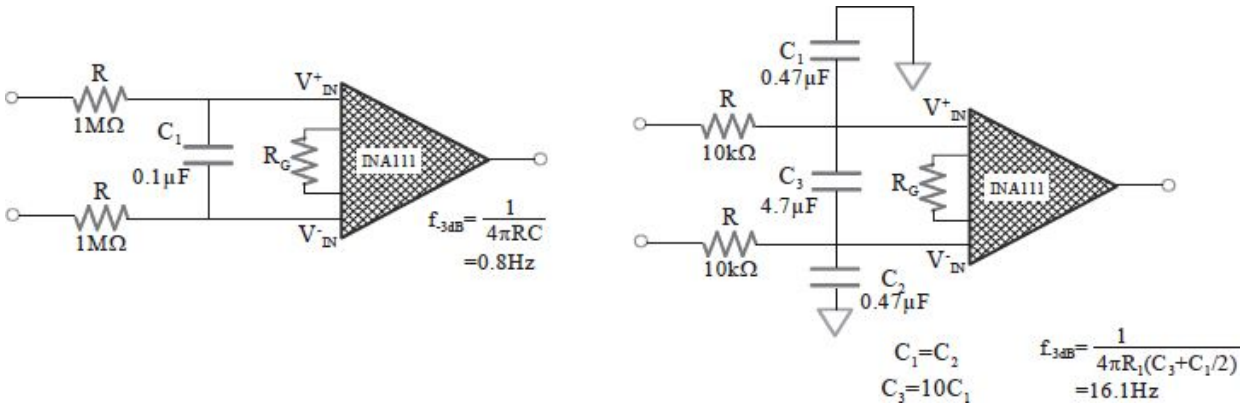


Fig. 4.8: Differential low-pass (on the left) and common mode (on the right) filtering.

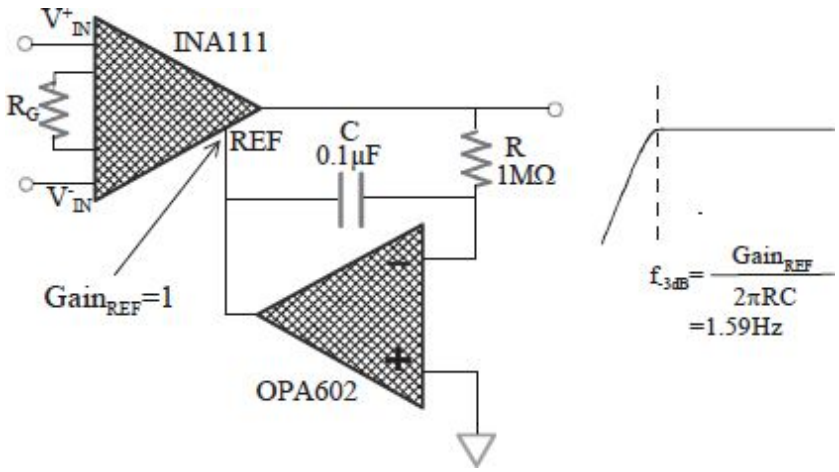


Fig. 4.9: High-pass filtering.

The gain *GainRef* is between the pin REF and the output of the INA. Nevertheless, if the DC input of the INA multiplied by the gain and then added to the DC common mode is high enough to saturate the INA, it is necessary to use the classical configuration shown in Fig. 4.10. Naturally, the mismatch problem of the components reappears. Another possibility is to act on the external gain resistance (Fig. 4.11, in which the two internal resistance values are supposed to be 20kΩ).

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.

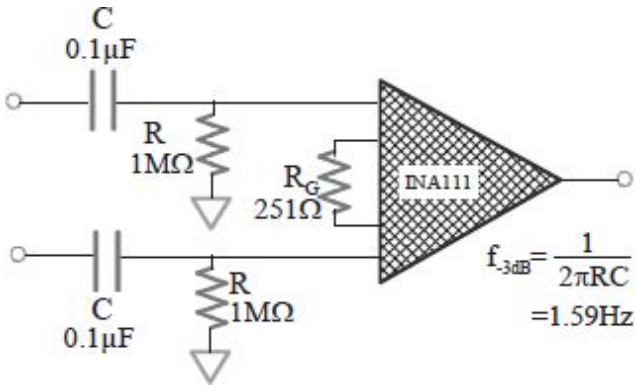


Fig. 4.10: Classical solution for high-pass filtering.

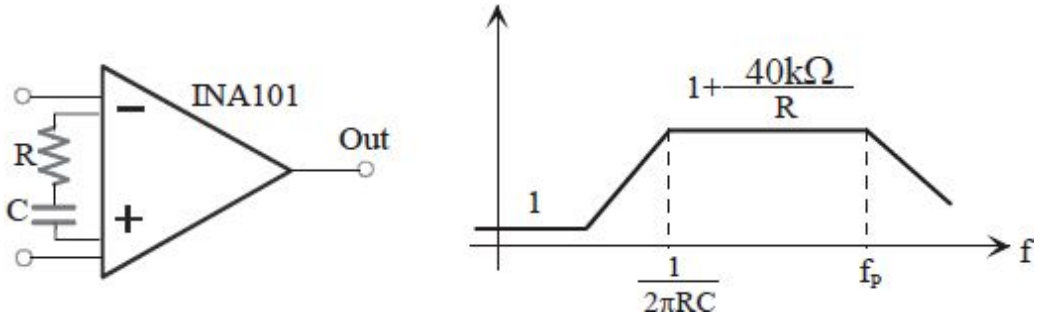


Fig. 4.11: Filtering obtained acting on the external gain resistance.

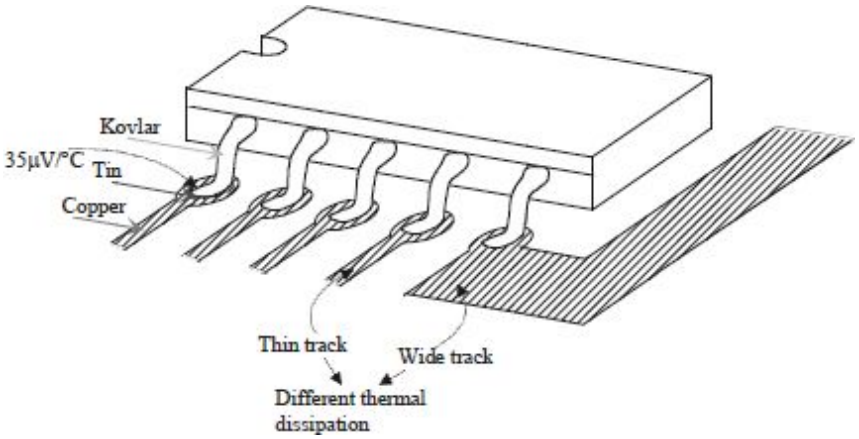
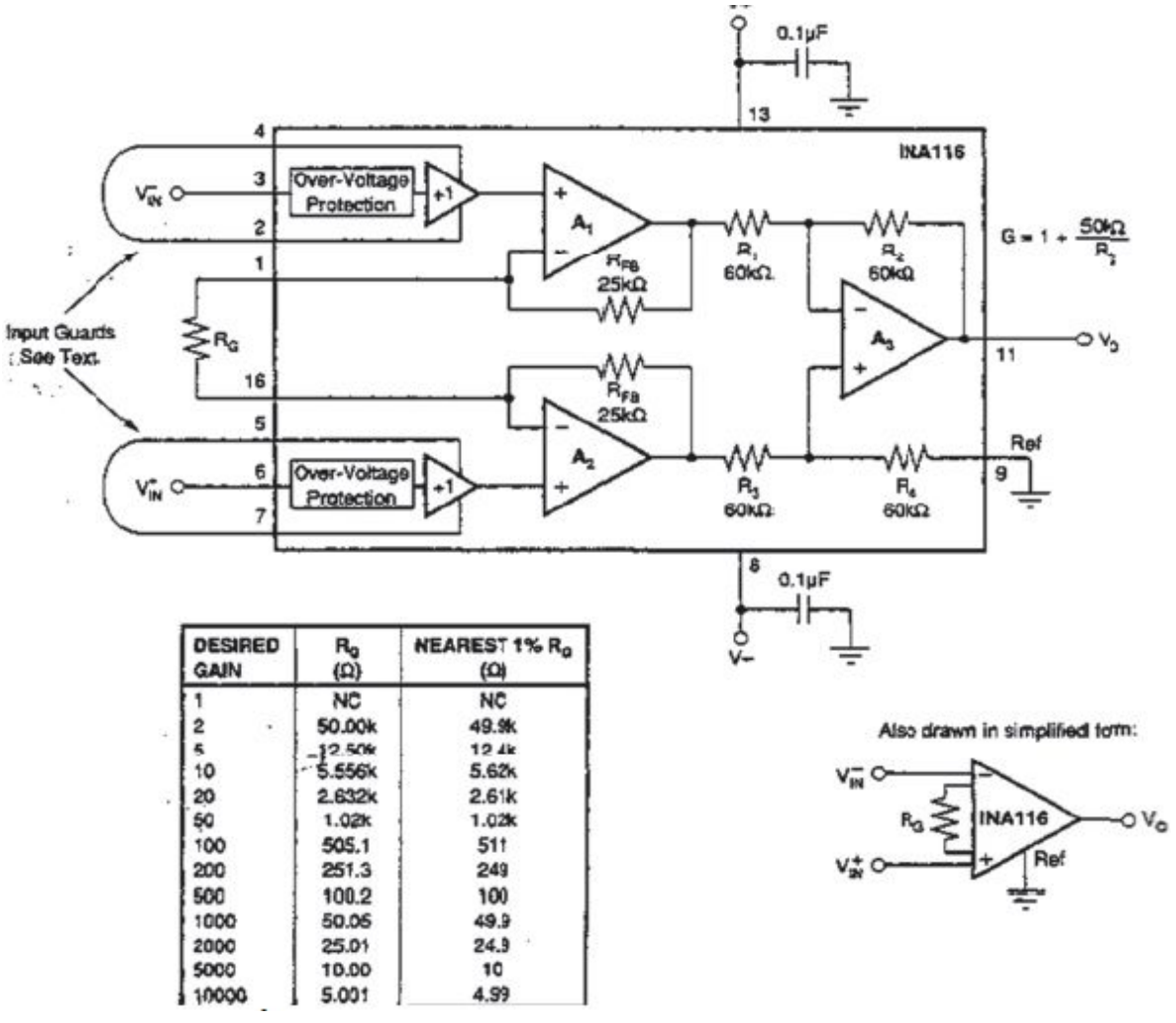


Fig. 4.12: Thermocouple problem due to the IC welds.

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.

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.



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

Fig. 4.13: Fast signal Op-Amp.

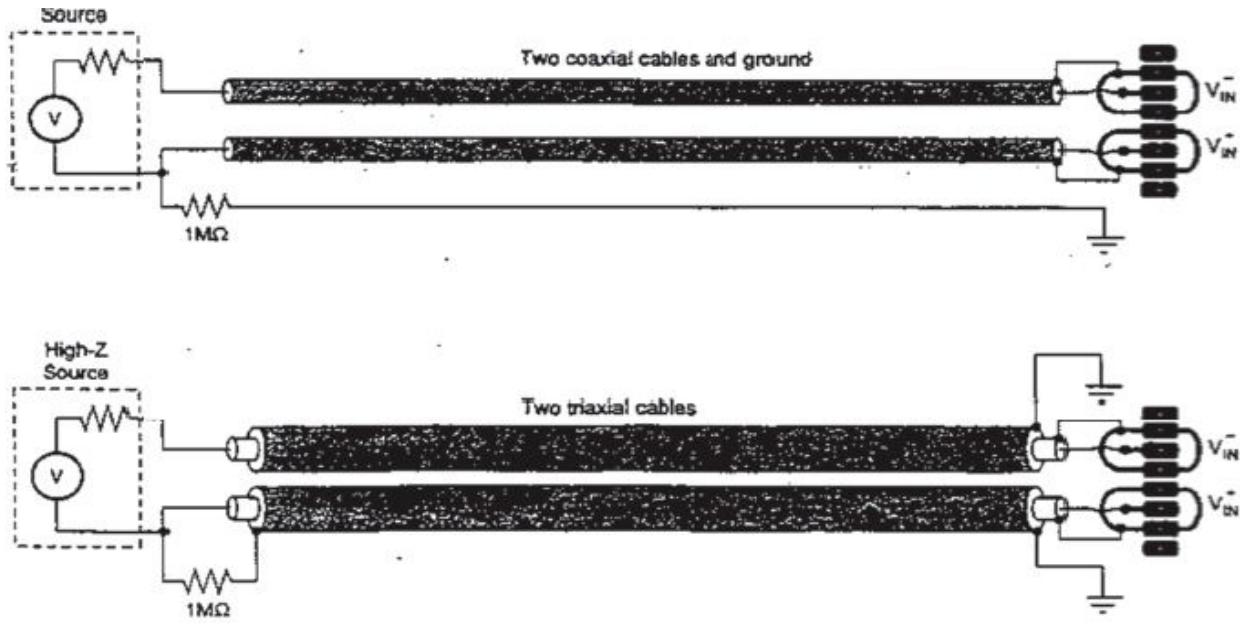


Fig. 4.14: Connection that ensures that the bias current goes forward and then returns.

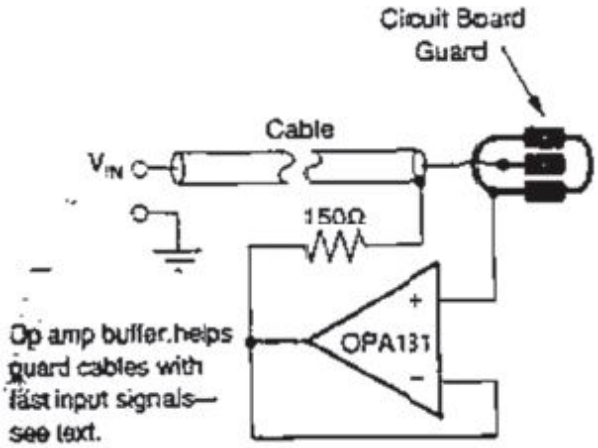


Fig. 4.15: Fast connection buffer.

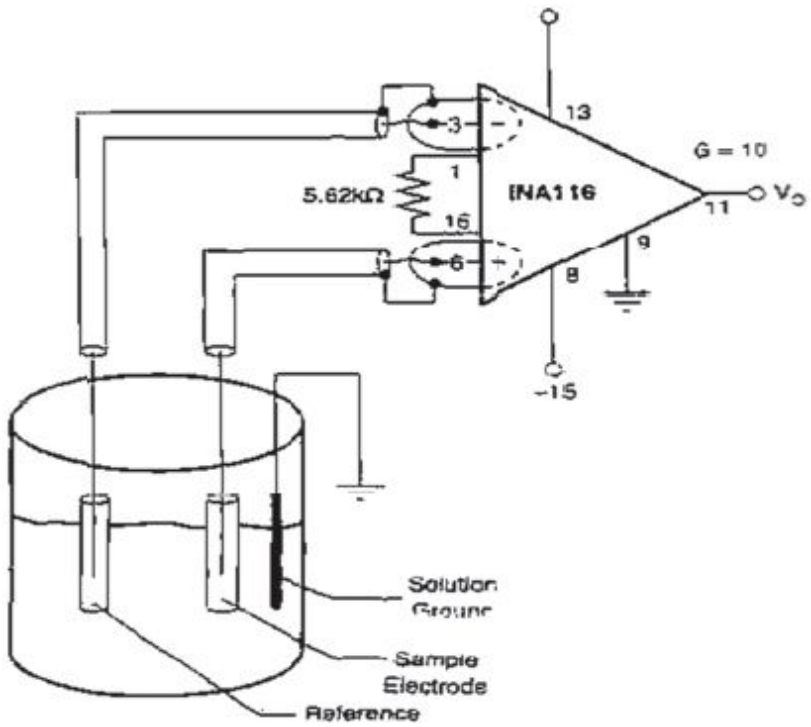


Fig. 4.16: Example of connections to ensure bias current loop.

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.

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.

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.

4.2. Isolation Amplifiers (ISO)

4.2.1 Introduction

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.

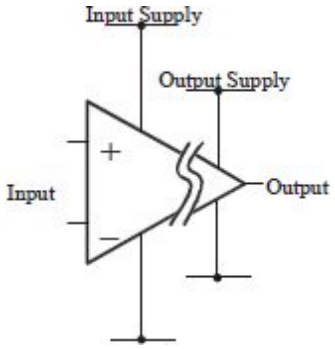


Fig. 4.17: Isolation Amplifier scheme.

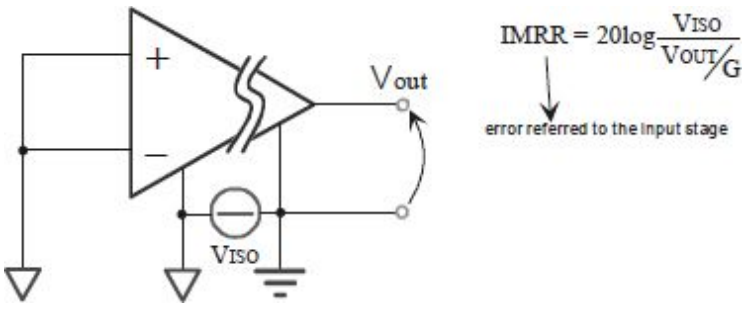


Fig. 4.18: Definition of IMRR.

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.

4.2.2 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.

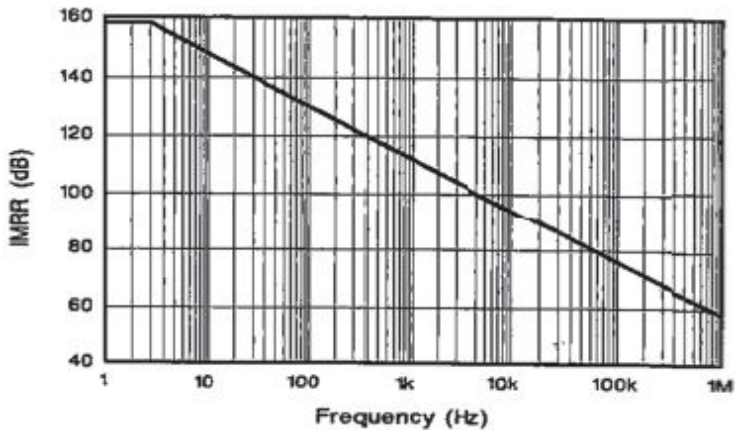


Fig. 4.19: IMRR vs. frequency.

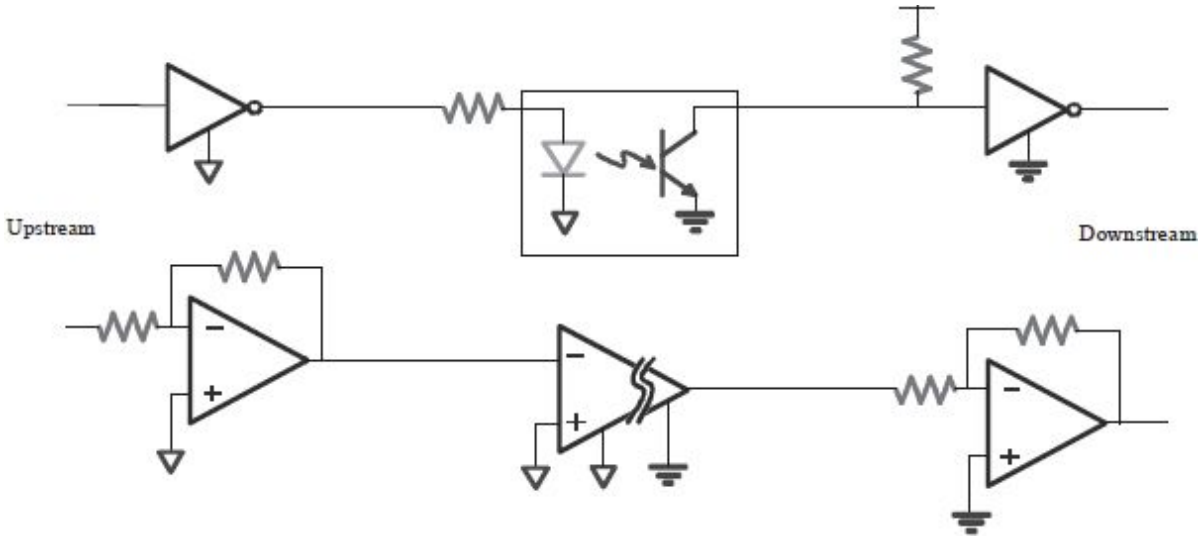


Fig. 4.20: Optical coupling scheme for digital (top) and analog (bottom) information transfer.

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.

The disadvantage of the optocoupler is the detector mismatch which causes the poor linearity between the ISO input and the output. Another problem is the amplifier's internal noise, the resistance noise, the output stage noise, and the noise of the couple photodiode/LED. This noise can be reduced by filtering the ISO output.

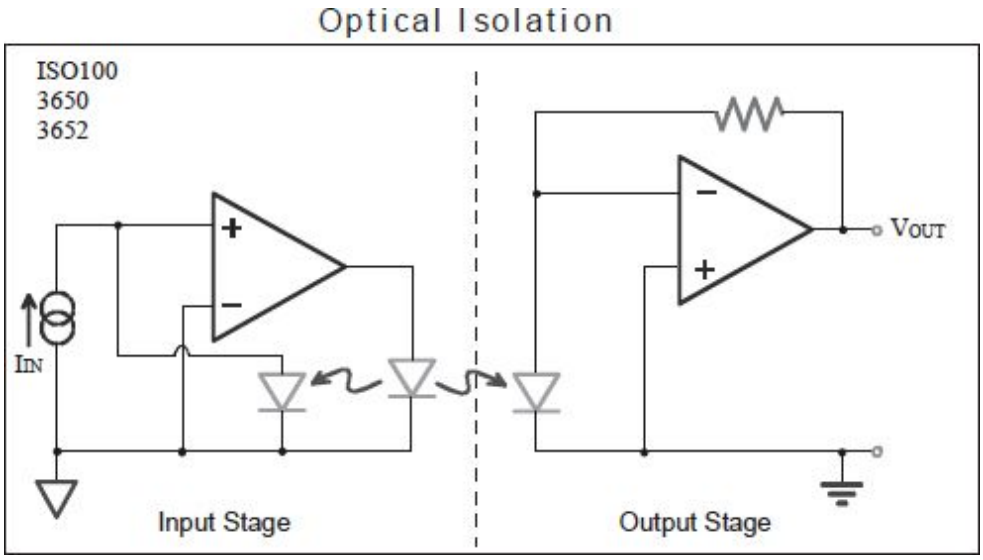


Fig. 4.21: Analog optocoupler based on a negative feedback through a double optocoupler.

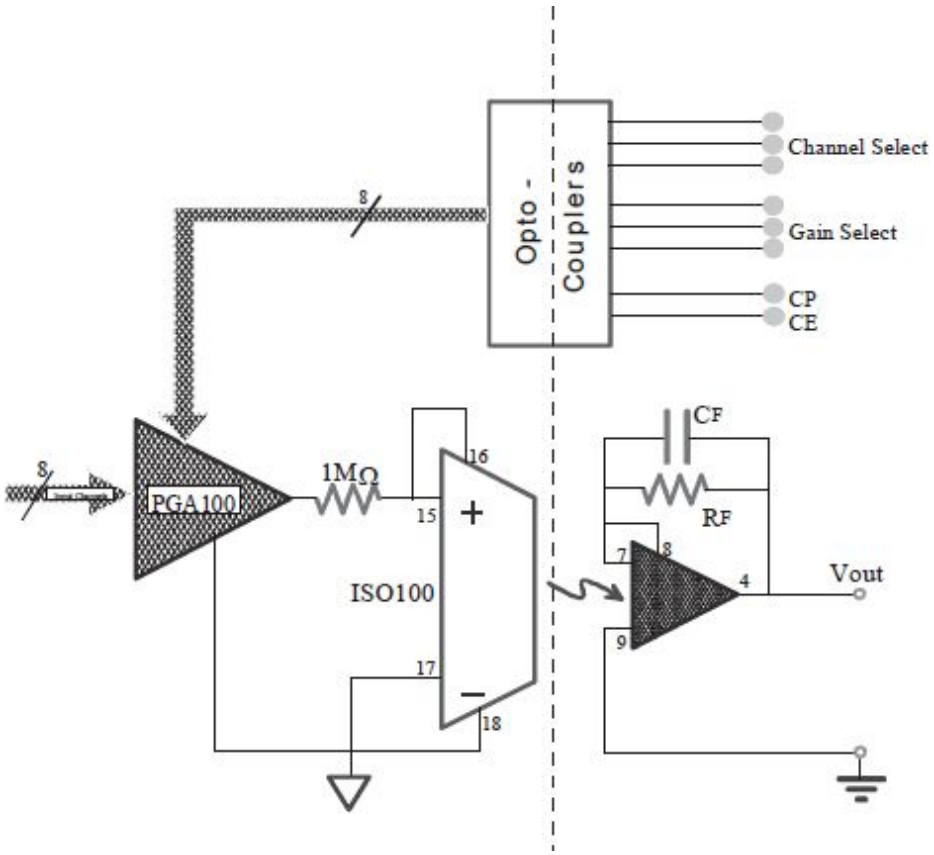


Fig. 4.22: Signal acquisition and precondition system.

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.

4.2.3 Transformer Isolation

To transfer analog information in this way, a chopper within the ISO is required to perform a signal amplitude modulation (DC eventually). On the secondary winding of the transformer, the signal must be rectified or demodulated with the correct phase. In this way, it is possible to obtain an isolated galvanic connection that is quite tolerant to RF noise. The saturation is rarely a problem, and the signal path quickly recovers any overload.

Some ISOs, as the 3656 and the ISO212, have an internal DC/DC converter to internally generate the supply voltages (isolated) for the two worlds (Fig. 4.23). Connecting the output OpAmp, it is possible to realize various circuit configurations.

The transformer is realized within the IC that is not monolithic, but has a hybrid structure (however, always contained in a single package). The main disadvantage of using a transformer is the reduced signal bandwidth because the windings are made with bondings on a toroidal nucleus as in Fig. 4.24.

The problem for this kind of isolation is, in addition to the amplifier internal noise and resistance noise; there exists the output residual ripple. This is a product of the ISO modulation/demodulation that has the fundamental tone at the modulation frequency f_{mod} over the signal bandwidth. Increasing the frequency, we have a better linearity and a reduction in the gain error, as shown in Fig. 4.25. Particularly, with a linearity error of 50m%=0.05%·FSR, the producer states that it is suitable for 11 bits of accuracy (in fact ½11=0.048%FSR).

Fig. 4.26 shows the effect of the residual output ripple at various input frequencies and amplitudes.

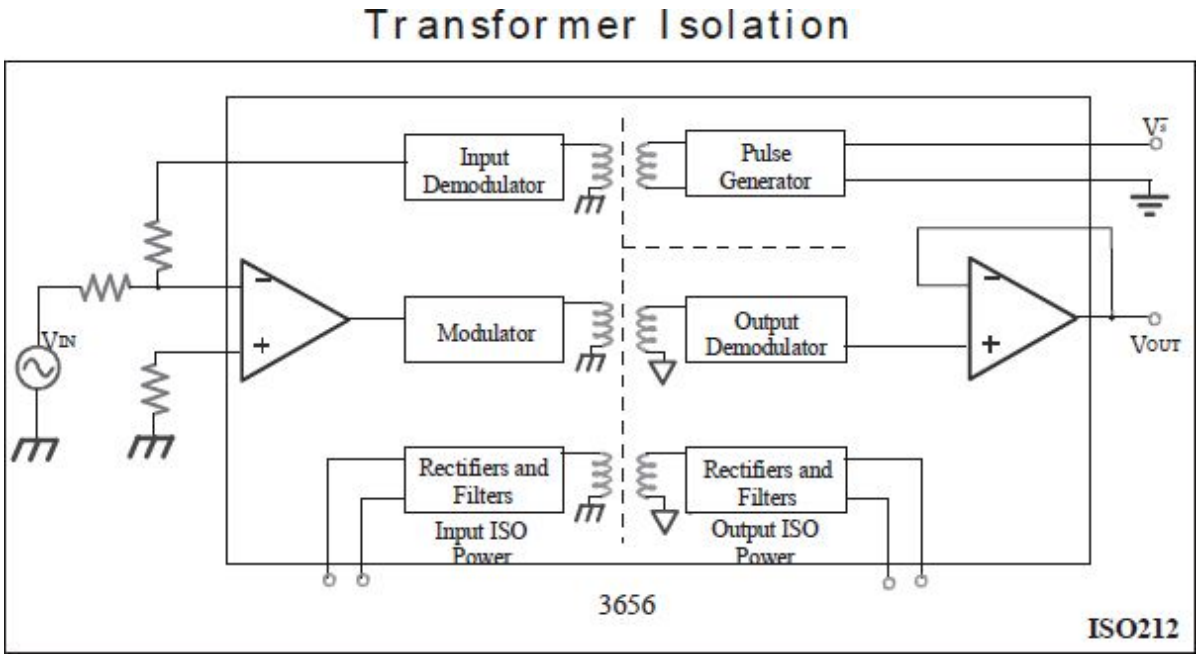
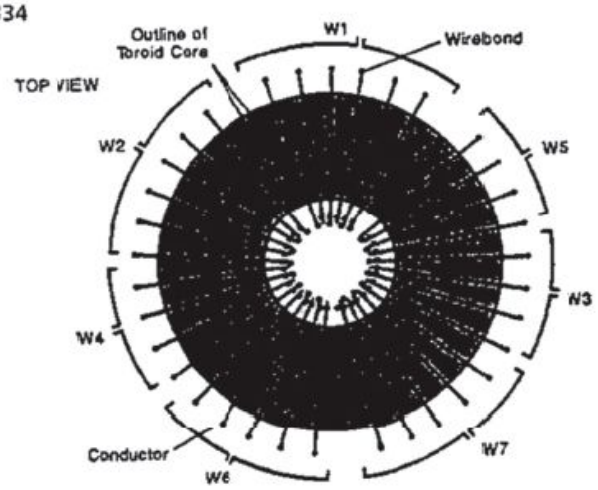
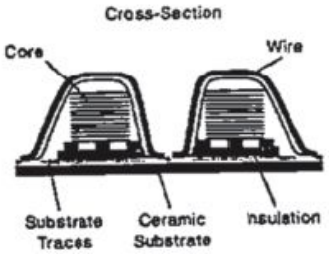


Fig. 4.23: ISO3656 and ISO212.





BURR-BROWN®

ISO212KP

ISO212P

Low Cost, Two-Port Isolated, 1500Vrms ISOLATION AMPLIFIER

FEATURES

- 12-BIT ACCURACY
- 2.5mA (typ) QUIESCENT CURRENT
- LOW PROFILE (LESS THAN 0.5" HIGH)
- SMALL FOOTPRINT
- EXTERNAL POWER CAPABILITY (±5V at 5mA)
- "MASTER/SLAVES" SYNCHRONIZATION CAPABILITY
- INPUT OFFSET ADJUSTMENT
- LOW POWER (75mW)
- SINGLE 10V TO 15V SUPPLY OPERATION

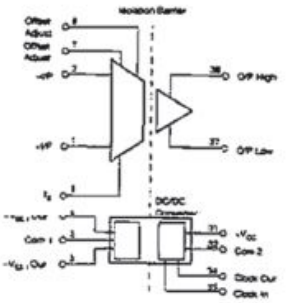
APPLICATIONS

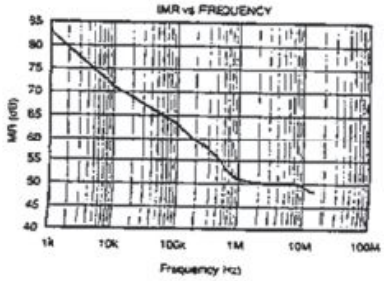
- INDUSTRIAL PROCESS CONTROL: Transducer Channel Isolator for Thermocouples, RTDs, Pressure Bridges, Flow Meters
- 4mA TO 20mA LOOP ISOLATION
- MOTOR AND SCR CONTROL
- GROUND LOOP ELIMINATION
- ANALYTICAL MEASUREMENTS
- POWER PLANT MONITORING
- DATA ACQUISITION/TEST EQUIPMENT ISOLATION
- MULTIPLEXED SYSTEMS WITH CHANNEL TO CHANNEL ISOLATION

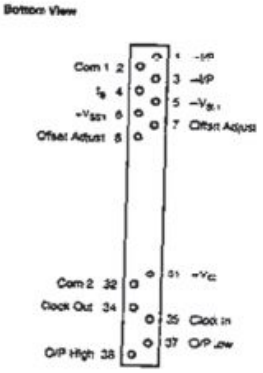
DESCRIPTION

The ISO212P signal isolation amplifier is a member of a series of low-cost isolation products from Burr-Brown. The low-profile SIL plastic package allows PCB spacings of 0.5" to be achieved, and the small footprint results in efficient use of board space.

To provide isolation, the design uses high-efficiency, miniature toroidal transformers in both the signal and power paths. An uncommitted input amplifier and an isolated external bipolar supply ensure the majority of input interfacing or conditioning needs can be met. The ISO212P accepts V_{IN} input voltage range of ±5V for single 15V supply operation or ±3.0V for single 10V supply operation.







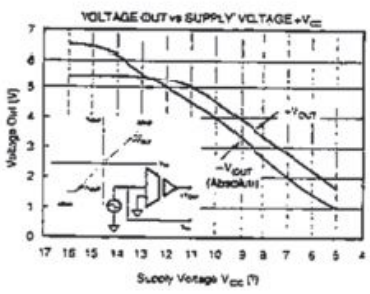


Fig. 4.24: The windings limit the bandwidth.

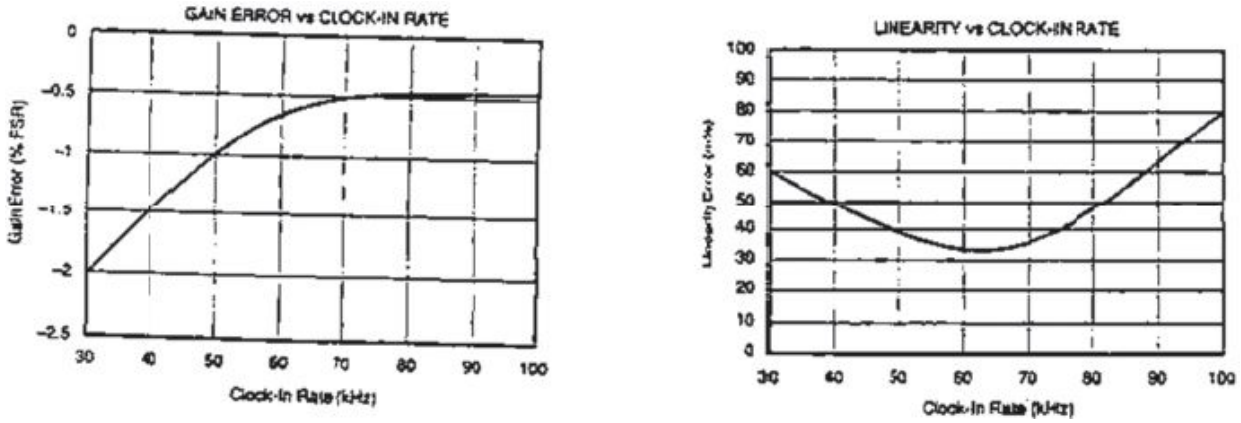


Fig. 4.25: Trends of gain error and the linearity vs. frequency.

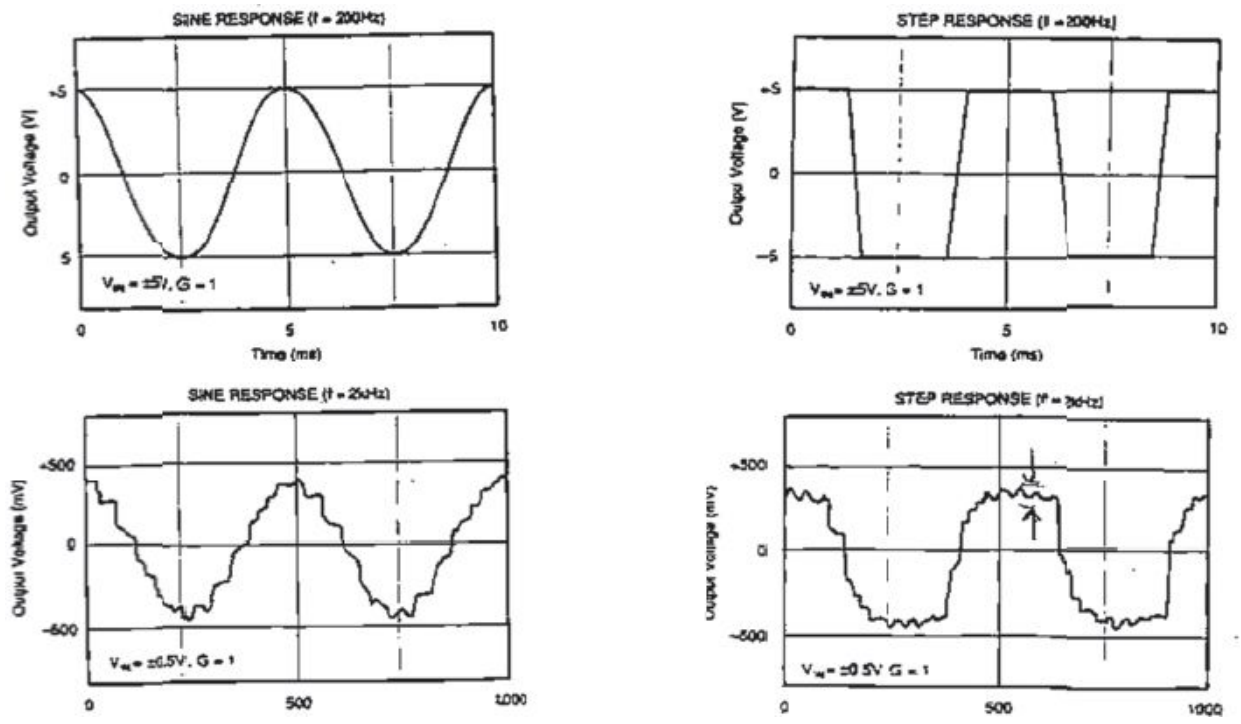


Fig. 4.26: Residual ripple effect.

4.2.4 Capacitive coupling

It is used as a modulation technique with variable duty-cycle or frequency modulation to transmit a signal through the “barrier”. The output stage

receives and demodulates the signal to analog domain, rejecting the common-mode signal.

These ISOs can have (Fig. 4.27) or not (Fig. 4.28) a DC/DC converter for the internal and external supply voltages. For multichannel applications, the ISO with capacitive coupling and the DC/DC converter can be synchronized to reduce the noise along the path and the number of various oscillation frequencies on the board. Compared to the ISO103 and ISO107, the ISO113 has the “Rectifier Filters” and “Driver” blocks.

The capacitive coupling problems are the resistors and amplifier noises, both of which are always present, the comparator noise jitter, and the residual output ripple noise due to the modulation. If we use, for example, a f_{in}=20kHz, we notice a ripple net error (Fig. 4.29).

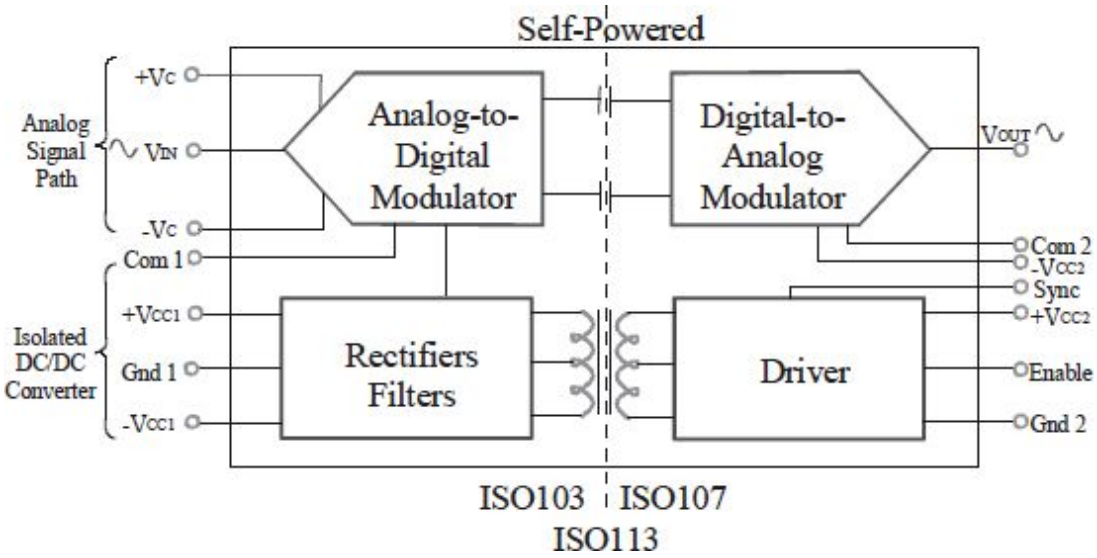


Fig. 4.27: ISO with an internal DC/DC converter.

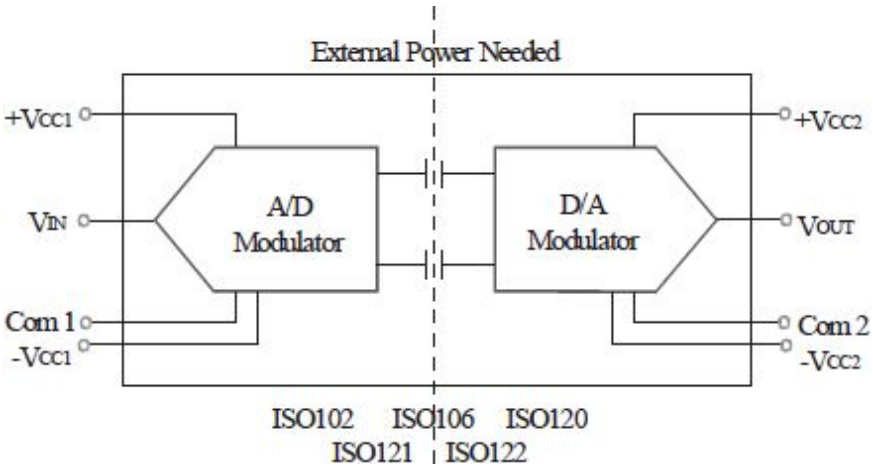


Fig. 4.28: ISO without an internal DC/DC converter.

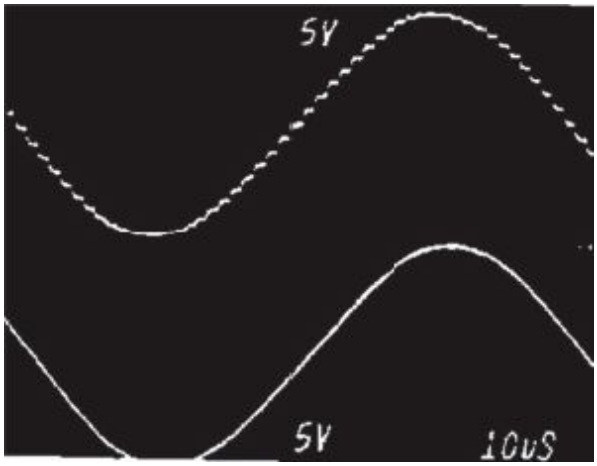


Fig. 4.29: Ripple at the output of an ISO122, without (top) and with (bottom) output filtering.

To reduce this error, it is necessary to filter the output that, as we can see, has the typical ripple at 500kHz of the PWM modulation. A Sallen-Key low-pass filter can be used for avoiding this ripple without compromising the 50kHz bandwidth of the amplifier. This filter and those within the ISO122 are three Butterworth poles, as shown in Fig. 4.30.

Thanks to the very low coupling capacitance values, this method is very noise insensitive because the demodulator receives only the differential edges. Incidentally, RF interference in the MHz range can interfere with the transmission or also overload the receiver, causing an error on the channel.

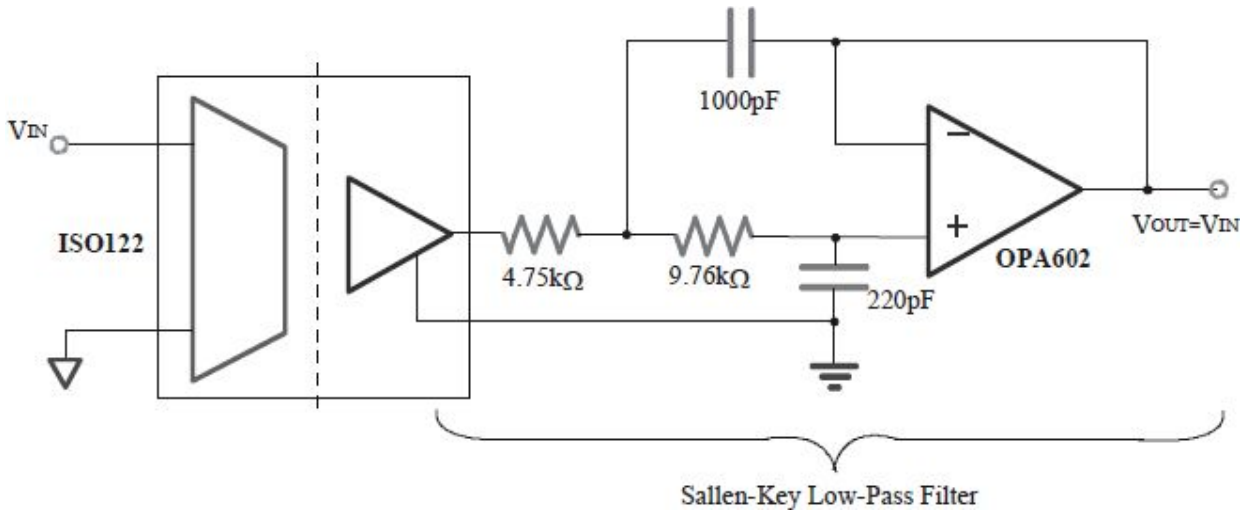


Fig. 4.30: Ripple filtering through output low-pass filter.

4.3. Current Feedb Amplifier (CFA)

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.

4.3.1 Voltage-Mode and Current-Mode approaches

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 ($Rout_1=R_{C1}$

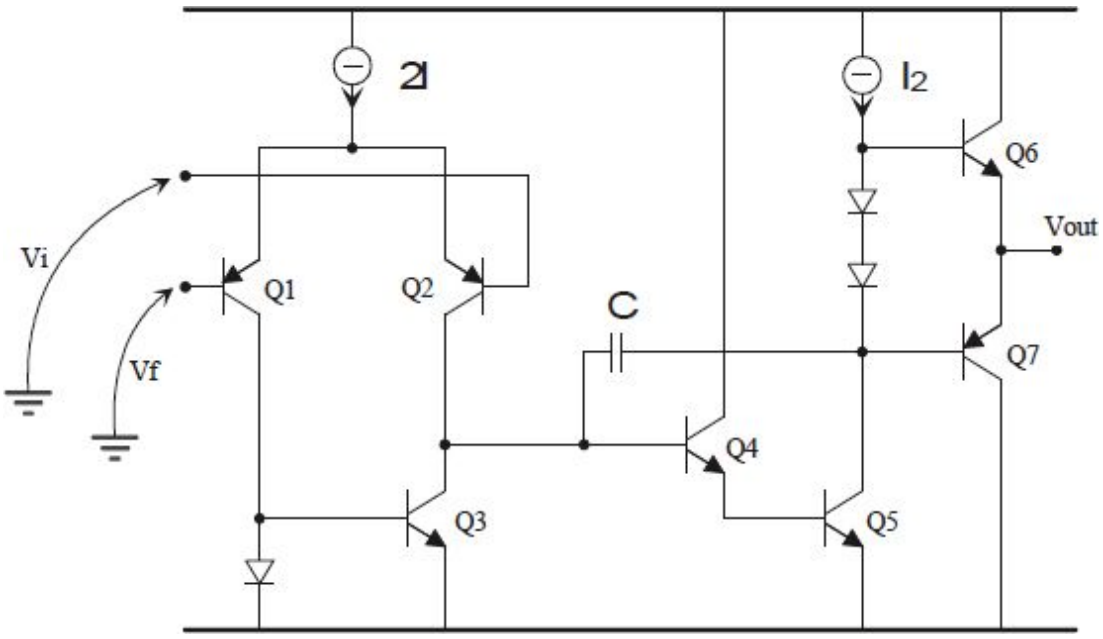


Fig. 4.31: VOA simplified schematics.

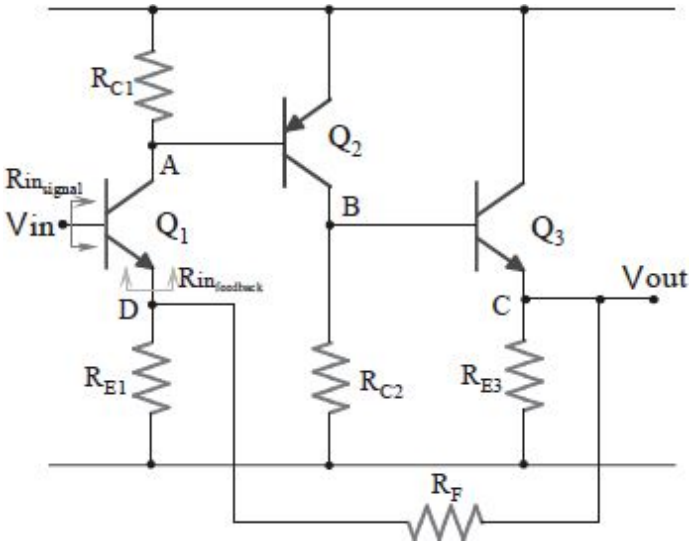


Fig. 4.32: Simple voltage amplifier with more stages.

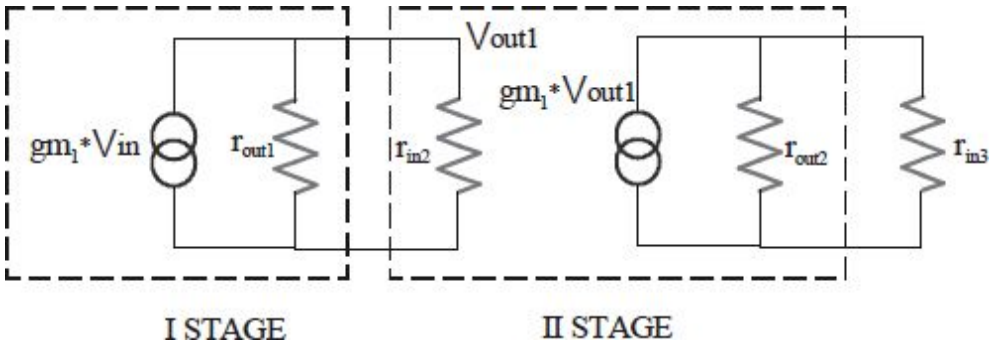


Fig. 4.33: Amplifier scheme: two transimpedance stages.

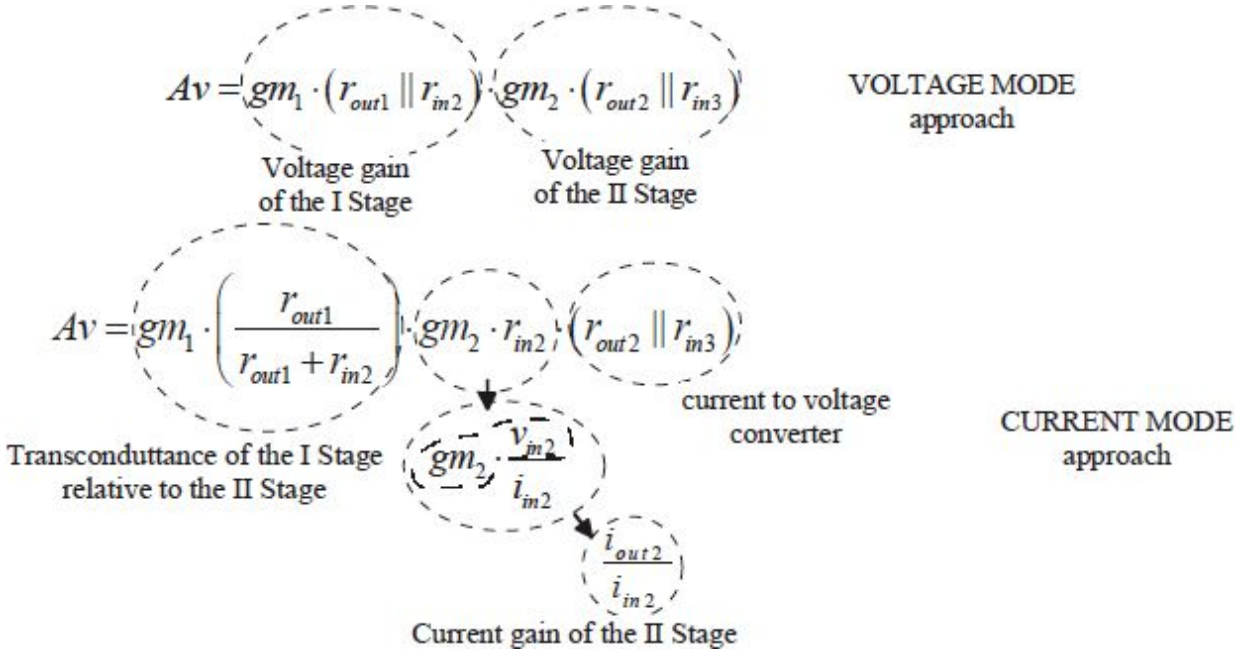
Let us think **in terms of currents**. The first stage produces a current proportional to the input voltage signal (*transconductance stage*). This second stage output current can follow three paths: re-flows into r_o of Q_1; flows through R_{C1}; enters into the base of Q_2. According to the hypothesis, 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 $\mathbf{R_{IN_{2°STAGE}} < Rout_1}$.

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 $R_{IN_{2°STAGE}}$.

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.

The circuit can be modeled as shown in Fig. 4.33 with two possible interpretations on the basis of the voltage or current mode.



In the preceding equations, we can notice that the second term is the transistor Q_2 h_{fe}, thus a constant term; the third term is fixed, too. We can act only on the first term to improve the gain: we need to lower r_{in2} with respect to R_{out1}. Only by making R_{out1} much greater than r_{in2} the fractional term can tend to the unity, and we reach an optimal condition for the maximum gain. Pay attention because, concerning the absolute value, it is more convenient that the resistance values will be small, thus the parasitic capacitance values should see low resistance values. In this way, the time constant should be small, ensuring greater bandwidth.

Nonetheless, what has been said thus far is valid only if it is possible to "connect" the real circuit to the theoretical schemes just discussed. In fact, it often happens that only considering the voltage mode approach, a better result is reached. This occurs, for example, in the case of circuits, such as the μA741 where, because of the presence of complex circuit elements, for instance, Darlington (required for better performance), what we examined differs from the principle scheme, and it happens that you come to the maximization of the gain only if it is the high impedance output node of the first stage, which is obvious if we are using the voltage mode approach.

However, we can say that using the current-mode approach leads to a design with low impedance nodes, which leads us to believe that the time constants of the circuit will tend to be small, so it is reasonable to hope that they can achieve broadband amplifiers.

4.3.2 Current Feedback Amplifier

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.: $\text{Gain}_{\text{closed loop}} \cdot \text{Bandwidth}_{\text{closed loop}} = \text{Bandwidth}_{\text{open loop}}$

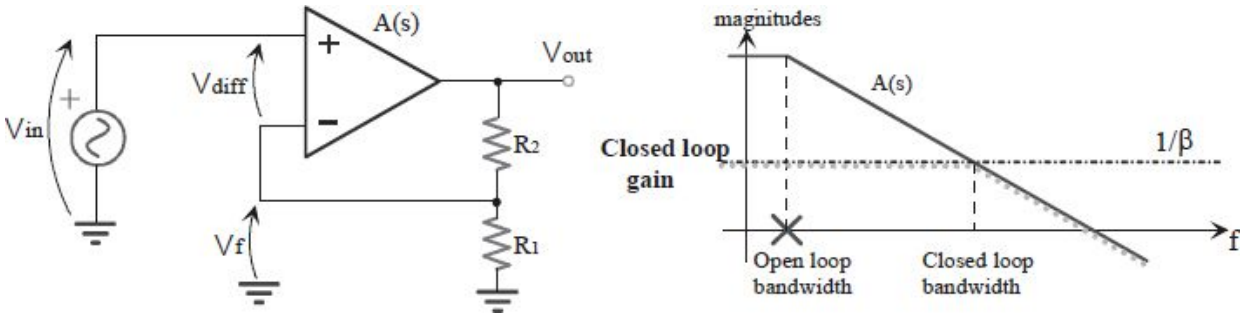


Fig. 4.34: Voltage Amplifier with a VOA: the GBWP is constant.

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.

Consider again the discrete component circuit seen before and shown again in Fig. 4.35 for your convenience. Observing the feedback, we realize that this is a current feedback configuration. In fact, the comparison is between current signals on the base-emitter junction of Q_1. Let us see how it works: given an input voltage signal Vin, a current signal is produced in the BE junction; the output signal Vout is translated into a current on R_F (the node D has low impedance, around 1/gm), which flows totally into the Q_1 emitter where it is compared with the input current signal.

Ideally, with closed-loop, the amplifier voltage gain will be: $G_c = 1 + R_F/R_{E1}$. To obtain the loop gain, we can cut on the node D and give a current signal. This current flows through the loop and arrives, modified by the path, to the output terminal where the resistance R_F is upstream the cut. If the following condition is verified $R_{E1} >> R_{in_{feedback}}$, than the feedback current flows into the Q_1 emitter, and then we can obtain the fundamental result that the circuit G_{loop} does not depend on R_{E1}. This means that this is a dominant pole system, and in first approximation, it is possible to modify the ideal gain (changing the R_{E1} value) without a corresponding change in bandwidth (because G_{loop} should not vary).

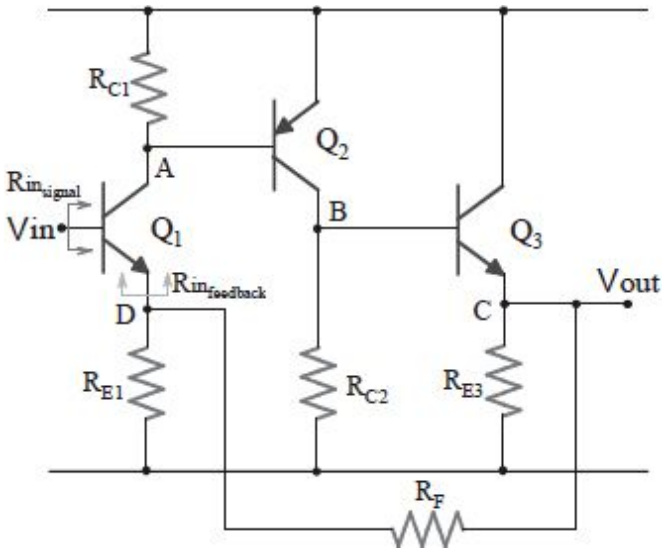


Fig. 4.35: Amplifier with the current feedback highlighted (CFA).

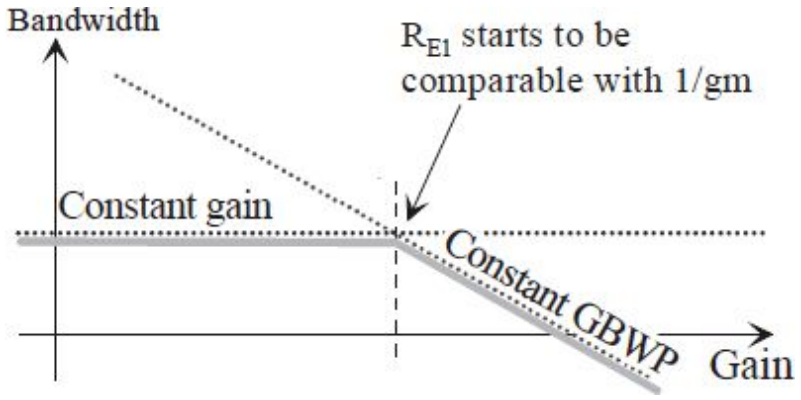


Fig. 4.36: Working conditions acting on R_{E1} to vary the gain.

This is substantially the idea behind the current feedback amplifier that deeply differs from the VOA. In fact, an amplifier with constant bandwidth does not have a constant GBWP with varying gain because the two terms are independent. Actually, what has been said is not completely accurate. In fact, the resistance R_{E1} can be varied as long as it is greater than the resistance seen within the emitter, otherwise the current will not completely flow in $R_{INfeedback}$ anymore. While the condition is not met, the loop gain depends on R_{E1}. For this resistance modifying the ideal gain, the bandwidth is modified, too. Actually, R_{E1} intervenes always in the G_{loop} expression, but its effect is negligible under the preceding condition (that all the current flows in $R_{INfeedback}$).

Such a system is in practice very unlikely to work. In fact, excessively changing R_{E1} to increase or decrease the gain, the bias conditions change so completely unacceptable in order to ensure the correct operation of the system.

4.3.3 CFA principle scheme

With a different designing approach, these amplifiers have the interesting property of allowing the user to vary the gain, obtaining the same bandwidth for each gain value. In many applications, it is very useful. However, the chance to change the gain (but keeping constant the bandwidth) occurred

only within a range of gain values (Fig. 4.36), when R_{E1} stays large enough compared to 1/gm.

To derive a general model of CFA, we can make some observations on the circuit depicted in Fig. 4.37. The first stage can be seen as a common emitter degenerated due to necessity (because only by degenerating it we can establish the feedback and use the BE junction of Q_1 as comparing elements between the signal and the feedback). However, it can also be seen as an emitter follower. In this case, its task is to follow the feedback node of the input signal. The important thing is that this first stage has high input impedance whereas the input impedance of the feedback must be low. An element that meets these characteristics is, for example, a voltage buffer.

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.

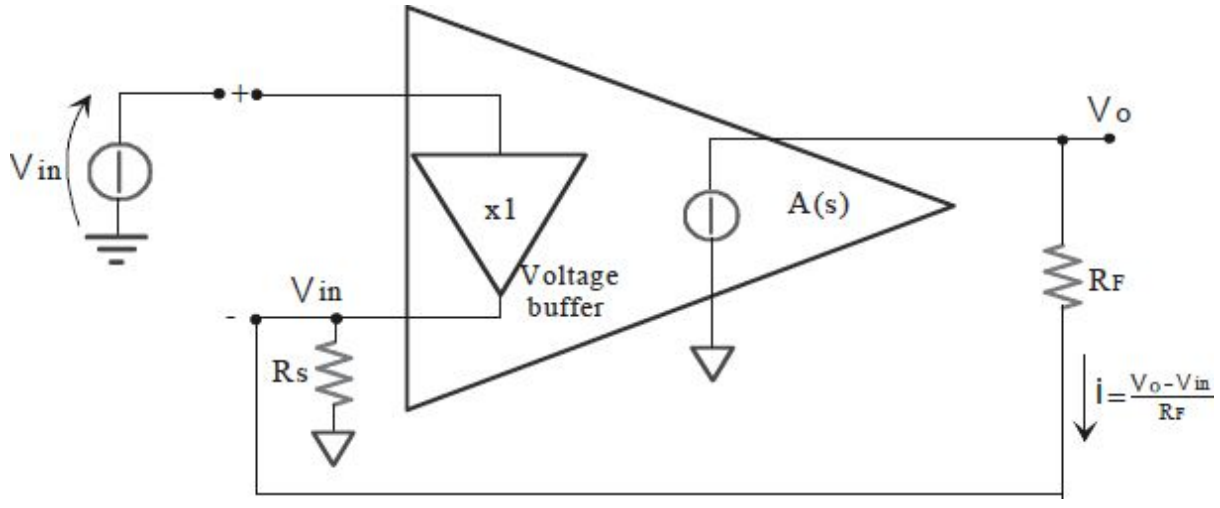


Fig. 4.37: Principle scheme for a CFA.

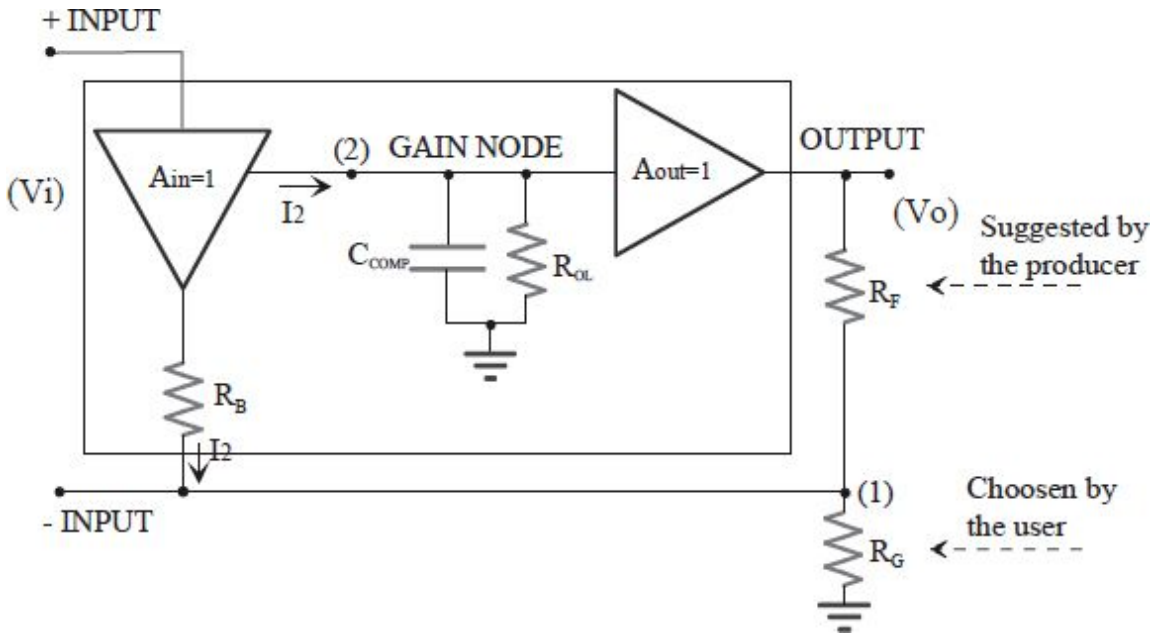


Fig. 4.38: Principle scheme for the CFA.

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: $\frac{V_1}{R_G} + \frac{V_1 - V_O}{R_F} + \frac{V_1 - V_{in}}{R_B} = 0$
from which

$$V_1 = \frac{\frac{V_{in}}{R_B} - \frac{V_O}{R_F}}{\frac{1}{R_F} + \frac{1}{R_G} + \frac{1}{R_B}}$$

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 $I_2 = \frac{Vin - V_1}{R_B} = V_1 \cdot \left(\frac{1}{R_G} + \frac{1}{R_F}\right) - \frac{Vout}{R_F}$ and because $V_{out} = A_{out} \cdot V_2 \cong V_2$ 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:

$$f_{pole} \cong \frac{A_{out}}{2\pi \left[R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B\right] \cdot C_{comp}}$$

For low gain, i.e. for: $1 + \frac{R_F}{R_G} << \frac{R_F}{R_B}$ it is simplified as: $f_{pole} \cong \frac{1}{2\pi \cdot R_F \cdot C_{comp}}$

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.

The inequality can be written as: $$R_B << \frac{R_F \cdot R_G}{R_F + R_G} = R_F \parallel R_G$$

It means simply that the output impedance of the input buffer must be less than the parallel of the two resistors placed outside the OpAmp by the user.

For great gain (higher than 50), the terms $(1+R_F/R_G) \cdot R_B$ becomes dominant, and the amplifier assumes a constant GBWP asymptotically because: $$GBWP = \frac{A_{out}}{2\pi R_B \cdot C_{comp}}$$

Observe Fig. 4.39. It is interesting to observe the results obtained by applying the theory of the classical feedback to the gain node (2). We have that: $$pole_{open} \cong \frac{-1}{C_{comp} \cdot \beta_{11} \cdot \beta_{14} \cdot R_F} \qquad Gloop \cong -1 \cdot \beta \cdot \beta$$

from which we can obtain the close loop pole:

$$pole_{closed} = pole_{open} \cdot \left(1 - G_{loop}\right) = \frac{-1}{C_{comp} \cdot R_F}$$

as obtained before in the case of low gains at closed-loop (Fig. 4.40).

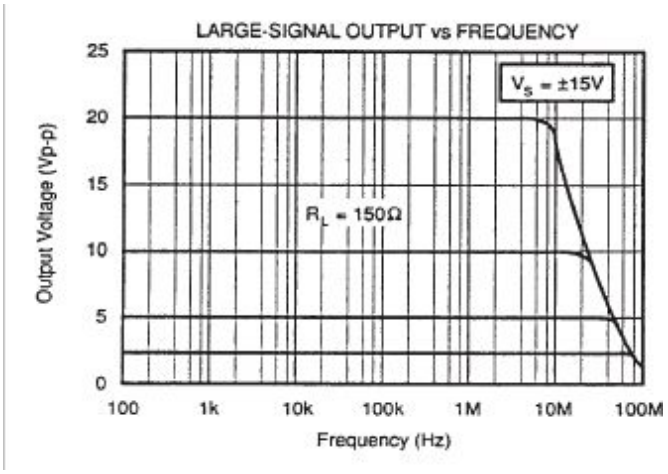


Fig. 4.39: Closed-loop gain.

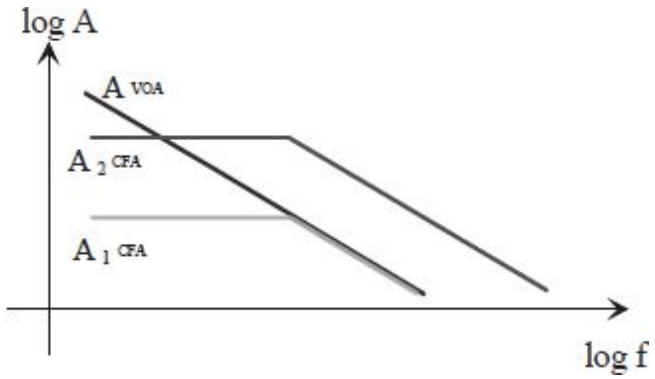


Fig. 4.40: Frequency responses.

Actual CFA structure

Once the current feedback principle has been understood, we can analyze the actual CFA amplifier as shown in Fig. 4.41. Although it seems like a complicated circuit, it is immediately attributable to the conceptual scheme just discussed.

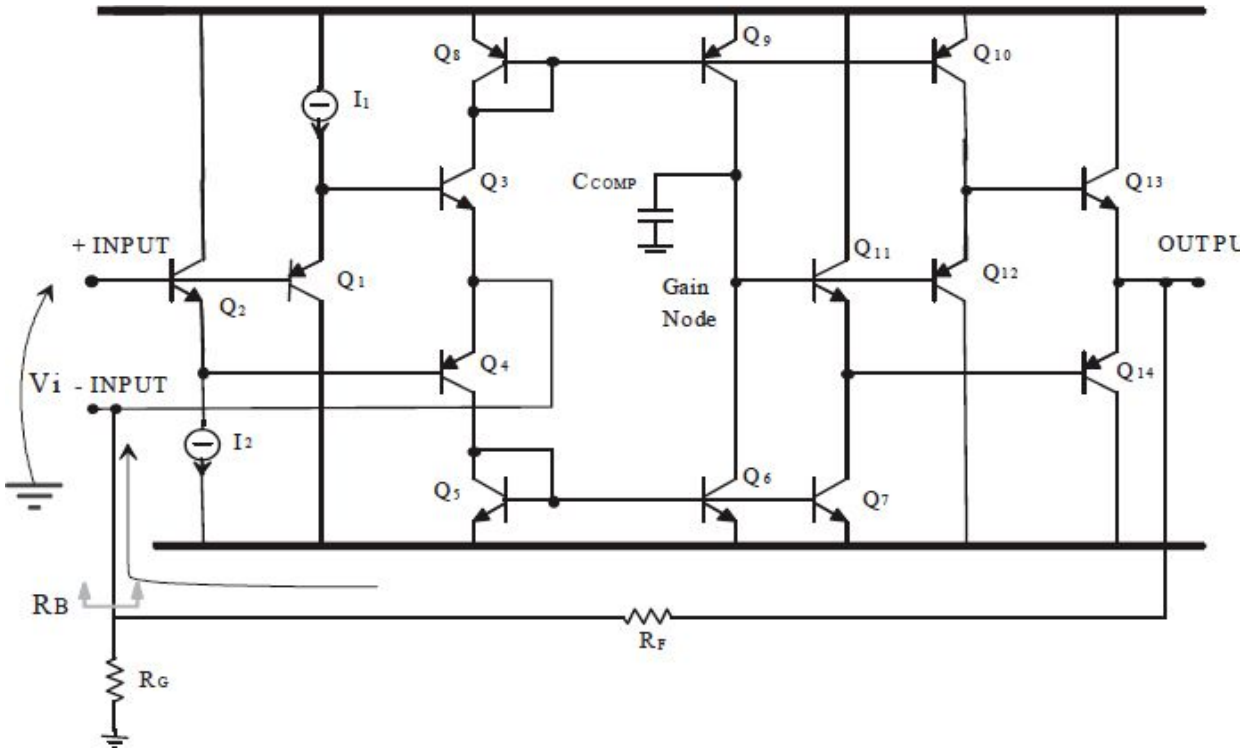


Fig. 4.41: Actual CFA simplified scheme.

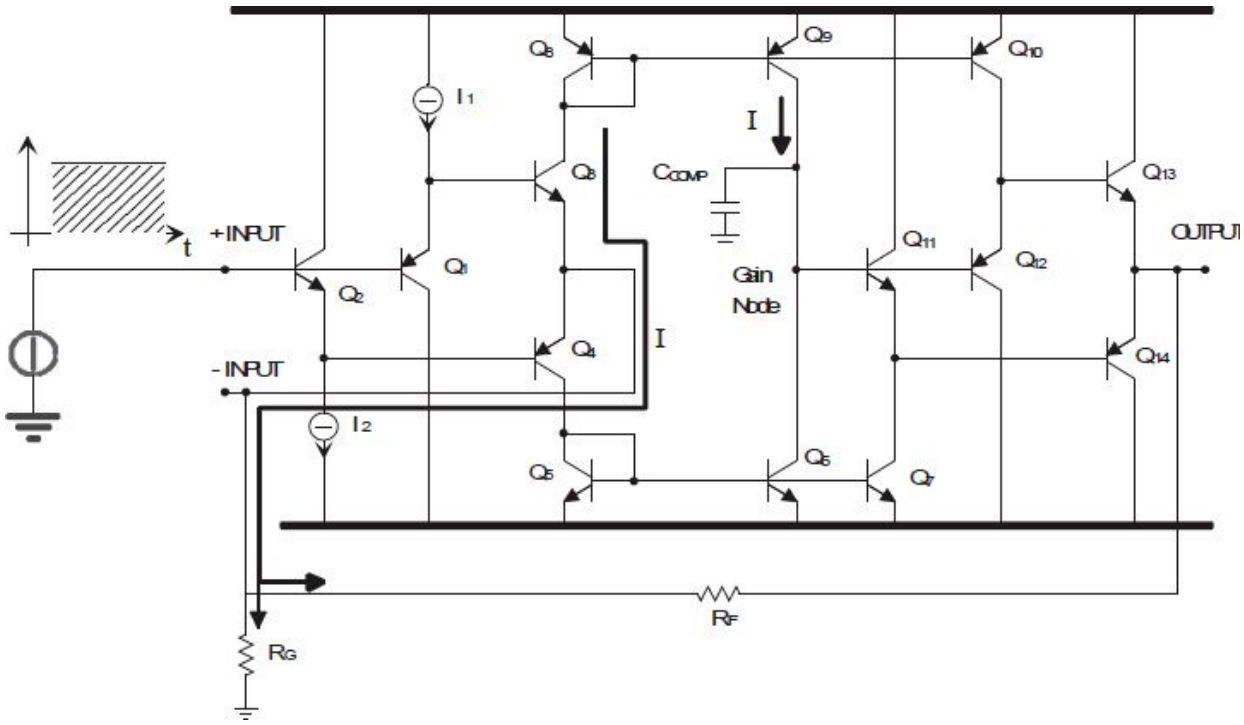


Fig. 4.42: Transient operation in response to a positive step input.

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:

$$SR = \frac{dV}{dt}\Big|_{\max} = \frac{I_{\max}}{C_{comp}} = \frac{1mA}{1pF} = 50\frac{kV}{\mu s}!!!$$

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.

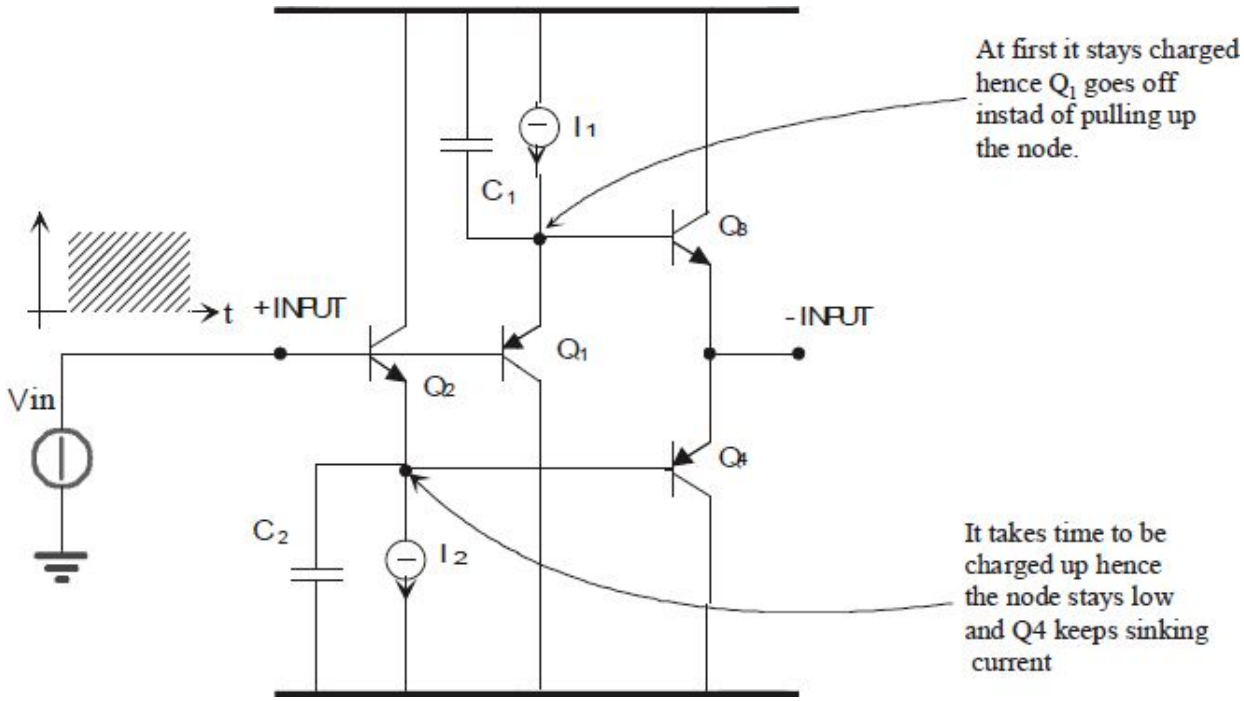


Fig. 4.43: Second order effects limiting the Slew Rate.

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.

		SR [V/µsec]	Bandwidth [MHz]	Offset [mV]
VOA	OP07CN	0.3	0.6	0.5
	µA-741CN	0.5	1	2
	AD818AN	2	2.7	3
CFA	EL2165	500	30	5
	LT1227	1100	140	10
	HFA1100IP	2500	850	8

Fig. 4.44: Comparison between CFA and VOA in terms of Slew Rate, Bandwidth, and Voltage Offset.

4.3.4 CFA drawbacks

CFAs, thanks to their performance, are used in high frequency circuits, such as the video amplifier. However, there are other contraindications compared to the voltage mode amplifier. For example, because of the input

stage asymmetry with respect to the differential one, CFAs have a CMRR worse than that of VOAs.

Here is another sore point: the **voltage offset**. If there were not the structure set up by the series npn-pnp (e.g. Q_2-Q_4), we would have an input offset voltage equal to V_{BE}, hundreds of mV! With this structure, the two junctions are automatically compensated, so the offset is neither very large nor very small. The problem is that it relies on the bipolar V_{BE} and, worse still, the "series" of an npn and a pnp (characterized by different Gummel numbers although it makes use of sophisticated and expensive technology).

The **offset thermal drift** is high. For example, if in the circuit presented before the positive input potential increases, V_{CE2} decreases, Q_2 dissipates less power, and the value of V_{BEnpn} increases significantly. On the other hand, V_{CE4} increases, then V_{BEpnp} decreases because Q_4 heats up. Unfortunately, the thermal drifts do not cancel each other. Assuming a temperature coefficient of 70°C/W, with an input of 1V, we have power dissipation that changes of 1mW if the stage is biased at 1mA. This results in a change in temperature of 0.14°C. With a typical coefficient of -1.7 mV/°C for V_{BE}, there is a variation of ΔV_{BE}=240μV, resulting in a double-error total of 0.48mV per volt.

Another problem is due to the **input bias currents**, which are different: at the positive terminal, I_{Bpnp} is subtracted from I_{Bnpn}, but such a subtraction is imperfect. Typically, there remains a difference in the value of 5μA flowing out of the OpAmp. This current may result in a significant voltage drop (50mV on a 10KΩ source resistor). Different and worse still is the case with the negative terminal where these currents can reach values up to 100 times higher.

Another disadvantage is the limited CMRR of the CFA. In fact, changing the voltage at the positive node, the negative node voltage changes (determining Vout which depends on R_G and R_F) and determines a difference current in Q3 and Q4 transistors due to r_o. Only for $V_{in}^+ = 0$, we will have an output voltage equal to zero (over the offset). For all the other values, the output will be slightly different from the theoretical value because of the Early effect. We can demonstrate that:

$$CMRR = 20\log\left(\frac{V_{EARLY}}{2V_{th}}\right)$$

that is the same expression as that seen in the differential stages with single-ended output. For V_{EARLY}=100V, we will have a CMRR equal to only 66dB, very lower than 100dB obtainable by a simple VOA. The obvious remedy would be to add an input cascade that would introduce slew rate problems.

A further error source is the R_{OL} transresistance that is present in the following expression:

$$\frac{R_F + \left(1 + \frac{R_F}{R_G}\right) \cdot R_B}{R_{OL}}$$

which is not infinite. A simple current mirror usually has $r_o \cong 100 \div 200 k\Omega$. To increase it, we could use a Wilson mirror ($R_{OL} \cong 1M\Omega$) that improves the mirroring precision ($1/\beta^2$) (as shown in Fig. 4.45).

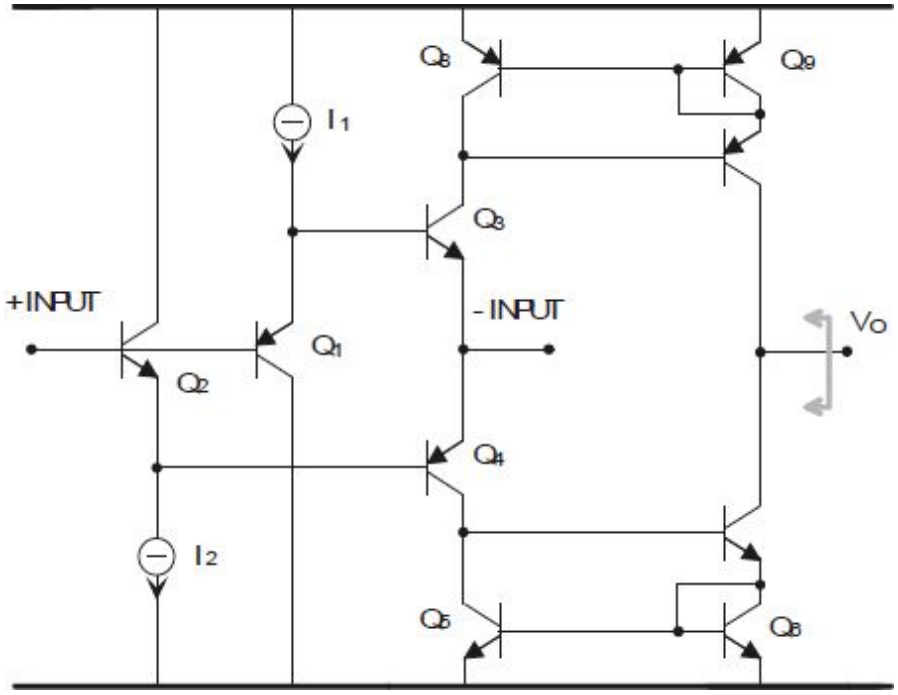


Fig. 4.45: The R_{OL} transresistance generates error.

Because typically the feedback resistor R_F varies in the range 250Ω÷1kΩ, the error on the gain varies in the range 0.1%÷0.7%.

The term containing R_B (in the order of 10÷25Ω) is almost irrelevant for low gains while, for high gains, R_B tends to assume a value close to R_G and therefore can cause problems. However, keep in mind that the CFA is usually used for low gains to take advantage of the constant bandwidth characteristics!

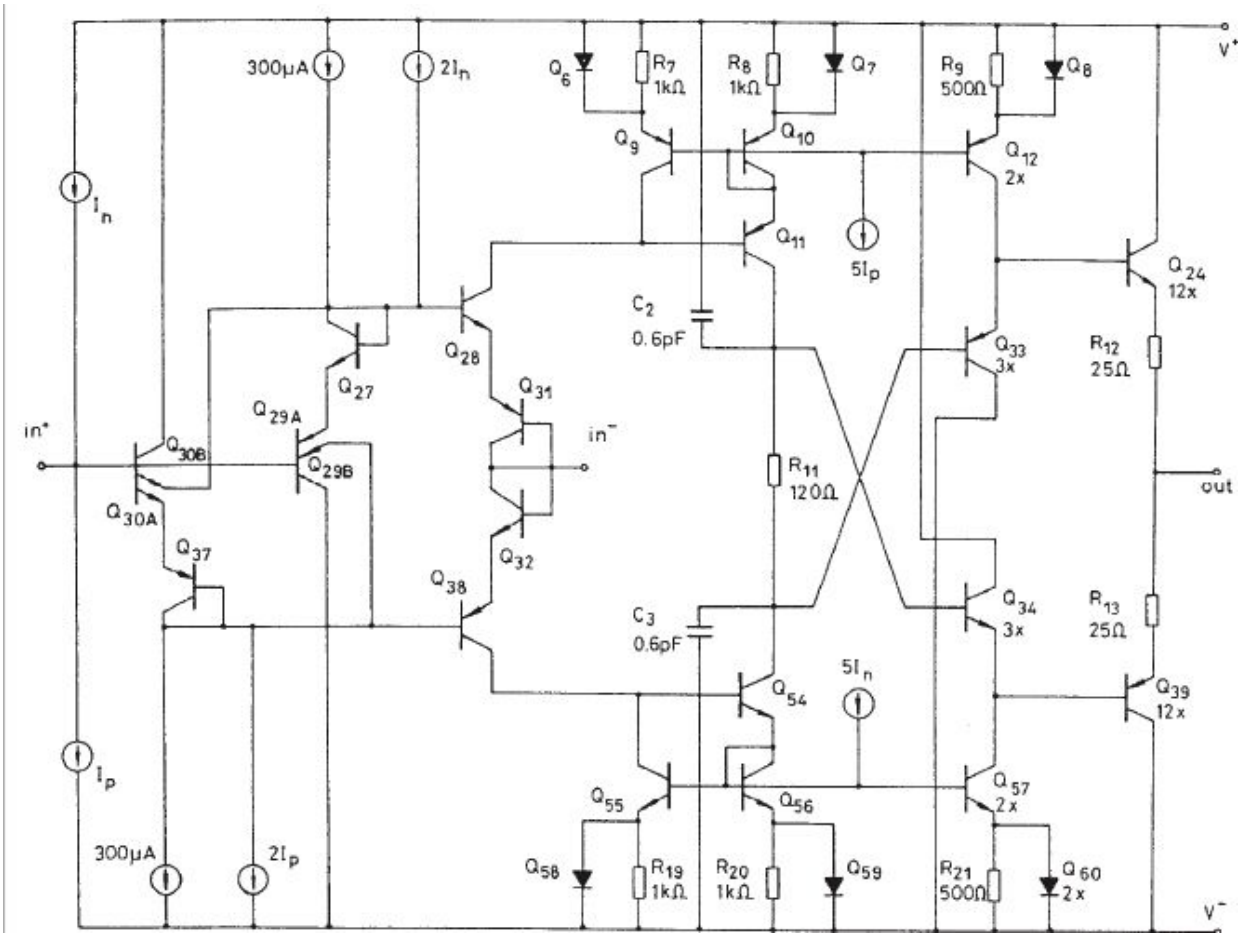


Fig. 4.46: OPA 260.

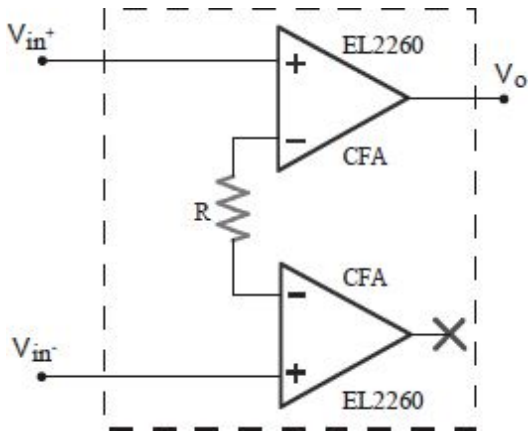


Fig. 4.47: Configurations to improve the CFA accuracy.

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.

To improve the accuracy, we can use dual configurations. We can first start to consider a circuit purely conceptual, not real (Fig. 4.47). The two CFAs work as a normal “voltage mode”. The lower CFA offset voltage will cancel (quasi) the upper CFA offset. Only the CFA mismatch will remain; the bandwidth does not degrade and depends only on R, which behaves as the feedback resistance.

Unfortunately, the circuit presented does not control the lower CFA output voltage that can vary between the power supplies, causing heavy imbalances internal to the OpAmp. Furthermore, the upper CFA negative input pin current produces a non-compensated offset. A natural development of this

circuit is shown in Fig. 4.48. In this way, the output of the bottom CFA of Fig. 4.1 is stabilized: the output follows the negative terminal as a buffer. The problem of current offset, however, remains. To remedy this problem, we can also make a further change to the circuit (Fig. 4.49).

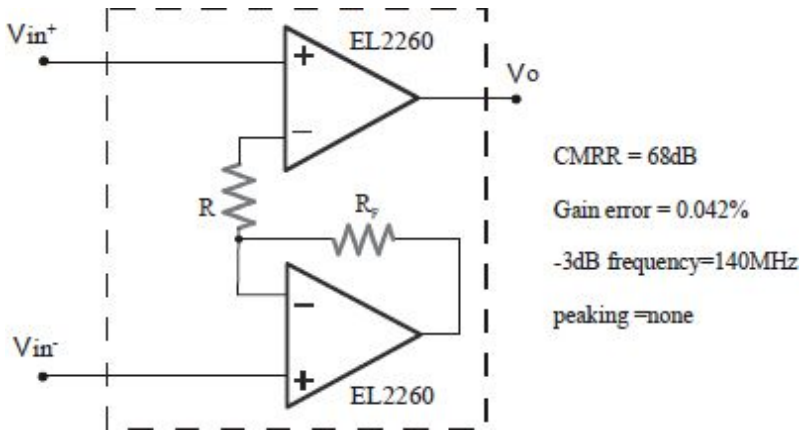


Fig. 4.48: Improved circuit (compared to that of Fig. 4.47).

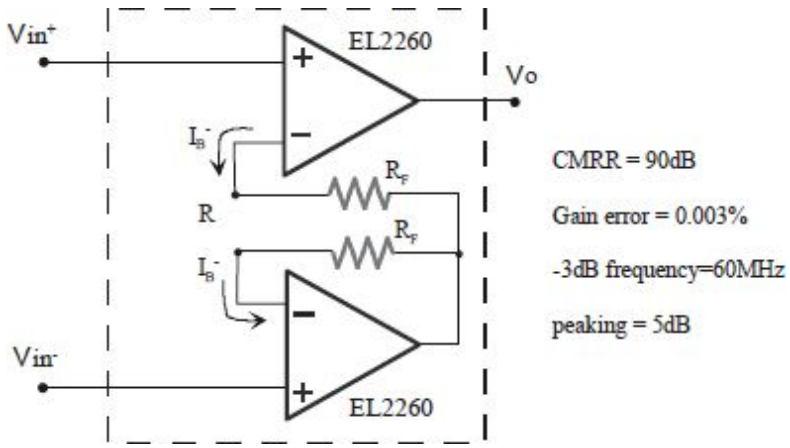


Fig. 4.49: Scheme to resolve the offset problem.

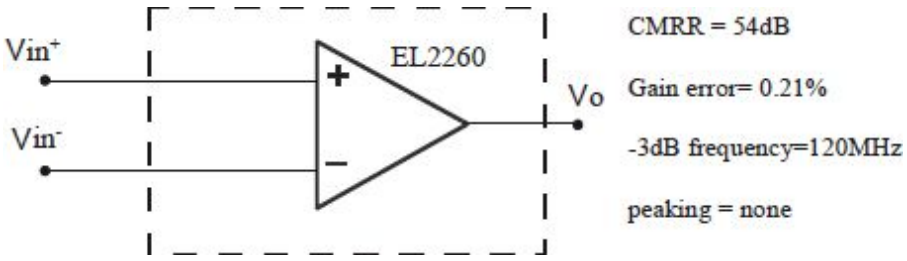


Fig. 4.50: Single CFA and relative values with principal characteristics.

The output voltage of the lower CFA follows the actual offset voltage up to less than $R_F \cdot IB^-$ of both the upper and lower CFAs. These contributions cancel the offset voltage of the upper CFA. We observe that the upper CFA has a gain dependent on the external network while the lower one will always have a gain of 1. Consequently, the gain will be accurate even if the upper CFA is a buffer; otherwise, the error due to R_{OL} will not be canceled. The three circuits have performance gradually improved in terms of accuracy in DC while bandwidth slightly gets worse than in the corresponding single-CFA (Fig. 4.50), and, in the third case, a peak appears as a stability problem, as shown by the data presented Fig. 4.49, on the right.

Let us look at the CFA in terms of noise. Even for the current feedback, we can model the equivalent input noise generators. To estimate Vn and In for the CFA, it is necessary to take into account different noise sources from those in the case of the VOA since the two circuits are internally very different. However, once in possession of the equivalent circuits, it is possible to use the scheme shown in Fig. 4.51.

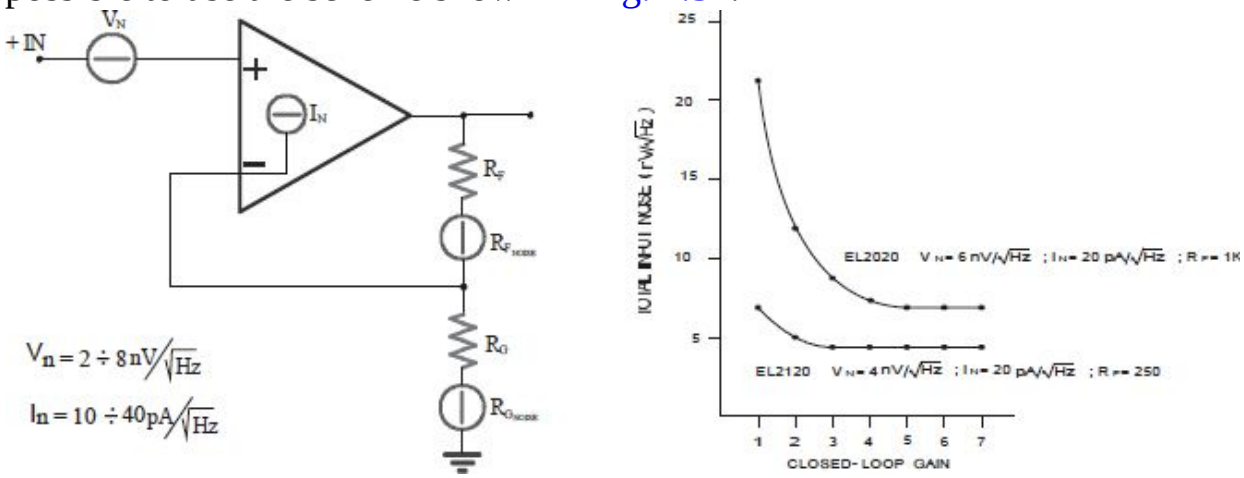


Fig. 4.51: Equivalent noise source for the CFA (on the left) and total input voltage noise for the CFA as a function of the gain (on the right).

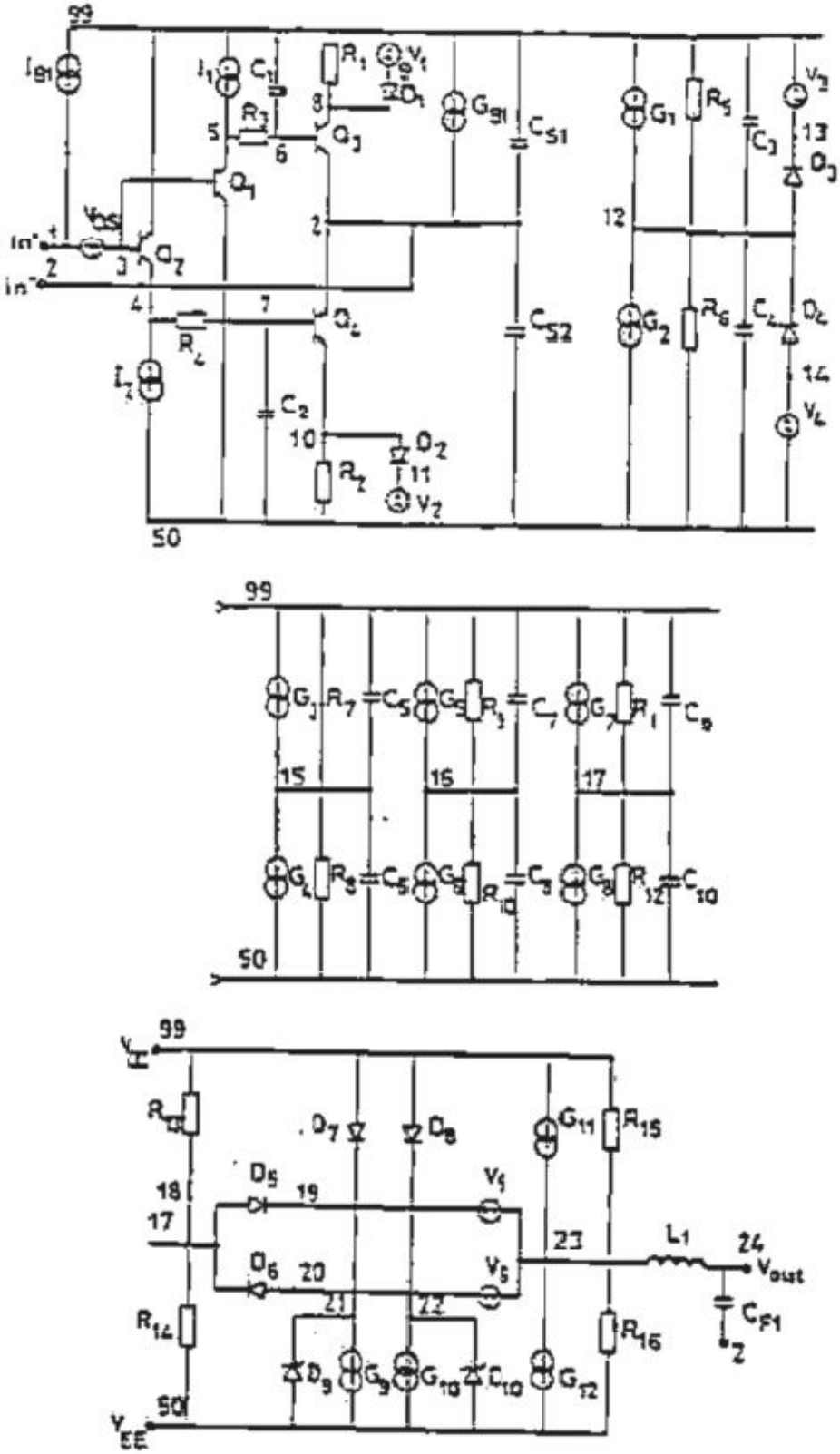
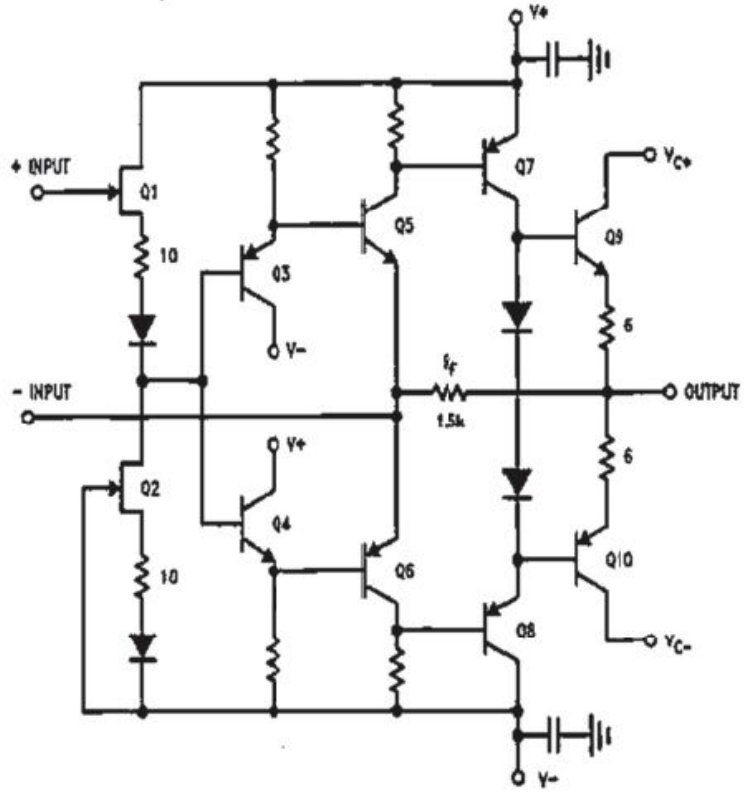


Fig. 4.52: SPICE models for CFA.

The same Fig. 4.51 shows the graph of the total input voltage noise (non-inverting configuration) for two different CFAs with slightly different performance varying the R_G (and therefore the gain since R_F is recommended by the manufacturer). The deterioration for low gain is due to high R_G.

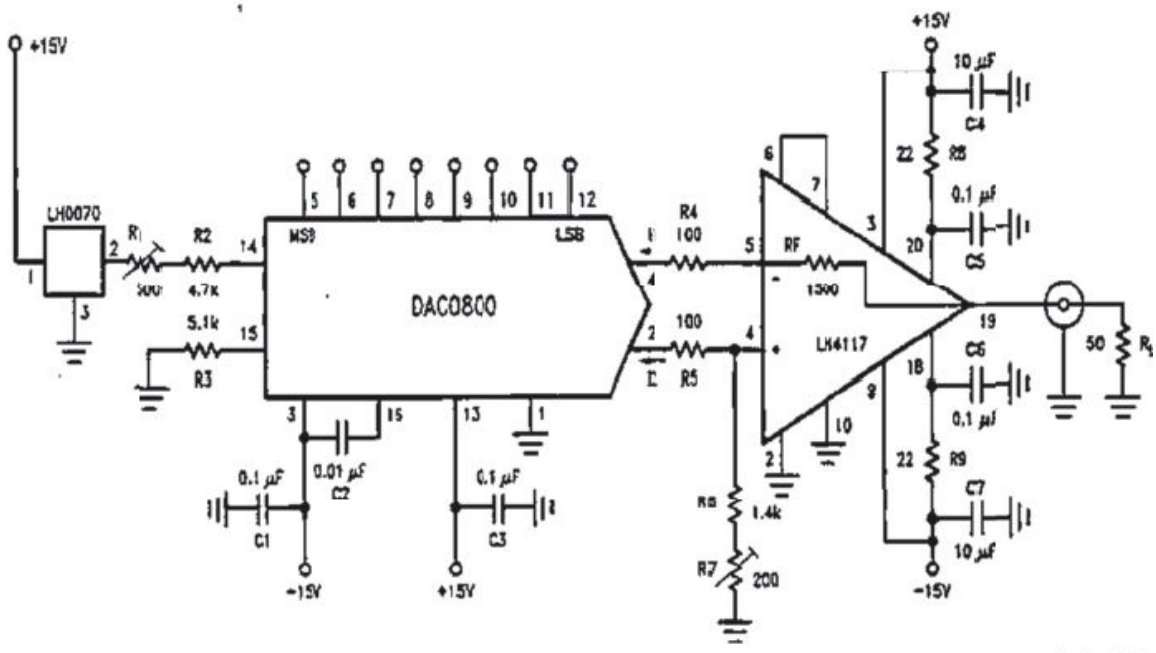
As for the VOAs, also the CFAs must be correctly modeled in SPICE; otherwise, we will find unreliable results. The Fig. 4.52 shows some sample schemes.

In some CFAs, the resistance R_F is included in the IC, as in the case of Fig. 4.53. For other CFAs, the IC internally has the optimal R_F at which another resistance can externally be added (Fig. 4.54).



TL/K/10377-9

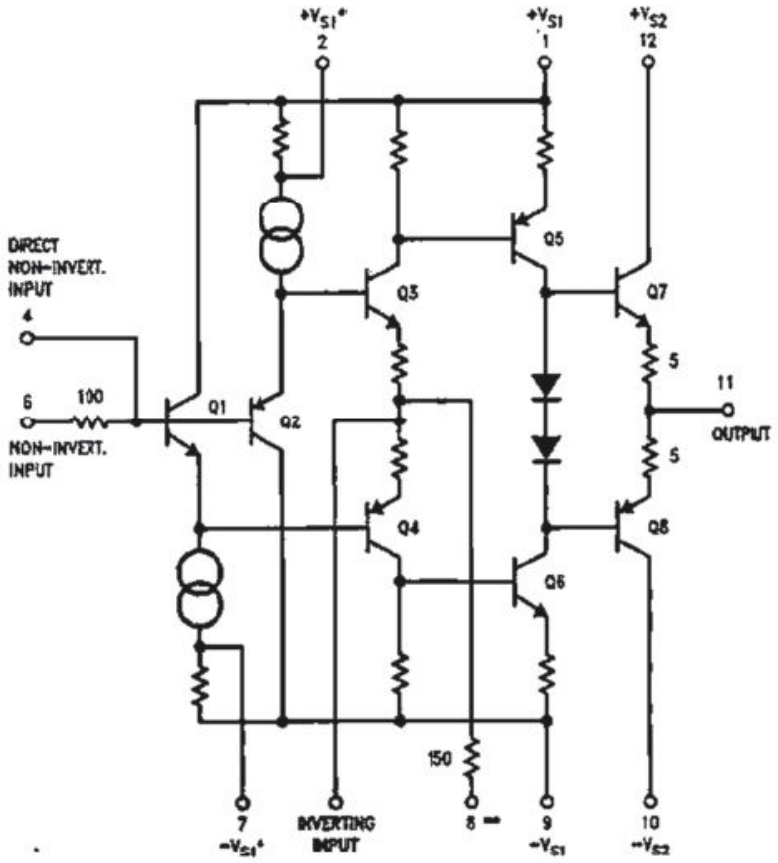
FIGURE 6. Simplified Schematic of the LH4117



TL/K/10377-

FIGURE 7. Fast Current-to-Voltage Current Mode DAC Amplifier

Fig. 4.53: Example of a commercially-available CFA with internal R_F.

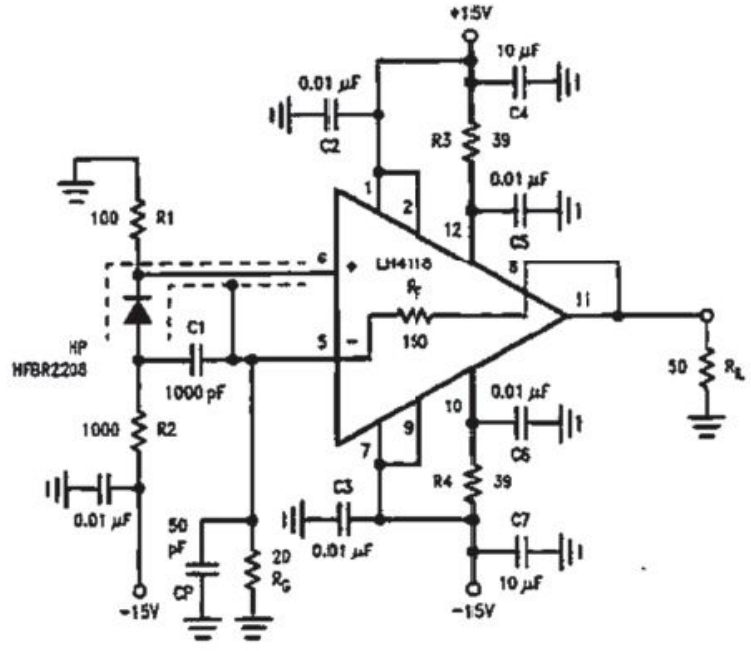


TL/K/10377-12

*Pins 2 and 7 can also be left disconnected (floating).

**The built-in 150Ω can be used as feedback resistor for $A_V = 1$.

FIGURE 9. LH4118 Simplified Schematic



TL/K/10377-13

FIGURE 10. Fiber Optic Receiver Using the LH4118

Fig. 4.54: CFA with optimal internal R_F resistance.

4.4. Current Mode Amplifier (NORTON)

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.

4.4.1 Current Amplifier Input Stage

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 $I_{in} << I_0$, 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.

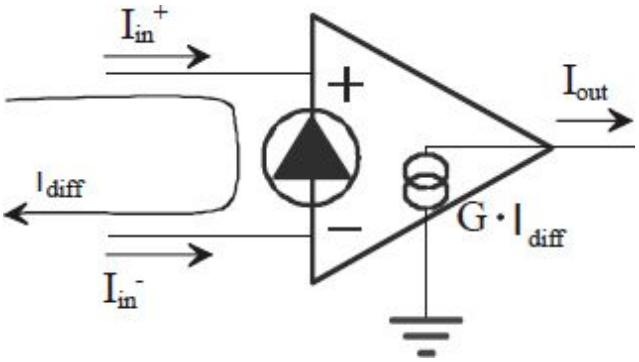


Fig. 4.55: Current Mode Amplifier: both inputs have low impedance values.

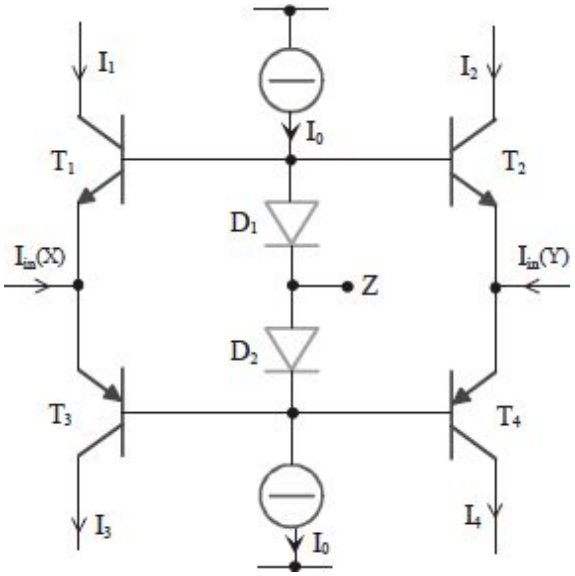


Fig. 4.56: Internal structure of the Current Mode Amplifier.

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. 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.

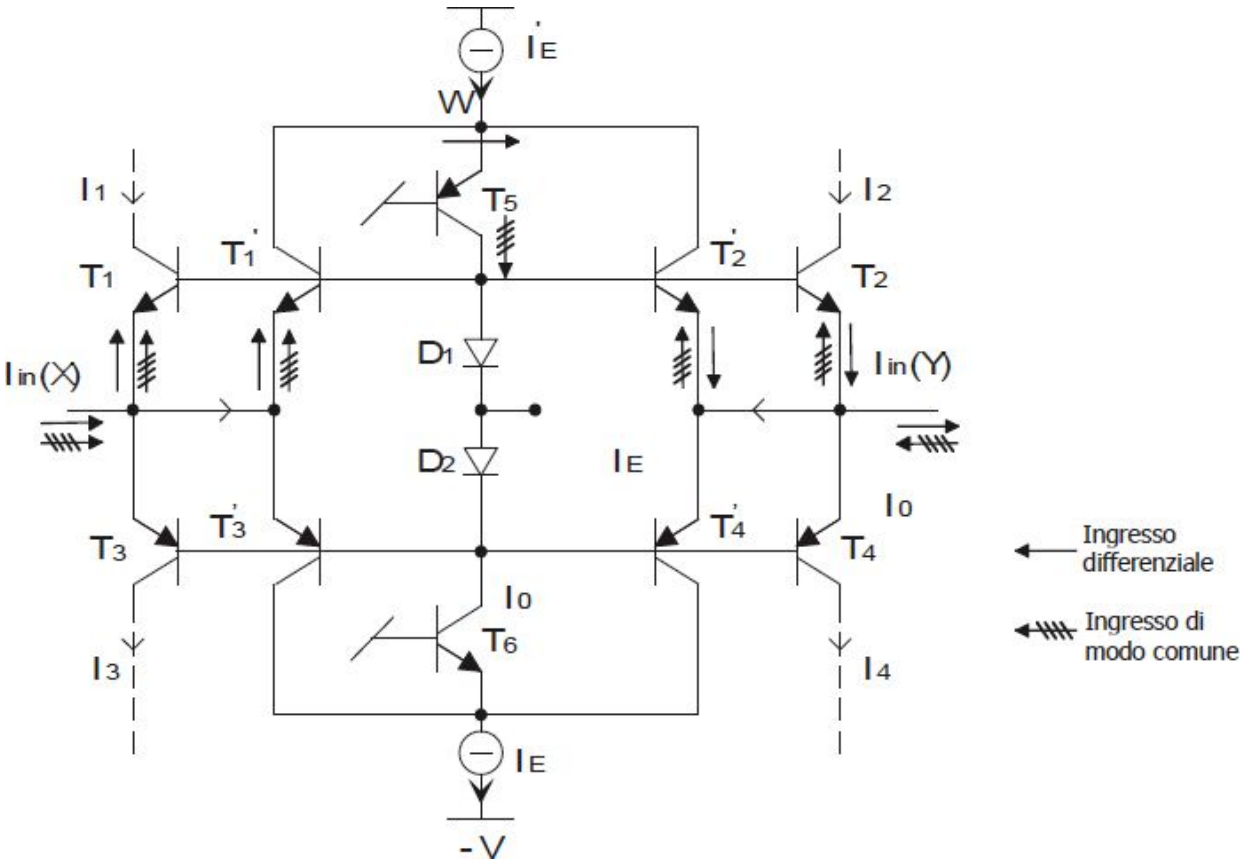


Fig. 4.57: Input stage with an improved CMRR.

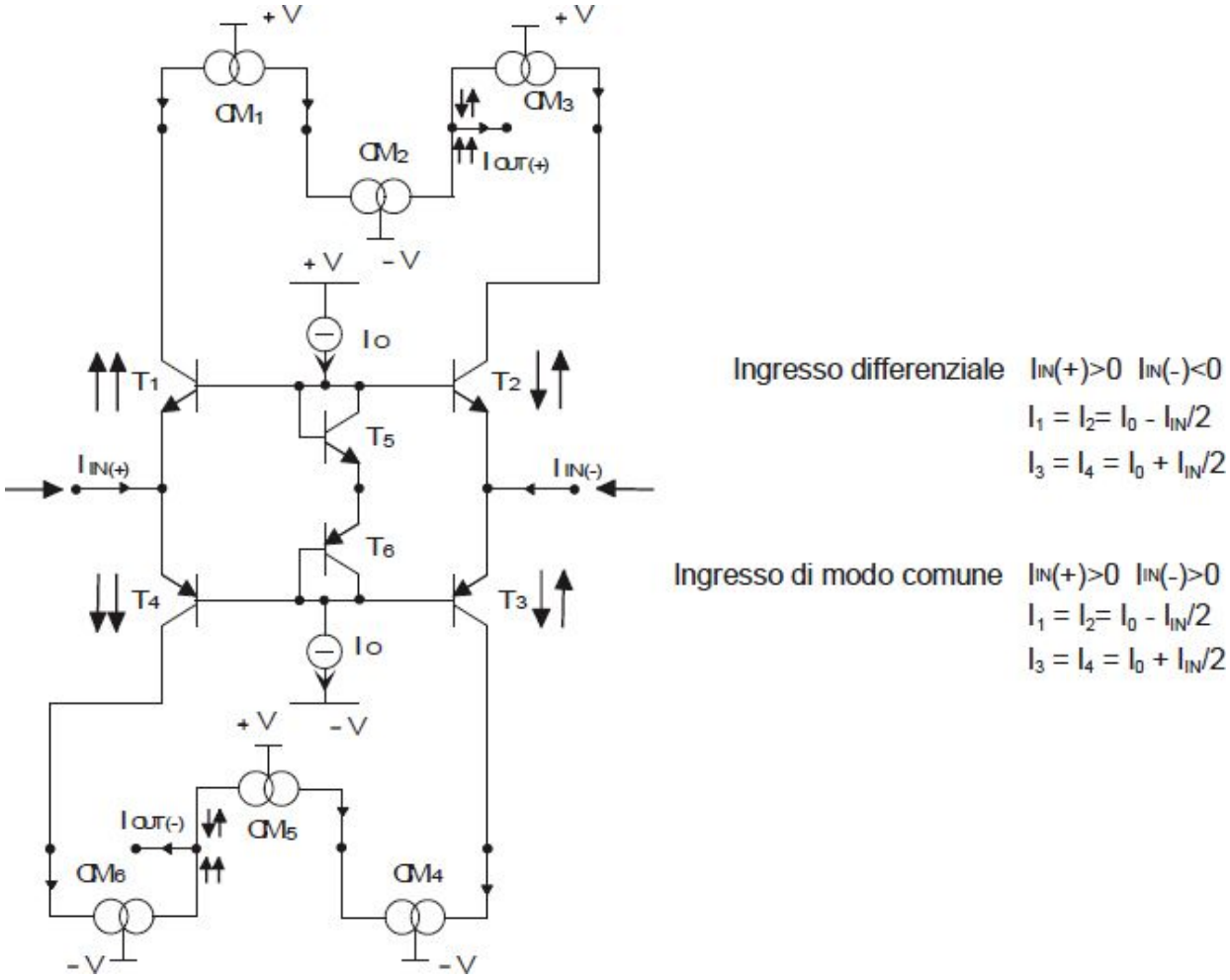


Fig. 4.58: Input stage with six current mirrors.

To amplify the current signal, we could send the signal in the base of a bipolar transistor. The problem is that the BJT base is not a good current well. Another mode is to use Gilbert cells that are circuits with four BJTs cross connected as shown in Fig. 4.59.

4.4.2 Gilbert Cell

Gilbert Cells were originally introduced to multiply signals. The fundamental equations are the following: $i = gm \cdot v_{BE}$ and $gm = \frac{I_C}{V_{th}}$. By linearly varying the transconductance by means of another signal (to have $gm = k \cdot v_{sec}$), the output will be $i \propto v_{BE} \cdot v_{sec}$, i.e. proportional to the product of the two signals.

The Fig. 4.59 shows the Gilbert cell scheme. Notice that the transconductance is determined by the tail that provides the bias current. Thus, using another signal, we can control I_{POL}: we reach the goal. A current proportional to a voltage signal can be simply obtained: you can think, for example, an emitter degenerated current generator driven by the signal $i \approx v/R_E$. This circuit is an analog multiplier. Let us see how it works. The transistors Q_2 form a differential stage. The BJT couple of Q_1 is biased with a current I whose value is due to a reference voltage applied to the bases.

We apply a differential current i_1 on the external branches. This current, in first approximation, flows in the emitters of Q_1 (less resistive than the bases of Q_2). We can focus our attention on the right side: i_1 corresponds to a reduction in the current that Q_1

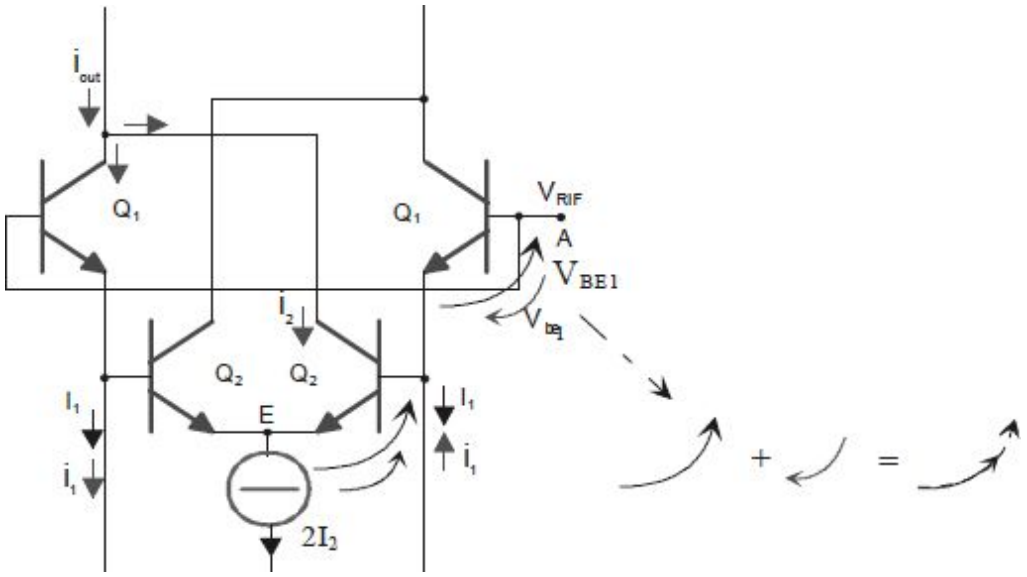


Fig. 4.59: Gilbert cell.

brings; thus, we have a corresponding reduction in V_{BE1}. Because of, in first approximation, fixing the points A and E on a differential signal, we will have a corresponding increase in V_{BE2} which produces a current signal i_2. The equations of the behavior are the following:

a) $i_1 = gm_1 \cdot v_{BE1} \quad gm_1 = \frac{I_1}{V_{th}}$ b) $i_2 = gm_2 \cdot v_{BE2} = gm_2 \cdot v_{BE1} \quad gm_2 = \frac{I_2}{V_{th}}$

From the equation a) we obtain: $v_{BE1} = \frac{i_1}{gm_1}$

which substituted in the equation b) brings to:

$$i_2 = gm_2 \frac{i_1}{gm_1} = \frac{I_2}{V_{th}} \cdot \frac{i_1}{\frac{I_1}{V_{th}}}$$

from which we have:

$$G = \frac{i_2}{i_1} = \frac{I_2}{I_1}$$

Because the collectors of Q_1 and Q_2 are connected as in the figure, we have:

$$i_{out} = i_1 + i_2$$

thus, recovering some preceding relationship:

$$i_{out} = i_1 + \frac{I_2}{I_1} \cdot i_1$$

and substituting the latter in the gain expression, we have the required result:

$$G = \frac{i_{out}}{i_1} = \left(1 + \frac{I_2}{I_1}\right)$$

Examining the hypothesis made to reach this result, we have:

- small signal hypothesis (gm=I/Vth);
- hypothesis for which all the signal current “i” will flow in the emitter of Q_1 (equivalent to say that the base current of Q_2 is negligible).

The former is non-binding because we deal with a mesh and can use the transfer characteristics for large signal to obtain the same result. The latter implies that the Q_2 steady current is not too high, and then the gain is

limited. If we are satisfied with a factor of 10 between the base of Q_2 and the Q_1 emitter current, and if h_{FE2}=200, the maximum gain is G=1+20.

This is a cell with a current gain. To obtain a higher gain, we need to use transistors with a high gain. The obtainable gains are moderate, but using a cascade of cells, it is possible to obtain higher values. This circuits have wider bandwidth values and are often used in open-loop; for this reason, the required gain is not too high, generally few tens.

4.4.3 Current Amplifier

To have a high gain for the circuit depicted in Fig. 4.58, we must use Gilbert cells. We obtain the current amplifier of Fig. 4.60. In this way, the small signal gain is equal to: $$G = A_i = \left(1 + \frac{I_G}{2I_0}\right)$$
Therefore, varying the bias current I_G, we can modify the current gain. Moreover, I_G determines the maximum output current limit.

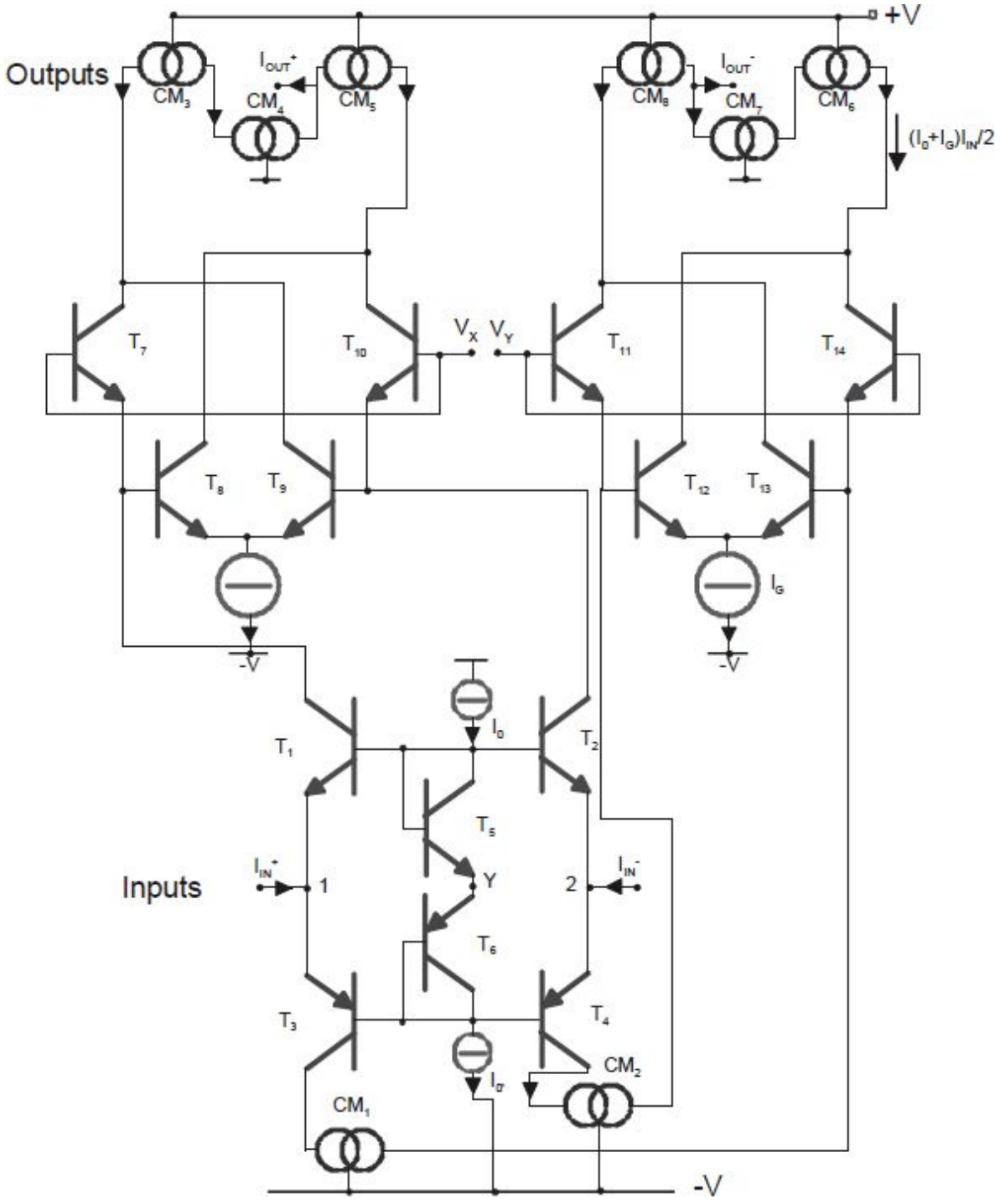


Fig. 4.60: Current amplifier.

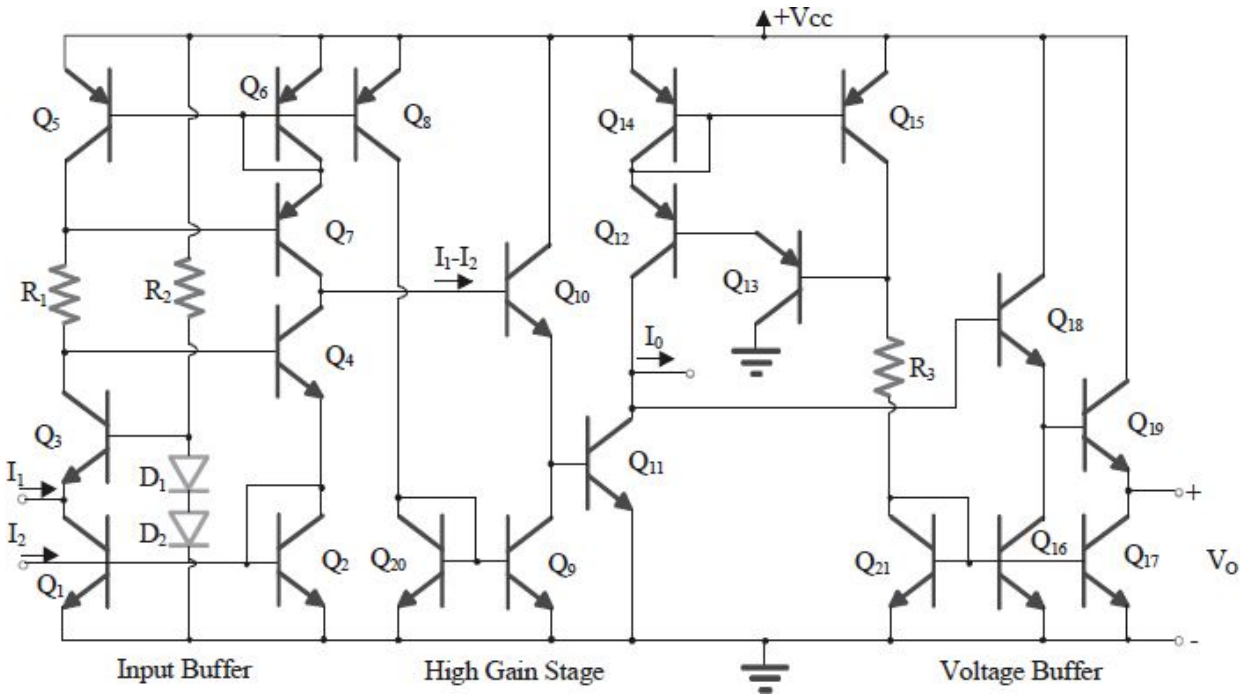


Fig. 4.61: Current amplifier (without Gilbert cell).

The voltages V_X and V_Y are the references of the Gilbert cell and must be externally set. Choosing $I_0 \cong 100\mu A$ and varying I_G between 1mA and 3mA, the current differential gain varies between 5 and 15. We can notice that each OpAmp internal node has very low impedance, up to 2/gm. This cause parasitic capacitors to determine high frequency poles.

In addition to this circuit, there are other circuit topologies which allow to make a difference between two input currents and give an output current I_0. An example is the circuit shown in Fig. 4.61. The output stage will provide the output voltage V_0. Another example is the circuit topology used by National Semiconductor in the LM359.

In literature, there is often much confusion of acronyms and terminology. It happens, for example, that these OpAmps, whose both input terminals have low impedance values (they act as current readers) and whose output can either be a current or a voltage, are named “NORTON”. For all these cases, with an external feedback of the OpAmp, the output will change in voltage to ensure that the differential input current is null (as for VOA with the differential input voltage). For the circuit in Fig. 4.62, despite the

similarity with a typical VOA, we must write the following equations:

$$I_1 = \frac{V_1}{R_1} \qquad I_2 = \frac{V_2}{R_2} + I_0 \qquad I_0 = \frac{V_0}{R_F}$$

in which we suppose that the OpAmp is built up with an input close to the ground (0V), and the impedance values are infinitesimal.

Because the OpAmp (open-loop) has a gain for which $I_0 = A_i \cdot (I_1 - I_2)$, we obtain an overall gain (closed loop) that is: $V_o = \frac{A_i}{1+A_i} \cdot \left(V_1 \cdot \frac{R_F}{R_1} - V_2 \cdot \frac{R_F}{R_2} \right)$

Therefore, the gain for the single inputs is R_F/R_1 or R_F/R_2 and, although they resemble the VOA equations, actually they are conceptually different. In fact, now there is a flowing current in both the input terminals.

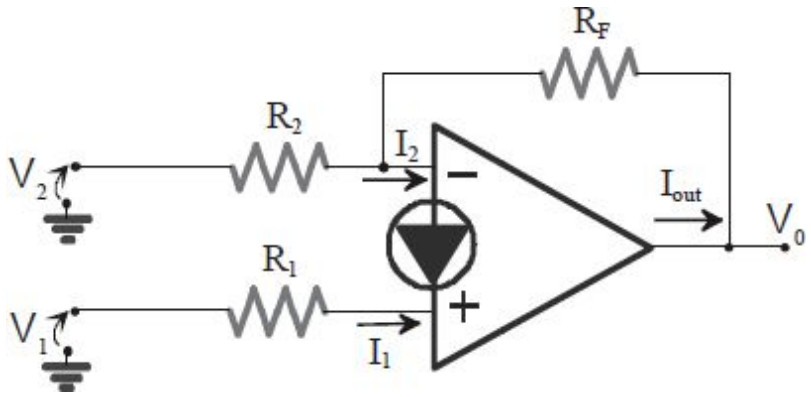


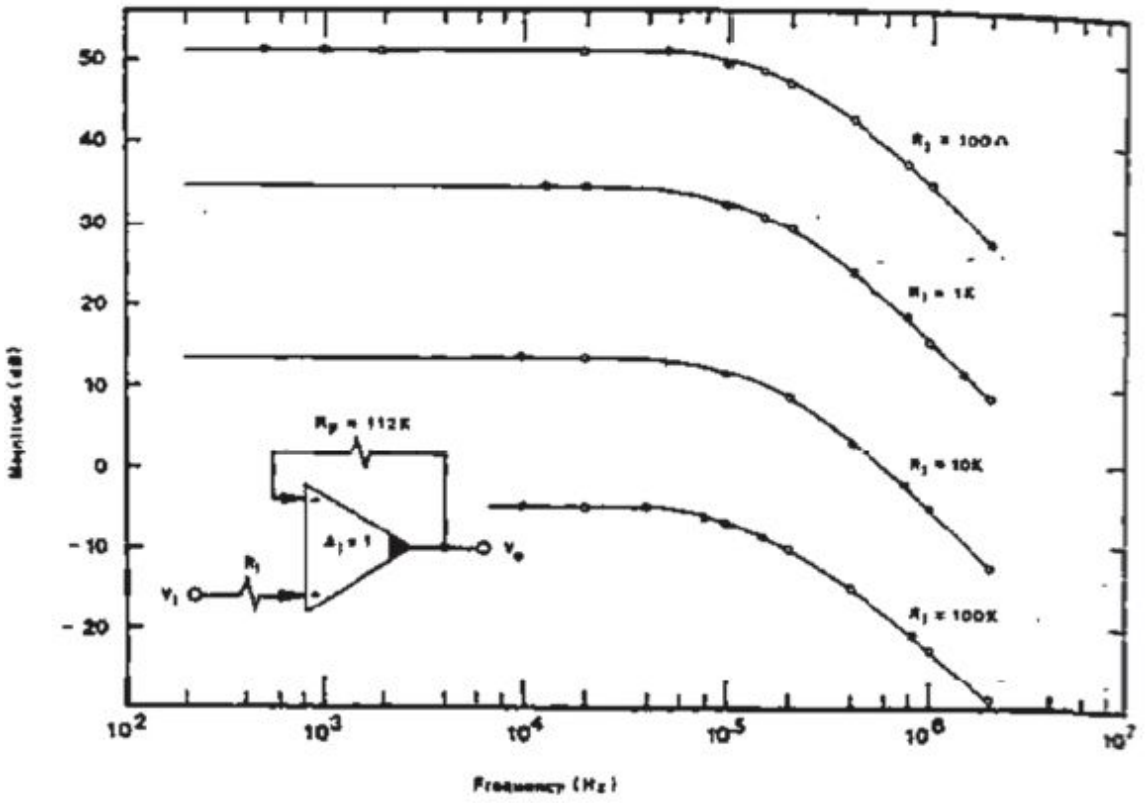
Fig. 4.62: "Norton" amplifier.

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: $\frac{R_F}{R_1} \cdot \frac{Ai}{1+Ai} = \frac{1}{2} \cdot \frac{R_F}{R_1} = \frac{112}{2} = 35dB$ 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.



Voltage amplifier frequency response using a unity gain current amplifier demonstrating the gain versus bandwidth independence.

Fig. 4.63: Independence between bandwidth and gain.

4.5. Operational Transconductance Amplifier (OTA)

4.5.1 Generalities

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. $g_m=I_E/V_{th}$.

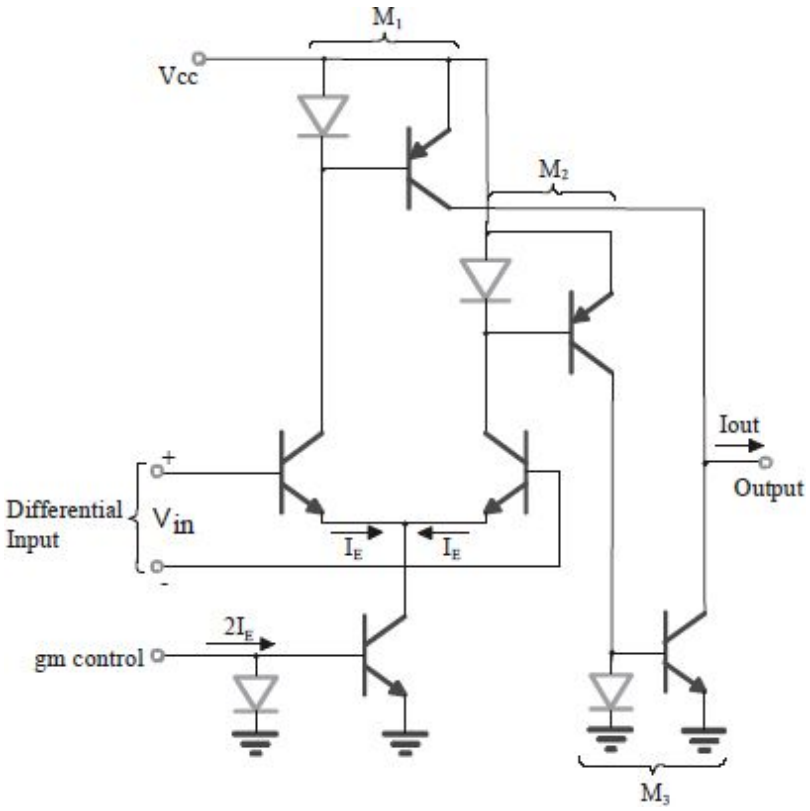


Fig. 4.64: Simplified OTA internal schematics.

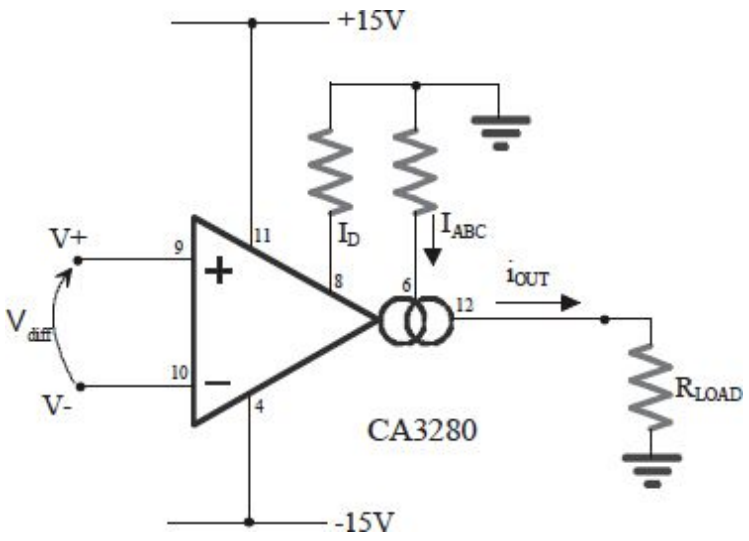


Fig. 4.65: Example of external connection of an OTA.

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).

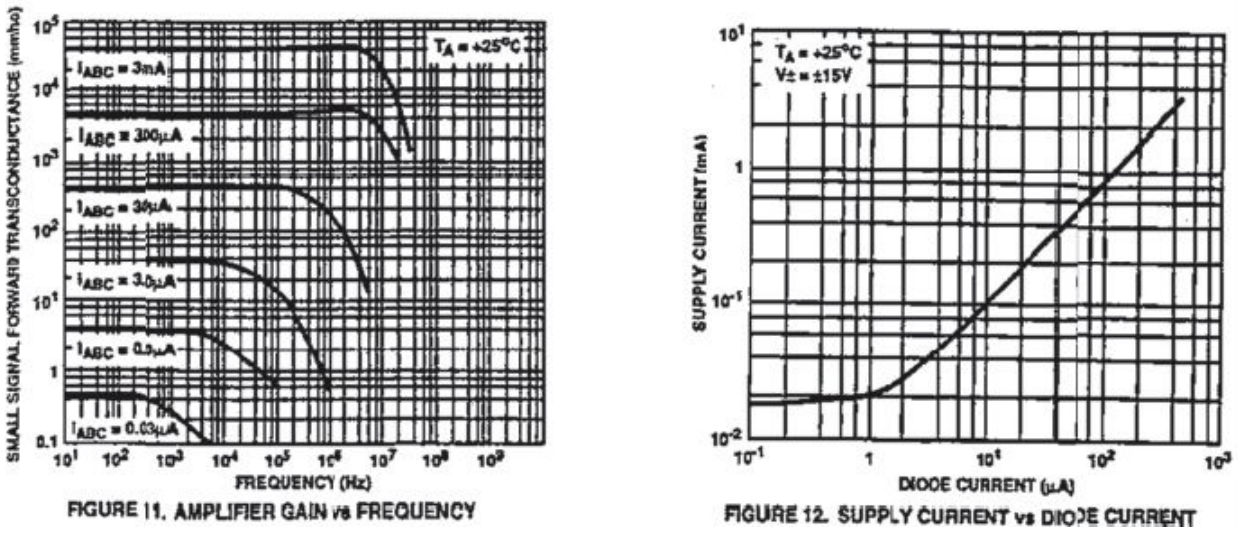


Fig. 4.66: Curves for various parameters of the OTA varying the current I_{ABC}.

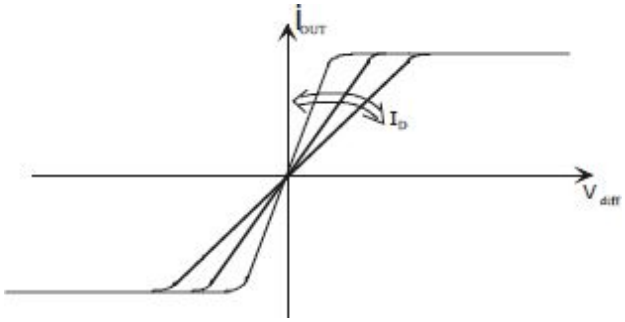


Fig. 4.67: Input-output characteristic.

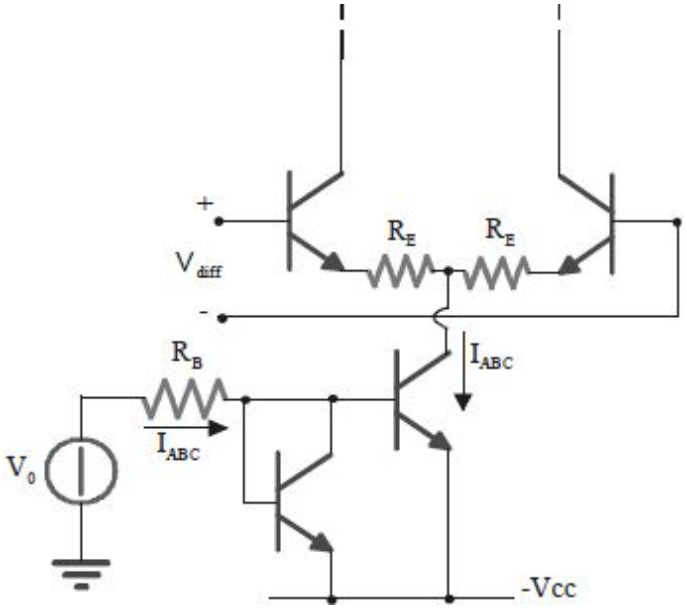
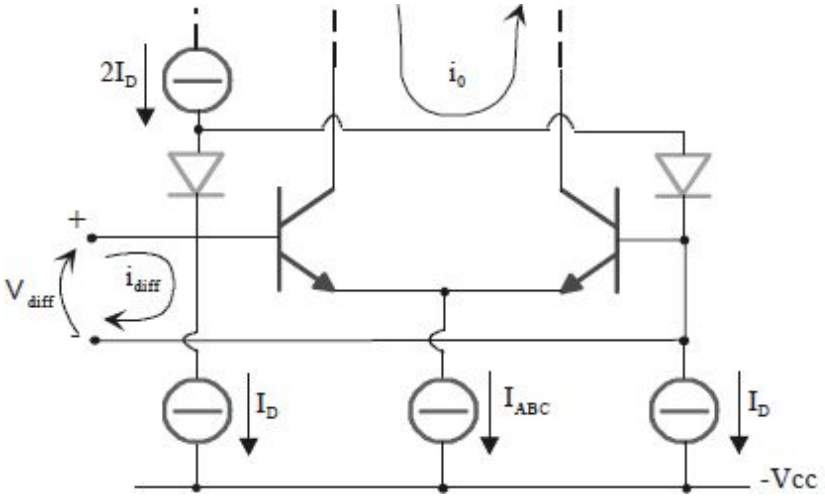


Fig. 4.68: Improving the linearity with emitter degeneration.



$$i_0 = \frac{I_{ABC}}{I_D} \cdot i_{diff}$$

Fig. 4.69: Linearity improvement with input diodes.

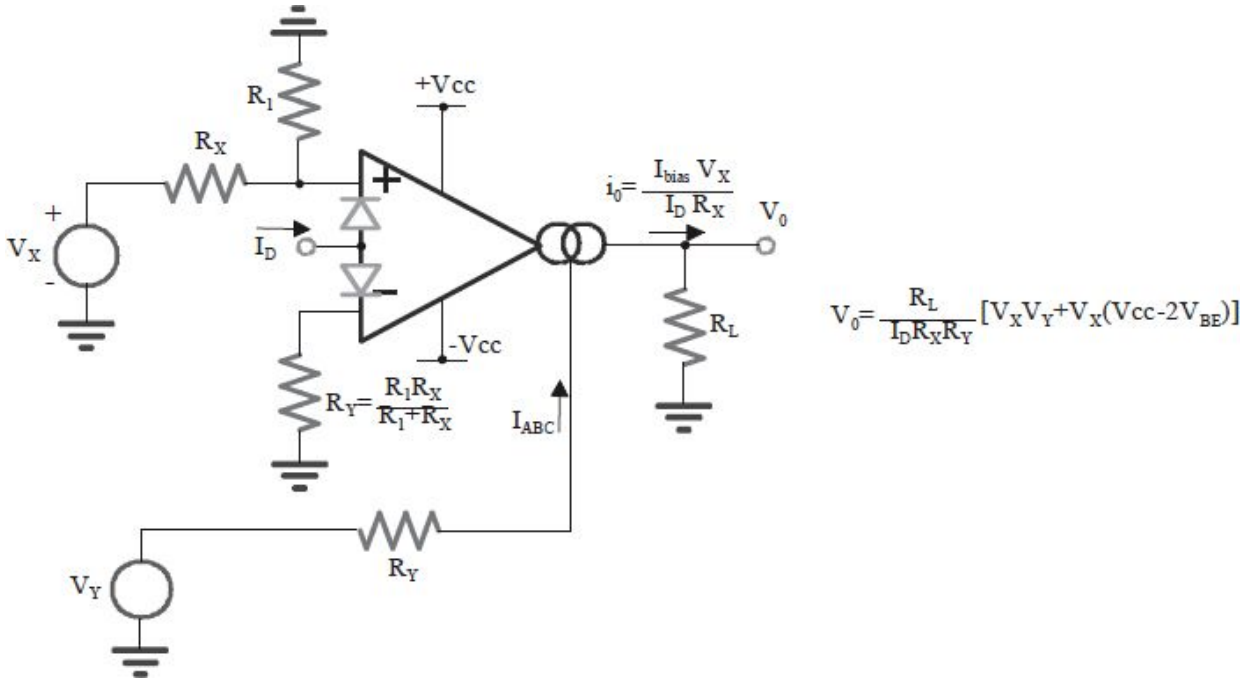


Fig. 4.70: Analog voltage multiplier, obtained by means of one OTA.

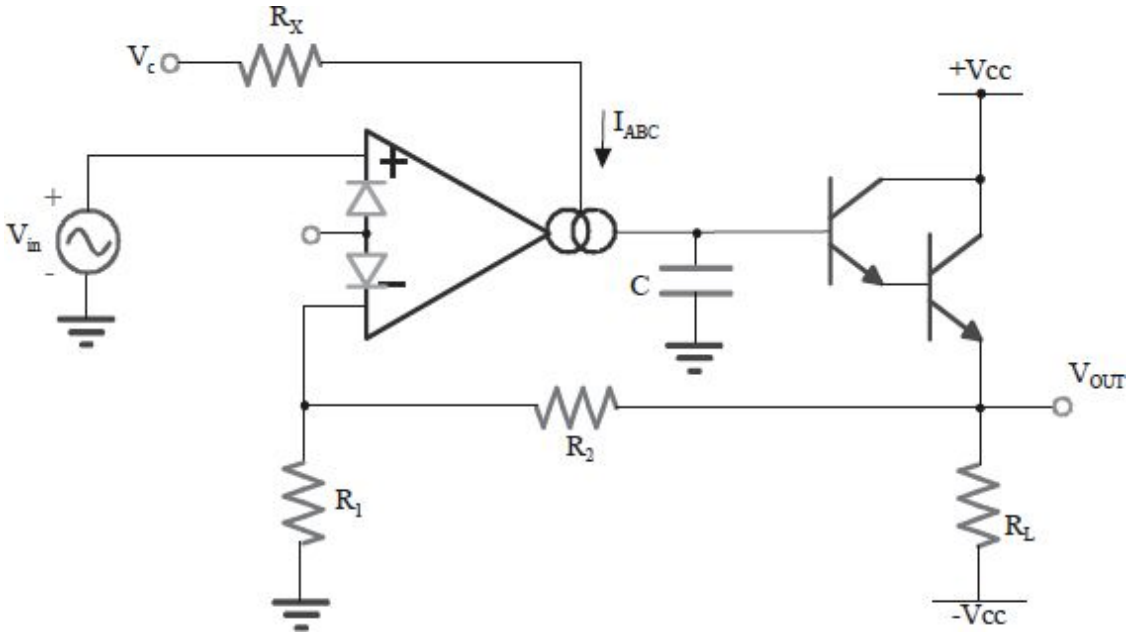


Fig. 4.71: Voltage controlled low-pass filter.

4.5.2 Applications

We can design various circuits, using OTAs. For example, it is possible to use the current bias pin of the OTA to realize an analog multiplier. Observe the Fig. 4.70. For a Wilson current source, we would have: $I_{ABC} = \frac{V_Y + V_{cc} - 2V_{BE}}{R_Y}$

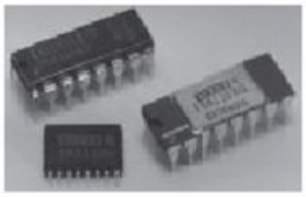
which will provide an output voltage given by the expression in Fig. 4.70, on the right. Translating V_Y, one can cancel the troublesome offset which is a function of V_X (see in the expression of V_0 the term $V_X \cdot (Vcc-2 \cdot V_{BE})$).

Another circuit that one can design with an OTA is an economic voltage-controlled low-pass filter, depicted in Fig. 4.71, in which the –3dB cut-off frequency is: $\omega_0 = \frac{gm \cdot R_1}{(R_1 + R_2) \cdot C} = \frac{q \cdot I_{ABC} \cdot R_1}{2kT \cdot (R_1 + R_2) \cdot C}$

4.6. DATA-SHEETS

In the followings, some commercial advanced OpAmps' data-sheets are reported.

BURR-BROWN®



INA110

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

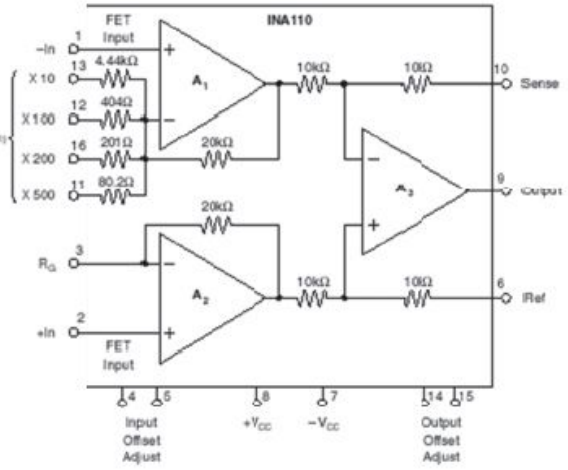
- MULTIPLEXED INPUT DATA ACQUISITION SYSTEM
- FAST DIFFERENTIAL PULSE AMPLIFIER
- HIGH SPEED GAIN BLOCK
- AMPLIFICATION OF HIGH IMPEDANCE SOURCES

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 $\pm V_{CC}$. 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.



NOTE: (1) Connect to R_G for desired gain.

International Airport Industrial Park • Mailing Address: PO Box 11400 • Tucson, AZ 85734 • Street Address: 6730 S. Tucson Blvd. • Tucson, AZ 85706
Tel: (520) 746-1111 • Twx: 910-952-1111 • Cable: BBRCORP • Telex: 066-6491 • FAX: (520) 889-1510 • Immediate Product Info: (800) 548-6132

PDS-645E

Printed in U.S.A. September, 1993

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

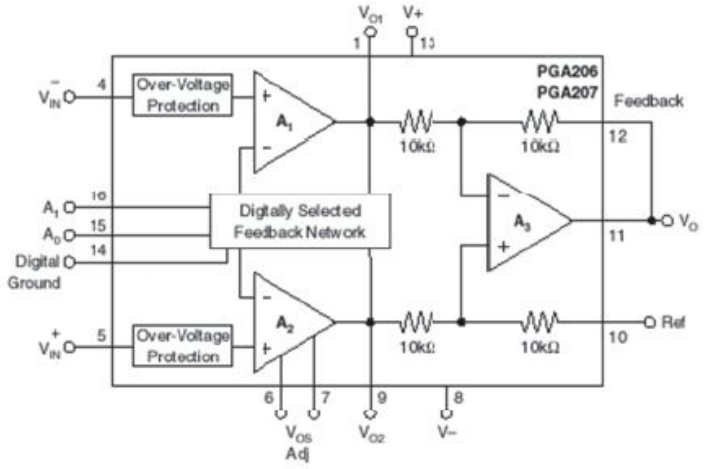
DESCRIPTION

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.



International Airport Industrial Park • Mailing Address: PO Box 11400, Tucson, AZ 85734 • Street Address: 6730 S. Tucson Blvd., Tucson, AZ 85706 • Tel: (520) 746-1111 • Twx: 910-952-1111
Internet: http://www.burr-brown.com/ • FAXLine: (800) 548-6133 (US/Canada Only) • Cable: BBRCORP • Telex: 066-6491 • FAX: (520) 889-1510 • Immediate Product Info: (800) 548-6132

PDS-1241B

Printed in U.S.A. May, 1995

BURR-BROWN®

ISO120
ISO121

Precision Low Cost ISOLATION AMPLIFIER

FEATURES

- 100% TESTED FOR PARTIAL DISCHARGE
- ISO120: Rated 1500Vrms
- ISO121: Rated 3500Vrms
- HIGH IMR: 115dB at 60Hz
- USER CONTROL OF CARRIER FREQUENCY
- LOW NONLINEARITY: ±0.01% max
- BIPOLAR OPERATION: V_O = ±10V
- 0.3"-WIDE 24-PIN HERMETIC DIP, ISO120
- SYNCHRONIZATION CAPABILITY
- WIDE TEMP RANGE: –55°C to +125°C (ISO120)

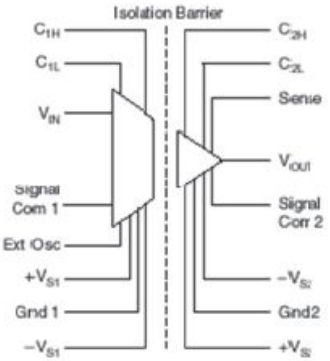
APPLICATIONS

- INDUSTRIAL PROCESS CONTROL: Transducer Isolator for Thermocouples, RTDs, Pressure Bridges, and Flow Meters, 4mA to20mA Loop Isolation
- GROUND LOOP ELIMINATION
- MOTOR AND SCR CONTROL
- POWER MONITORING
- ANALYTICAL MEASUREMENTS
- BIOMEDICAL MEASUREMENTS
- DATA ACQUISITION
- TEST EQUIPMENT

DESCRIPTION

The ISO120 and ISO121 are precision isolation amplifiers incorporating a novel duty cycle modulation-demodulation technique. The signal is transmitted digitally across a 2pF differential capacitive barrier. With digital modulation the barrier characteristics do not affect signal integrity, which results in excellent reliability and good high frequency transient immunity across the barrier. Both the amplifier and barrier capacitors are housed in a hermetic DIP. The ISO120 and ISO121 differ only in package size and isolation voltage rating.

These amplifiers are easy to use. No external components are required for 60kHz bandwidth. With the addition of two external capacitors, precision specifications of 0.01% max nonlinearity and 150µV/°C max V_{OS} drift are guaranteed with 6kHz bandwidth. A power supply range of ±4.5V to ±18V and low quiescent current make these amplifiers ideal for a wide range of applications.



International Airport Industrial Park • Mailing Address: PO Box 11400 • Tucson, AZ 85734 • Street Address: 6730 S. Tucson Blvd. • Tucson, AZ 85706
Tel: (520) 746-1111 • Twx: 910-952-1111 • Cable: BBRCORP • Telex: 066-6491 • FAX: (520) 889-1510 • Immediate Product Info: (800) 548-6132

PDS-820D
Printed in U.S.A. March, 1992

INITIAL RELEASE
Final Electrical Specifications

LT1815

6.5mA, 220MHz, 1500V/µs Operational Amplifier with Programmable Current

January 2001

FEATURES

- **220MHz Gain-Bandwidth Product**
- **1500V/µs Slew Rate**
- **7mA Maximum Supply Current**
- **Space Saving SOT-23 Packages**
- **Shutdown or Programmable Current Option**
- **6nV/√Hz Input Noise Voltage**
- 450MHz –3dB Bandwidth ($A_V = 1$)
- Unity-Gain Stable with C_{LOAD} Up to 100pF
- 1.5mV Maximum Input Offset Voltage
- 8µA Maximum Input Bias Current
- 800nA Maximum Input Offset Current
- 50mA Minimum Output Current, $V_{OUT} = \pm 3V$
- ±3.5V Minimum Input CMR, $V_S = \pm 5V$
- Specified at ±5V, Single 5V Supplies
- Operating Temperature Range: –40°C to 85°C

APPLICATIONS

- Wideband Amplifiers
- Buffers
- Active Filters
- Video and RF Amplification
- Communication Receivers
- Cable Drivers
- Data Acquisition Systems

DESCRIPTION

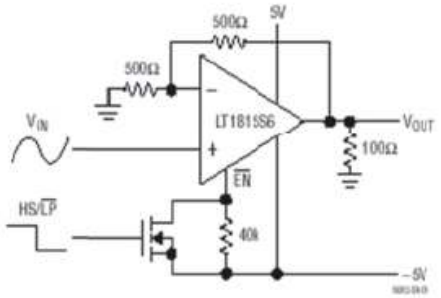
The LT®1815 is a low power, high speed, very high slew rate operational amplifier with excellent DC performance. The LT1815 features higher bandwidth and slew rate, much lower input offset voltage and lower noise than devices with comparable supply current. A programmable current option (LT1815S6) allows power savings and flexibility by operating at reduced supply current and speed as well as a complete shutdown reducing supply current to 150µA. The circuit topology is a voltage feedback amplifier with the slewing characteristics of a current feedback amplifier.

The output drives a 100Ω load to ±3.8V with ±5V supplies. On a single 5V supply, the output swings from 1V to 4V with a 100Ω load connected to 2.5V. The amplifier is stable with a 100pF capacitive load, which makes it useful in buffer and cable driver applications.

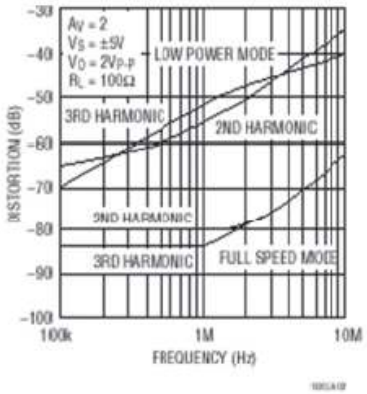
The LT1815 is manufactured on Linear Technology's advanced low voltage complementary bipolar process and is available in space saving 5-lead and 6-lead SOT23 packages, as well as in an SO-8.

TYPICAL APPLICATION

Programmable Current Amplifier Switches from Low Power Mode to Full Speed Mode



Distortion vs Frequency



LT1395/LT1396/LT1397

Single/Dual/Quad 400MHz Current Feedback Amplifier

FEATURES

- **400MHz Bandwidth on ±5V (A_V = 1)**
- **350MHz Bandwidth on ±5V (A_V = 2, –1)**
- **0.1dB Gain Flatness: 100MHz (A_V = 1, 2 and –1)**
- High Slew Rate: 800V/µs
- Wide Supply Range: ±2V(4V) to ±6V(12V)
- 80mA Output Current
- Low Supply Current: 4.6mA/Amplifier
- LT1395: SO-8, SOT23-5 and SOT23-6 Packages
  LT1396: SO-8 and MSOP Packages
  LT1397: SO-14 and SSOP-16 Packages

APPLICATIONS

- Cable Drivers
- Video Amplifiers
- MUX Amplifiers
- High Speed Portable Equipment
- IF Amplifiers

DESCRIPTION

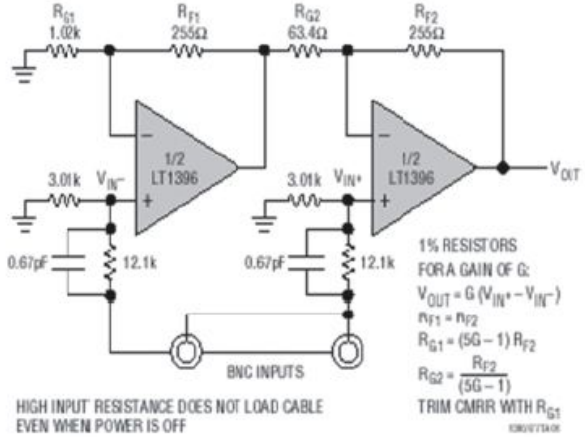
The LT®1395/LT1396/LT1397 are single/dual/quad 400MHz current feedback amplifiers with an 800V/µs slew rate and the ability to drive up to 80mA of output current.

The LT1395/LT1396/LT1397 operate on all supplies from a single 4V to ±6V. At ±5V, they draw 4.6mA of supply current per amplifier. The LT1395CS6 also adds a shutdown pin. When disabled, the LT1395CS6 draws virtually zero supply current and its output becomes high impedance. The LT1395CS6 will turn on in only 30ns and turn off in 40ns, making it ideal in spread spectrum and portable equipment applications.

The LT1395/LT1396/LT1397 are manufactured on Linear Technology's proprietary complementary bipolar process. They have standard single/dual/quad pinouts and they are optimized for use on supply voltages of ±5V.

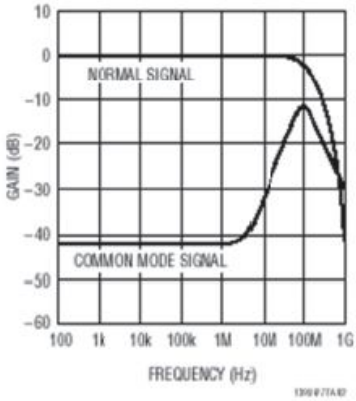
TYPICAL APPLICATION

Unity-Gain Video Loop-Through Amplifier



HIGH INPUT RESISTANCE DOES NOT LOAD CABLE EVEN WHEN POWER IS OFF

Loop-Through Amplifier Frequency Response



Data Sheet | September 2000 | File Number 1174.5

Dual, 9MHz, Operational Transconductance Amplifier (OTA)

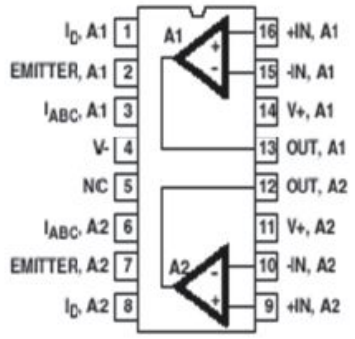
The CA3280 and CA3280A types consist of two variable operational amplifiers that are designed to substantially reduce the initial input offset voltage and the offset voltage variation with respect to changes in programming current. This design results in reduced "AGC thump," an objectionable characteristic 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.

Pinout

CA3280
(PDIP, CERDIP)
TOP VIEW



Features

- Low Initial Input Offset Voltage: 500 V (Max) (CA3280A)
- Low Offset Voltage Change vs I_{ABC}: <500 V (Typ) forAll Types
- Low Offset Voltage Drift: 5 V/°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

- Votage Controlled Amplifiers
- Votage Controlled Oscillators
- Multipliers
- Demodulators
- Sample and Hold
- Instrumentation Amplifiers
- Function Generators
- Triangle Wave-to-Sine Wave Converters
- Comparators
- Audio Preamplifier

Ordering Information

PART NUMBER	TEMP. RANGE (°C)	PACKAGE	PKG. NO.
CA3280AE	-55 to 125	16 Ld PDIP	E16.3
CA3280E	0 to 70	16 Ld PDIP	E16.3
CA3280AF3	-55 to 125	16 Ld CERDIP	F16.3

5

Exercises on OpAmps

“Here I am
Just like I said I would be
I’m your friend
Just like you think it should be
Did you think I would stand here and lie
As our moment was passing us by
Oh I am here
Waiting for your change of heart
It just takes a beat
To turn it around
Yes I’m waiting for your change of heart
At the edge of my seat
Please turn it around”

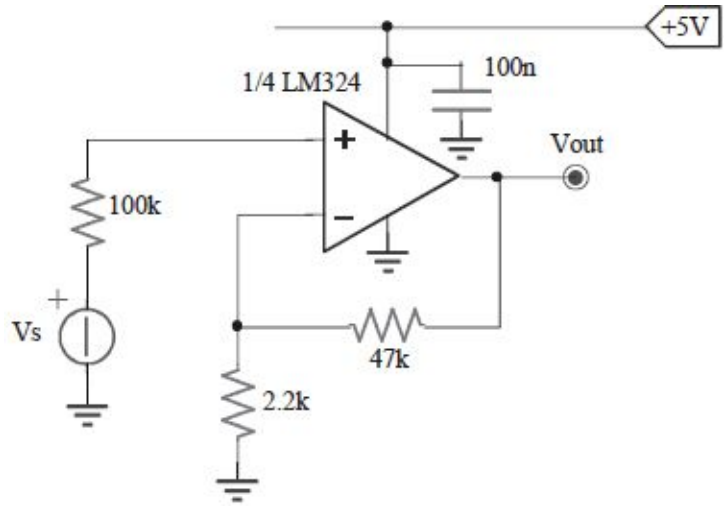
“Change of Heart”, Cindy Lauper

5.1.

The operational amplifier has *IB=50nA*, *IOS=10nA*, *VOS=5mV*, pnp input stage, *A0=100V/mV, PSRR=60dB*; the weight sensor provides *vS=450µV/g*.

a) Introduce a resistance (*2.2kΩ*, *47kΩ* or *100kΩ*) to reduce the effect due to *IB* and calculate the errors due to *IOS* e input *VOS*, expressed in "*grams*".

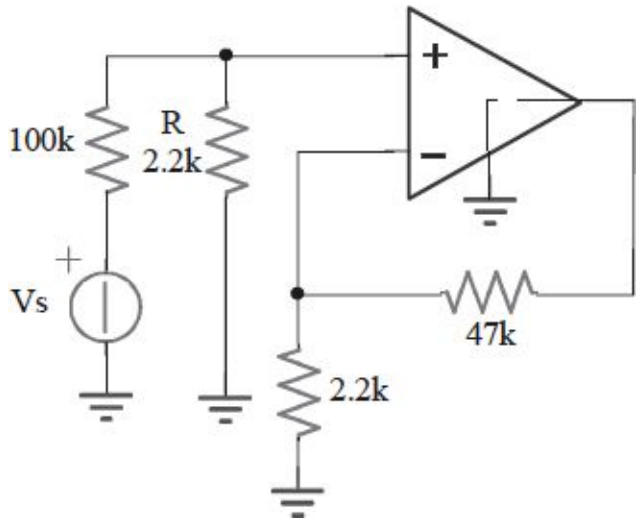
b) Calculate the supply fluctuation that causes an error of ±*1g*.



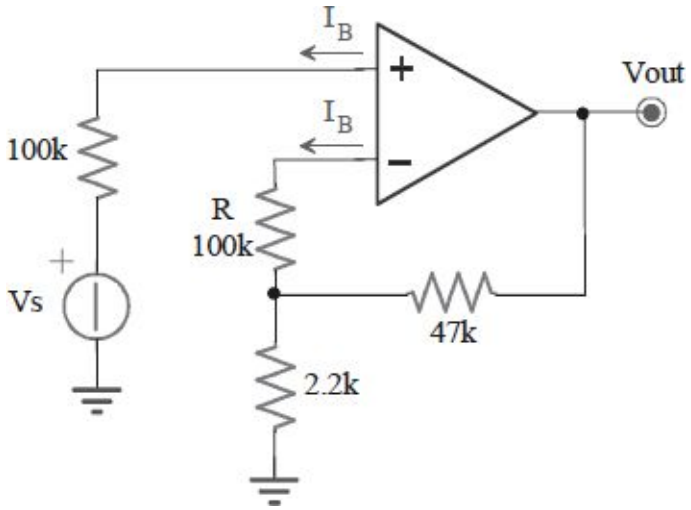
5.1.a

The stage has a voltage gain $A_V=(1+47k/2.2k)=+22.4$. The output sensitivity is, therefore, equal to $v_s \cdot A_V=10mV/g$.

The input differential stage of the OpAmp is a PNP one. Consequently, the bias current will be outgoing, as one can verify consulting the LM324 data-sheet. To balance the effect of this mean current I_B, the 2 input terminals must see the same equivalent resistance to ground. The input resistance of $100k\Omega$ is connected to the non inverting terminal, while the parallel between $2.2k//47k=2.1k\Omega$ is connected to the inverting one. To balance the two branches one needs to increase the value of the resistance seen from the inverting terminal, assuming the reduction of the signal generator resistance to be impossible, because it is intrinsically linked to the source itself. It is meaningless to try to reduce it by adding a resistance R in parallel to the input terminal, to have $100k//R=2.1k\Omega$, that is to say $R=2145\Omega\approx2.2k\Omega$, as shown in the figure.



In fact, the input signal would be reduced by a factor $2.2k/(100k+2.2k)=0.02$, that would strongly modify the functionality of the whole stage. The correct solution is shown in the following. Between the proposed values for the resistances, $2.2k\Omega$, $47k\Omega$ and $100k\Omega$, one should better choose the $100k\Omega$.



Because of the imperfect matching of the resistances seen from the two input terminals, an output offset $V_{out,IB}$ will still be present because of the I_B:

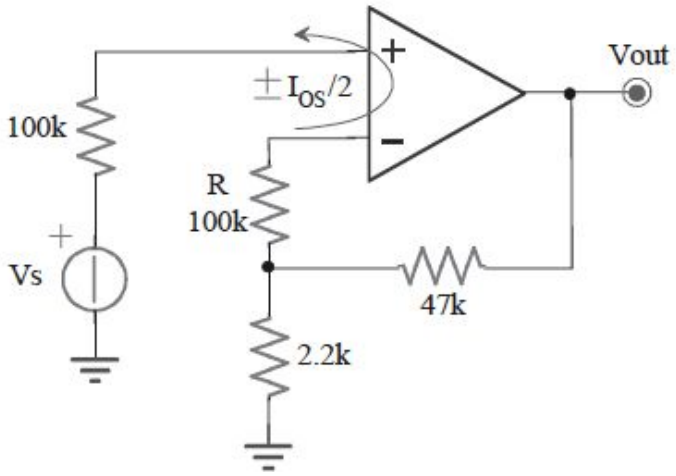
$$V_{out,IB} = I_B \cdot 100k \cdot \left(1+\frac{47k}{2.2k}\right) - \left[\left(I_B + I_B \cdot \frac{100k}{2.2k}\right) \cdot 47k + I_B \cdot 100k\right] = -I_B \cdot 47k = -2.35mV$$

That will determine a reading error equal to *-2.35mV/10mV/g=-0.235g*.

The input bias currents of the OpAmp are not perfectly equal. Just for this reason, in addition to the mean value I_B, the Manufacturer provides the offset current $I_{OS}=|I_{B+}-I_{B-}|$ too. So it is also necessary to consider this additional contribute, that can be schematize as an equivalent generator to the value of $I_{OS}/2$, shown in the figure. The effect is equal to:

$$V_{out,Ios} = {I_{OS}}/{2} \cdot 100k \cdot \left(1 + \frac{47k}{2.2k}\right) + \left[\left({I_{OS}}/{2} + {I_{OS}}/{2} \cdot \frac{100k}{2.2k}\right) \cdot 47k + {I_{OS}}/{2} \cdot 100k\right] =$$

$$= {I_{OS}}/{2} \cdot \left(2 \cdot 100k + 2 \cdot 100k \cdot \frac{47k}{2.2k} + 47k\right) = +22.6mV$$



That corresponds to an error of *±2.26g*. It's worth remembering that the offset I_{OS} has unknown direction a priori, unlike the mean value I_B. So it's important to keep the sign unset (by indicate it with "±") to avoid roughly nullifying it with other output offset contributes.

The offset voltage of the OpAmp can be directly compared with the input signal: V_{OS}=*±5mV* corresponds to an error of *5m/450μV/g=±11.1g*. All considered, this is the most important contribute. As in the previous case this value is known up to the sign. If the given value of V_{OS} is the "maximum" granted in the data-sheet, this means the the value of the output voltage of the OpAmp will have a real value ranging from *–11.1g* and *+11.1g* with respect to the indicated value.

5.1.b

A supply fluctuation is not filtered from the by-pass capacity in the circuit, whose function, on the contrary, is to reduce any possible ripple at high frequency or current spikes. Thus this fluctuation can be transferred to the OpAmp output, through its inside parts.

The OpAmp ability to be immune to the supply fluctuation is measured by the *PowerSupply Rejection Ratio*, *PSRR*, defined as the equivalent voltage $V_{OS,supply}$ that must be applied to the OpAmp input to have in output the same spurious signal $v_{out,supply}$ that the supply fluctuation ΔV_{supply} would cause:

$$PSRR = \frac{V_{OS,\text{supply}}}{\Delta V_{\text{supply}}}$$

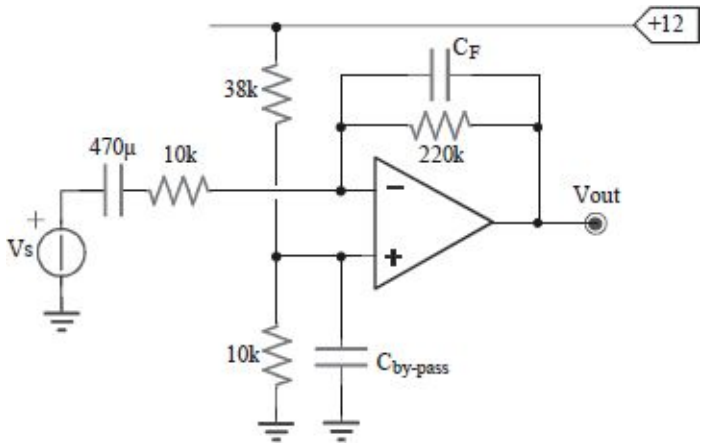
A value *PSRR=60dB* corresponds to a ratio *1/1000*. On the output, a spurious signal corresponding to *±1g* will appear when a value $V_{OS,supply}$*=±450μV* would be present at the input, that is to say when the supply voltage would fluctuate of V_{supply}*=1000·450μ=0.45V*. Such a value is far from being rare; some cheap voltage stabilizers are often granted with a level of precision of *±10%* with respect to the face value, that is of *±0.5V* in front of the face *5V*. The residual ripple at *100Hz* (in the case of a double wave rectification) could provide similar values to the one we have just found, too.

5.2.

The 20Hz sinusoidal signal *vS* has 100mV peak. The OpAmp has *IB=50nA, VOS=5mV, A0=100'000*. Downstream there is a *10bit* ADC with *Vref=5V*.

a) If there is a disturbance on the *12V* power supply, characterized by a peak value of *0.6V* at the frequency of *100Hz*, choose the PSRR in order to reach a level of precision of *±1LSB*.

b) Estimate C*bypass* to have an error of *±1LSB* due to the interference.


+12
C_F
38k
220k
470µ
10k
Vout
Vs
10k
$C_{by\text{-}pass}$

5.2.a

The stage has a voltage gain at medium frequency equal to A_V*=-220k/10k=-22*. At high frequency also the filter capacity C_F behaves like a short-circuit and the stage causing the signal to fade to zero at the output (note that it does not become a buffer). However, for the aim of this analysis, C_F is assumed to be always an open circuit.

For polarization, the output is *12·10k/(38k+10k)=2.5V*; the greatest range of the input signal can move the output voltage of *±100m·(-22)=-(±2.2V)*, that is to say make it vary from *+0.3V* to *4.7V*, by far included in the dynamics of an ADC with a *5V FSR*. In the case a *10bit* ADC is used on the output, one *LSB* corresponds to $5/2^{10}$*=4.9mV*.

Actually, because of the bias current I_B of the OpAmp differential stage, the voltage reached by the output is slightly different. Let's suppose that the bias current is sourced (outgoing, e.g. due to a pnp input stage). Because of the I_B in the uninverting terminal, the output voltage moves of I_B*·10k//38k=+0.4mV*; instead, because of the inverting terminal, the output voltage decrease of a value -I_B*·220k=-11mV*. The overall effect *ΔVout=-11m+0.4m=-10.6mV* seems to be negligible, but in reality it corresponds to more than *2·LSB*. Also the offset voltage V_{OS} can be a cause of static error of, at most, *±5m/4.9m=±1.02·LSB*.

The limited open loop amplification of the OpAmp forces, in order to have on the output a scale range of *5V*, a difference of potential between the input terminals of $5/A_0$*=50μV*. We have not to compare this value with *1·LSB* on the output, equal to *4.9mV*, but with the equivalent input *1·LSB*, before the voltage amplification of *22*. Therefore also this OpAmp non ideality introduces an error of about *¼·LSB*.

The input decoupling capacity have to be chosen in order to determine a pole at a very lower frequency than the signal one, *20Hz*. The value quoted in the electrical scheme determines a time constant of *470μ·10k=4.7s*, equal to a pole at the frequency *34mHz*, more than two decades before the useful frequency. The disturbance on the power supply goes directly into the OpAmp, through its supply pins. Supposing to be able to take an output fluctuation of *±1·LSB*, that is to say of *±4.9mV*, this is equivalent to take an input fluctuation, obtained as

in the case of a voltage offset, of *4.9m/23=213μV*. In fact an interference at *100Hz* is amplified in voltage like the signal at *20Hz* is. The *PowerSupply Rejection Ratio*, *PSRR*, is defined as the equivalent voltage $V_{OS,supply}$ to apply to OpAmp input in order to obtain the same spurious output due to supply fluctuation ΔV_{supply}.

It is derived: $\mathrm{PSRR} = \frac{V_{OS,supply}}{\Delta V_{supply}} = \frac{213\mu V}{0.6V} = -69dB$

A more than reasonable value for many commercial and cheap OpAmps.

5.2.b

The same supply interference can reach also the non-inverting terminal, through the input voltage divider. Without the by-pass capacity, this fluctuation of *0.6V* would reach the input with a peak value of *0.6·10k/(38k+10k)=125mV* and would be amplified to the output by the gain *1+22*. The corresponding spurious fluctuation on the output of *125m·23=2.9V* would even cause the OpAmp exit its correct dynamics, making it saturate alternately to the extreme values of the supply, that is to say between about ground and about *5V*.

It's just tank to the $C_{by\text{-}pass}$ that this fluctuation can be reduced, (but never totally eliminated). If we want a spurious output ripple lower than *±1·LSB*, it will be necessary that the maximum peak value of the ripple on the non-inverting terminal will be *4.9m/23=213μV*. The parallel value $10k\Omega // C_{by\text{-}pass}$ should have, at *100Hz*, a impedance value for which it is valid the following:

$$\Delta V_{in} = \Delta V_{supply} \cdot \frac{Z}{Z + 38k} \approx \Delta V_{supply} \cdot \frac{1/2\pi 100 \cdot C_{by-pass}}{38k}$$

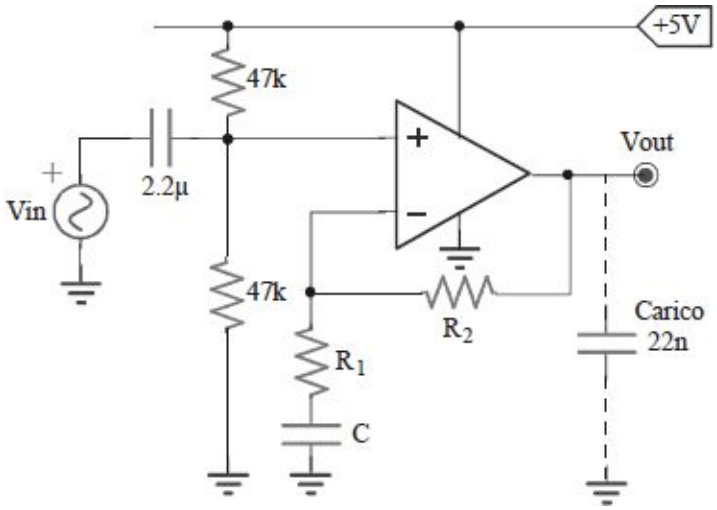
Since *Z* has to be very low, we can ignore the *10kΩ* resistance in parallel to the capacity and neglect the *Z* in series to the *38kΩ*. So doing, we obtain a value of $C_{by\text{-}pass} > 118\mu F$; we can choose a commercial value of *220μF*. Practically, we have realized a low pass filter with a pole at *0.1Hz*, three decades lower than the interference frequency *100Hz*.

5.3.

The OpAmp is rail-to-rail, *Rout=10Ω* and *FPBW=53kHz* with *Vsupply=±15V*. The input is *±500mV* with bandwidth *100Hz÷10kHz*.

a) Choose the values of *R1* and *R2* to exploit the output dynamics and compute the frequency from which the limitation due to *SR* exists.

b) Introducing the *22nF* load, determine the current *Io* that the OpAmp must be able to deliver and the minimum and maximum extremes of the *GBWP*.



5.3.a

This is an amplifier stage in an non-inverting configuration, with a level shifting to *2.5V*. The medium frequency gain, when both the capacitors are shorted, is of $1+R_2/R_1$. To make a sinusoidal input signal with a peak value of *500mV* exploit the entire dynamics available on the output, equal to *2.5V*, it is convenient to choose a gain of *2.5/0.5=5*, that is to say a ratio $R_2/R_1=4$.

Several selection criteria for the values of R_2 and R_1 can exist; for example, balancing the effect of the OpAmp bias current. The non-inverting input sees an equivalent resistance outwards of $47k//47k=23.5k\Omega$. Since the inverting input sees the only R_2, it is possible to choose $R_2\approx 23.5k\Omega$; whose nearest commercial value is $R_2=22k\Omega$. It's now possible to derive also the value of $R_1=22k/4=5.5k\Omega$; we'll choose $R_1=5.6k\Omega$, that provides an effective gain of $1+22k/5.6k=4.92$, as to exploit almost totally the dynamics, without the risk of saturation of the OpAmp. Alternatively, we could have connected two resistances of $2.2k\Omega$ and $3.3k\Omega$ in series, so obtaining a gain of *5*, not considering the tolerance of the resistances.

Given an input sinusoid with a maximum frequency f_{max}, the maximum slope of the signal on the OpAmp input node will be of $\omega \cdot V_p = 2\pi \cdot f_{max} \cdot V_p$. On the OpAmp output, the peak value of the sinusoid is equal to about *2.5V (*in theory). Moreover, to a first approximation, the maximum slope the OpAmp is able to assure is equal to its *SR* that, unfortunately, is not provided. In substitution, however, it is given the quoted value of the full-power bandwidth (*FPBW*), equal to the maximum frequency that, on the output, allows to reach the maximum range towards the specified supply. In this case *FPBW=53kHz* with $V_{aupply}=\pm 15V$. That is to say that the OpAmp is able to follow a sinusoid with a $V_p=15V$ and a frequency *f=53kHz*, corresponding to a maximum slope of $2\pi \cdot 53k \cdot 15 \approx 5V/\mu s$. This is just the limit value of the OpAmp, equal to its *Slew Rate*. Practically the formula is:

$$FPBW = \frac{SR}{2\pi \cdot V_{out,\max}}$$

Since in the case under consideration the OpAmp is polarized with a *5V single-power supply* and the output *swing* is assumed to be *rail-to-rail*, the maximum frequency that can be reached from the stage is higher, equal to:

$$f_{\max} = \frac{SR}{2\pi \cdot V_p} = \frac{5V/\mu s}{2\pi \cdot 2.5V} = 318kHz$$

If sinusoids at higher frequency are applied to the input, the output will be heavily limited from the OpAmp slew rate and therefore will be very distorted. In this case, the input has a bandwidth include in a range from *100Hz* and *10kHz* and consequently it will not be affected by such a phenomenon.

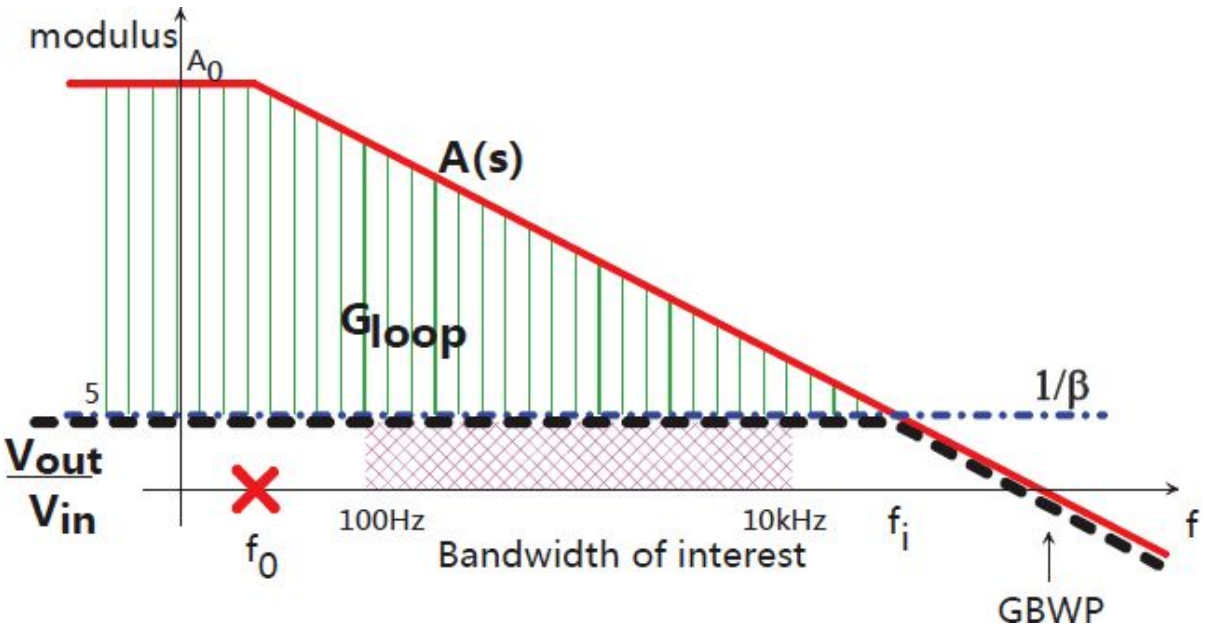
5.3.b

Adding a capacitive load to the output, the OpAmp will have to provide current to it too, besides the feedback resistance R_2. Usually, the contribute for the capacitor is quantitatively higher than that for the resistance. Let's begin to evaluate the last one that, even in the case of maximum output ranging, is equal just to $2.5V/R_2=113\mu A$, for sure negligible.

To be able to operate at the extreme values of the signal dynamics, that is to say at *10kHz*, the output potential must be able to vary with a slope $dV_{out}/dt=2\pi \cdot 10kHz \cdot 2.5V=0.157V/\mu s$. Since a capacitor is present, it will charge at the desired slope only if a sufficient amount of current is guaranteed to it, equal to:

$$i_c = C \cdot \frac{dV_c}{dt}$$

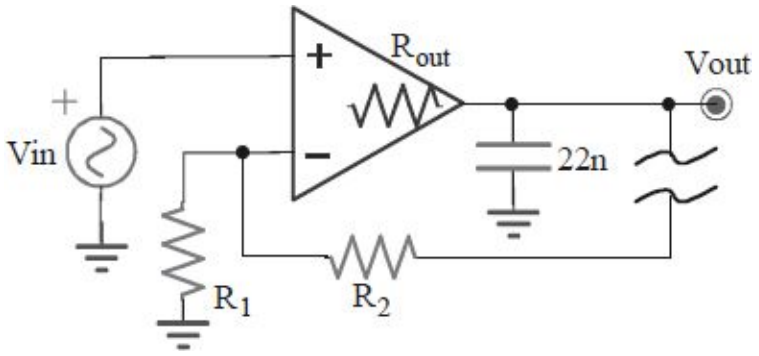
In the case under study, it will be necessary a charge current of $22nF \cdot 0.157V/\mu s=3.45mA$, far higher than that required from the feedback resistance. The OpAmp should therefore be able to deliver a sink and source current of at least $I_O=\pm 3.5mA$.



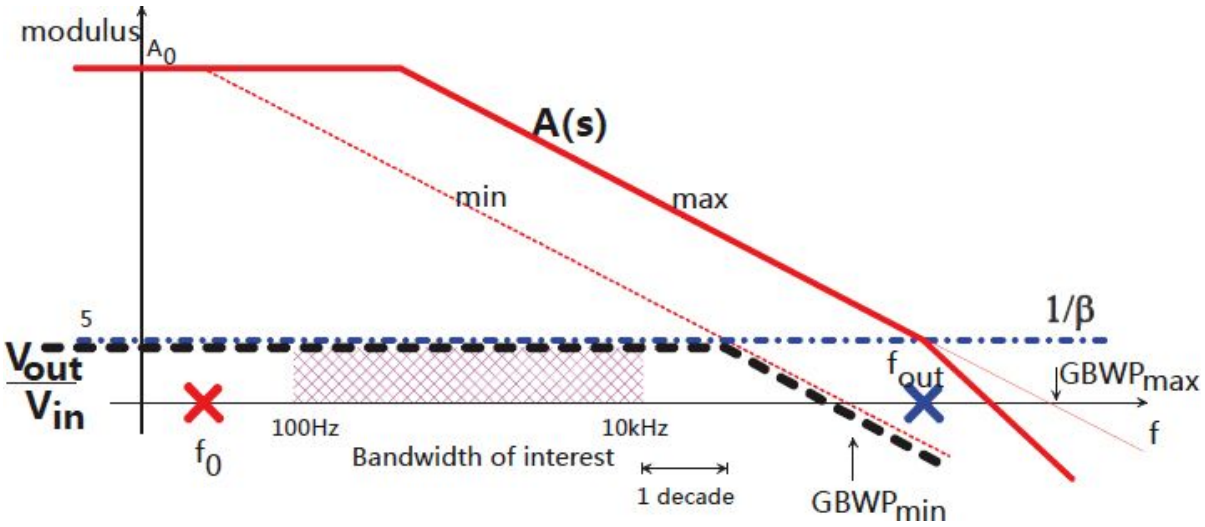
At medium frequency, the stage can be seen as a normal non-inverting configuration. So, the Bode diagram of the modulus of the forward gain of the OpAmp, *A(s)* and of the feedback, *β*, can be shown represented as in the figure.

In order to have a properly operating feedback, it is necessary that also at the maximum interest frequency, *10kHz*, the loop gain is sufficiently high, for example of a value of at least *10*. Since the loop gain is nothing but the vertical distance between the *A(s)* and the *1/β*, this means that it will be necessary to have a closed loop pole at least $f_i=10\cdot10kHz=100kHz$. So we have to choose an OpAmp with a minimum $GBWP_{min}=5\cdot f_i=0.5MHz$.

Actually, the OpAmp will have to drive a capacitive load of *22nF*. Because of the not-nill output resistance $R_{out}=10\Omega$, we'll have a pole in the forward block, with a time constant of $R_{out}\cdot C_{load}=10\cdot22n=220ns$, equal to *724kHz*.



The Bode diagram will be different from the previous one, as shown in the following figure, and a new problem regarding the stability of the amplifier stage could be established. To guarantee in any case a phase margin of *45°*, it is necessary that the intersection between *1/β* and the forward gain (*A(s)* with the addiction of the output pole) occurs in correspondence of the *724kHz* pole. This required a value of $GBWP_{max}=5\cdot724kHz=3.6MHz$. This is a superior limit for the OpAmp performances. On the contrary, if we want to increase the *GBWP* even further (for example to make the feedback better, increasing the gain loop) it is necessary to use a proper external compensation network.

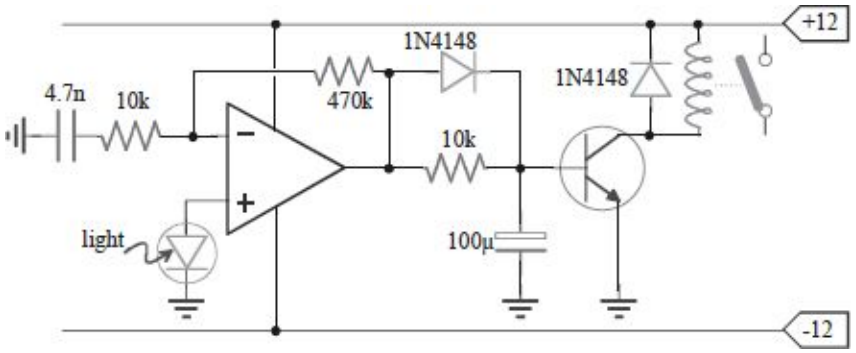

modulus
A0
A(s)
min
max
5
1/β
Vout
Vin
fout
GBWPmax
f
f0
100Hz
10kHz
Bandwidth of interest
1 decade
GBWPmin

5.4.

Given the circuit in the figure, that uses a photodiode:

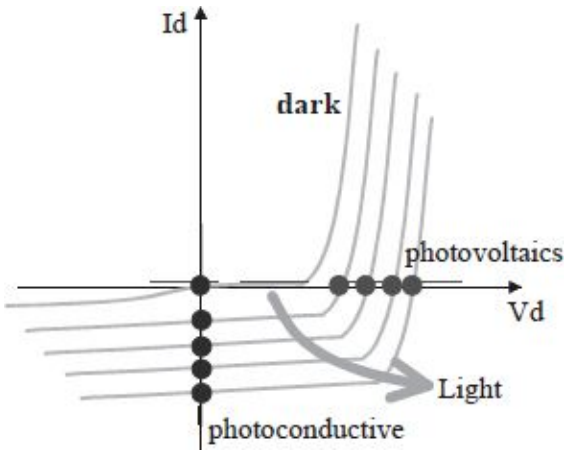
a) explain its functioning, drawing the wave forms.

b) provide the selection criteria for the OpAmp (type of input stage, *GBWP*, $I_{out,max}$, *SR*).

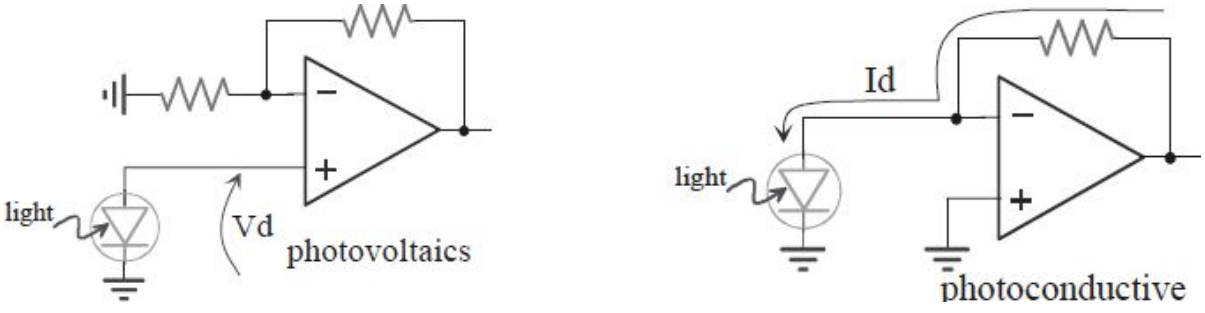


5.4.a

The voltage-current characteristic of a photodiode that is illuminated by light is different from the face static one, as shown in the following figure. The little translation of the characteristic determines an increase of the reverse current and of the direct voltage.

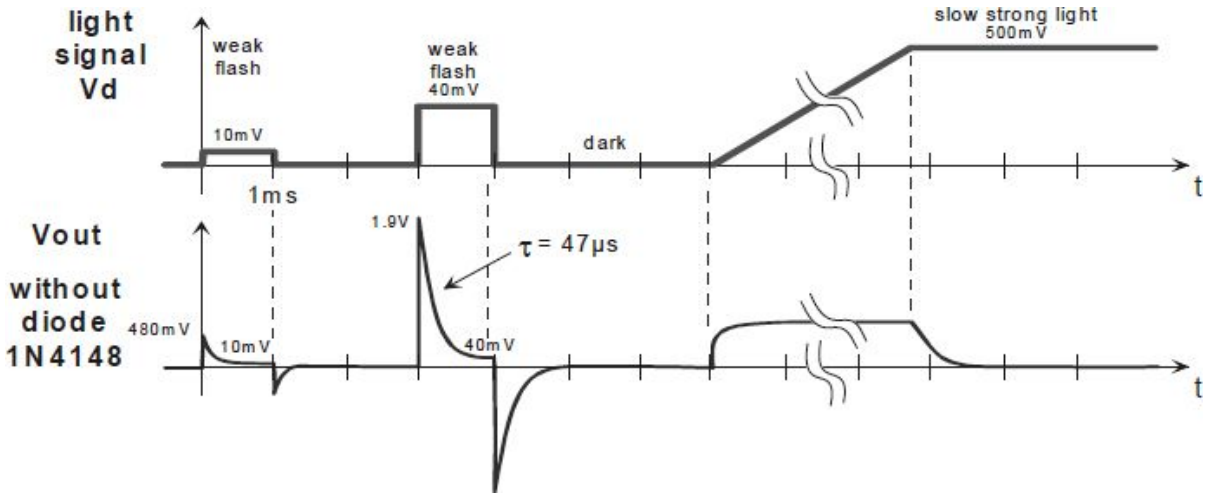


According to the chosen polarization used to drive the diode, two main working regimes are possible: photovoltaic, when any current can flow though the diode; photoconductive, when across the diode a voltage equal to zero is imposed. Despite a direct polarization of the diode through a resistance is possible, the two previous functioning modes are far away the most used, in the circuit configurations shown in the following figure.



In particular, the circuit under consideration collects the voltage across the photodiode in a photovoltaic mode. Moreover, the OpAmp is in a real (i.e. not ideal, hence approximated) differentiator configuration, tank to the *4.7nF* capacitor that in DC regime (f=0Hz) forced a gain of the stage equal to 1, while

for the quick fronts (ideally of infinite frequency) the gain is equal to *1+470k/10k=48*. The time constant of the stage is *4.7n·10k=47µs*, corresponding to a closed loop pole at the frequency of *3'388Hz*.

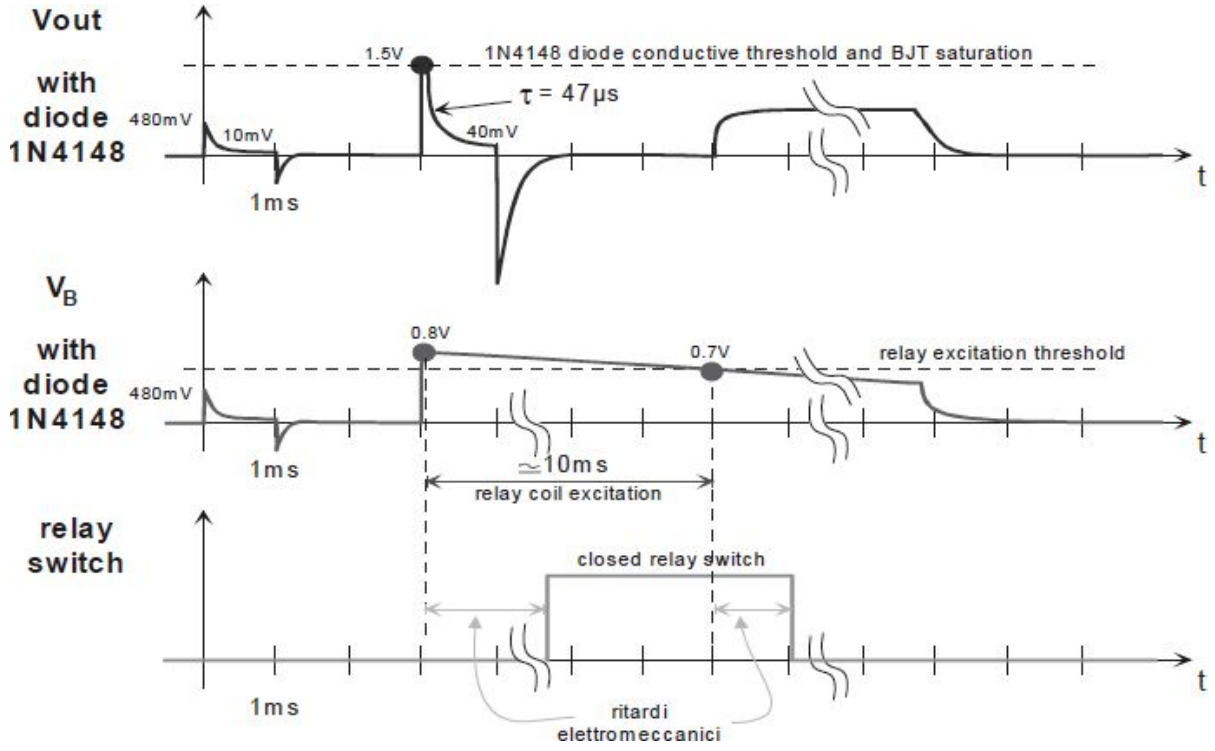


Assuming to have on the photodiode a light signal of *on/off* type, of variable amplitude, it is possible to draw the waveforms here reported. For convenience, we consider the only amplifier stage with OpAmp, without the resting output circuit (practically disconnecting the 1N4148 diode). As you can see, the pulse on the OpAmp output will be as much higher as higher the voltage signal slope (the derivative) of the photodiode is. In particular, in the case of step with amplitude V_d, the output voltage peak will be $V_d \cdot 48$, then decreasing with a time constant of *47µs* towards the steady-state value V_d. Instead, in the case of linear ramp (for example due to a light spot that progressively enlightens the active area of the photodiode) the output signal will be a step with amplitude proportional to the ramp slope.

Monitoring the OpAmp output to see when the pulse becomes higher than a certain threshold, it is possible to know when the optical input signal has been particularly rapid and intense. In this way, thanks to the high-pass action of the stage, one could trigger the downstream electronics when on the input a light signal not intense and constant is present (for example the environmental light on day-time), but a weak and rapid one (for example the flash of a lamp or the rapid reflection of a moving object in front of the photodiode).

This is just the functioning mode we want to obtain from the circuit under consideration. In fact the OpAmp output is connected to a transistor that, in turn, drives a relay. To make the BJT go rapidly into saturation a diode has

been inserted, that by-pass the *10kΩ* base resistance. In this way the threshold has been fixed, too, equal to $V_{BE,sat}+V_D \approx 0.8+0.7=1.5V$. Input signals with rapid fronts, of amplitude higher than *1.5/48=31mV*, will make the BJT saturate. The wave form on the OpAmp output will be different from that previously shown, because now the two junctions in series limit the positive ranging of the OpAmp to about *1.5V*. Practically, the current flowing into the 1N4148 diode will be limited from the OpAmp output source capacity, because there is no limiting resistance. It would have been better to put in series to the diode a resistance of a value of some tens of Ohms, in order not to make the OpAmp output stage suffer. The pulse on the OpAmp output lasts only some time constants. This duration (*47μs*) is certain to be insufficient to make the power contacts of the relay shorted, because of the inertia of its electromechanical control that needs an excitation of the coil that lasts at least some tens of milliseconds. This is why a *100μF* capacitor has been inserted that, in turn, once charged through the diode, will provide the needed base current to the BJT. Of course, when the pulse on the OpAmp output is going to finish, the diode will turn into reverse mode and won't absorb current from the capacitor anymore: in practice this is a peak stretcher. The *10kΩ* resistance is practically pointless and can be eliminated.



Supposing that the relay needs a current of about *100mA* and the BJT has a gain $\beta \approx 100$, the needed base current would be $I_B = 100m/100 = 1mA$. Assuming that the BJT keeps working in saturation when passing from $V_{BE,sat} = 0.8V$ to *0.7V*, the partial discharge of the capacitor will last $\Delta t \approx C \cdot \Delta V / I_B = 100\mu \cdot 0.1/1m = 10ms$, sufficient to correctly excite the relay. Despite that, as shown in the figure above, the closing of the power switch and its next opening, at the end of the excitation command, are delayed because of the electromechanical inertia of the moving parts inside the relay. To avoid these problems one could have replaced the relay with a more rapid and reliable power MOS or TRIAC, properly isolated from the OpAmp, by an opto-coupler.

5.4.b

Once the circuit functioning is understood, it is necessary to properly choose the components in order to assure the correct execution of the several functions. In this case, the most critical component is the OpAmp itself.

Since the input stage has the non-inverting terminal connected to the photodiode, it is necessary that this terminal doesn't absorb current. Therefore it is opportune to choose an OpAmp with a FET input stage (both MOSFET or JFET). In alternative, one can use a BJT input stage, but only of *PNP* type, to have a bias current outgoing from the input pins: in this case, the bias current will end up by polarizing the photodiode slightly in direct mode that, however, doesn't affect so much the photovoltaic functioning mode of the photodiode.

The output stage must be able to deliver the base current to the BJT, equal to about *1mA* because of what already exposed in the previous section, and, above all, the charge current for the *100μF* capacitor. Since the time constant of the OpAmp output discharge is of *47μs*, it is convenient that the capacitor is charged in a maximum time of some tens of microseconds. For example, if we want to charge it to *0.7V* in a time of *50μs*, it would be necessary to assure a current of $C \cdot \Delta V / \Delta t = 100\mu \cdot 0.7 / 50\mu = 1.4A$, absurd!

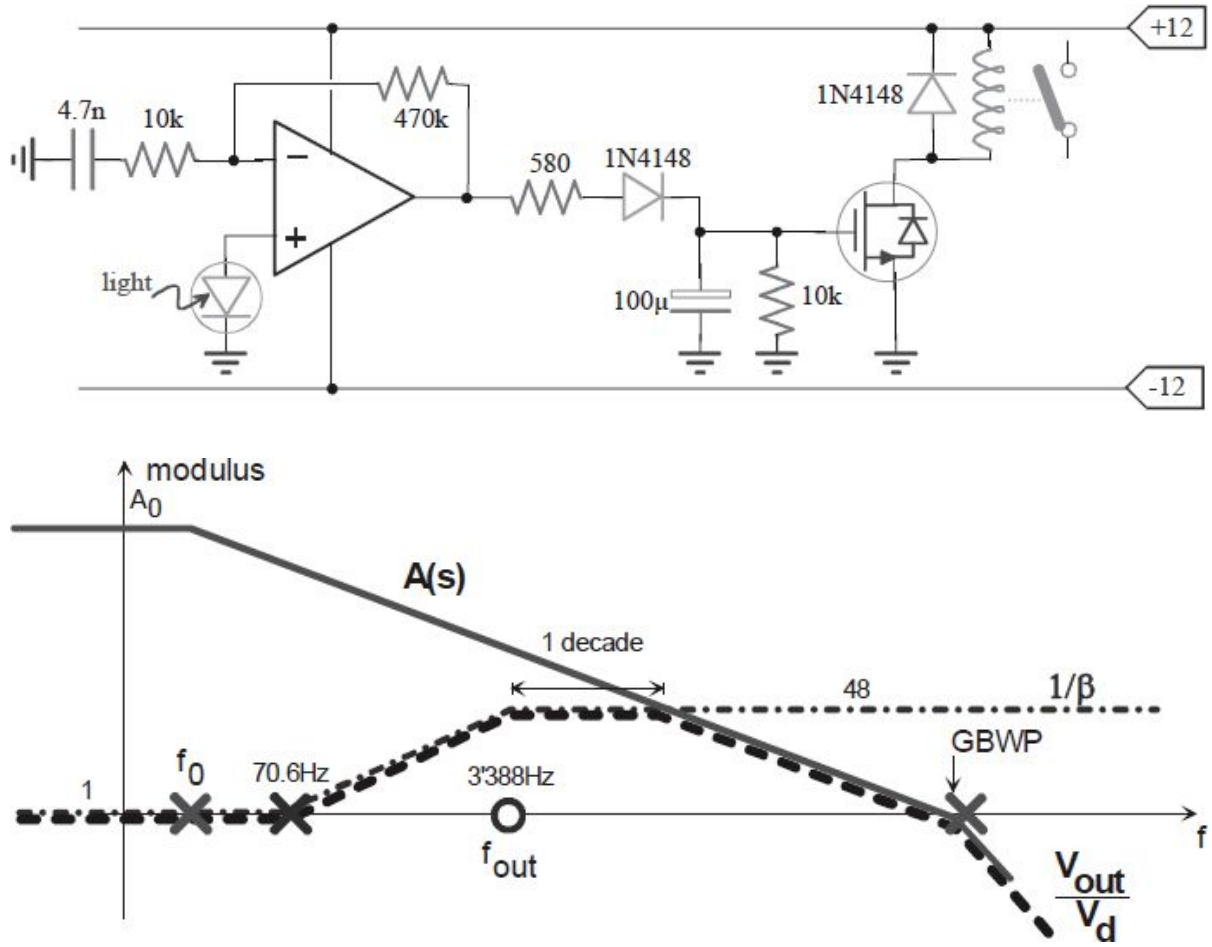
We need to redesign the circuit. The simpler solution is to reduce the value of the capacitor, but in this way it would be necessary to replace the BJT with a MOS that doesn't require any gate current, except during the switching on transient. If we want to limit the charge current to *20mA*, even in the case of maximum OpAmp output reaching *12V*, it is useful to insert a limiting resistance of about $12/20m = 600\Omega$, and we'll choose the nearest commercial value of *580Ω*. Now, the OpAmp output voltage won't saturate to *1.5V* anymore, but it will be able to reach the value imposed from the input. To make the capacitor discharge in about *1s* (and not in *10ms*, as in the previous case, with BJT) it is necessary that the voltage across it goes down the threshold voltage of the MOS in this time: we need to introduce a discharge resistance of about $R \cdot C \approx 1s$, that is to say $1/100\mu = 10k\Omega$. The modified circuit is reported in the following.

The OpAmp will swing its output at the requested speed, but always with the limitation of the available Slew-Rate. If we want an output swing of several Volts, more rapid than the time constant of the differentiator (for example *6μs*,

with respect to the *τ=47μs*), it is necessary to choose an OpAmp with *Slew Rate* of at least *12V/6μs=2V/μs*, an easily available value on the market.

Another consideration has to be done, regarding the OpAmp stability. Since this is a feedback configuration, the stage could show instability or, at least, overshooting responses, absolutely to avoid. To quickly analyze the phase margin of the circuit it is convenient to plot the open loop gain *A(s)* of the OpAmp and the inverse of the attenuation of the feedback block *1/β(s)*.

Assuming a compensated OpAmp, *A(s)* will show a single dominant pole pattern from *A(0)* till the *Gain-Bandwidth Product*, *GBWP*. The feedback network *β(s)* will have a pole due to the *4.7nF* capacitor, with a time constant equal to *4.7nF·(10kΩ+470kΩ)=2.3ms*,



at frequency f_{polo}*=70.6Hz*, and a zero at f_{zero}*=1/4.7nF·10kΩ=3'388Hz*. Because *1/β(s)* is not constant, the intersection with *A(s)* could occur with *40dB/dec* slope and, therefore, the circuit could be unstable. To guarantee a phase margin of *90°* it is necessary to place the zero at least one decade before

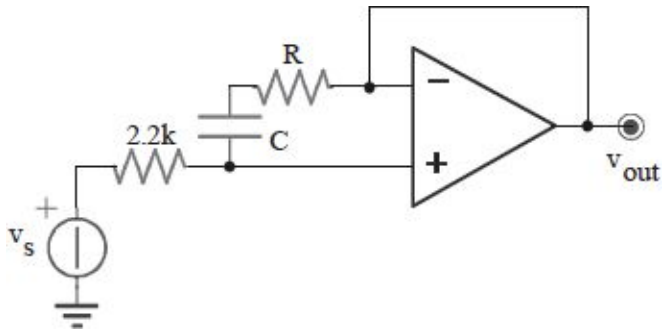
the crossing point. And since the intersection will occur at a gain equal to *48*, it will be necessary to choose an OpAmp with *GBWP 48* times higher than the frequency of the zero. In the end, it is necessary to choose a $GBWP=10 \cdot f_{zero} \cdot 48=10 \cdot f_{polo} \cdot 48 \cdot 48=1.6MHz$, reasonable value for the most of the commercial OpAmps.

5.5.

The high frequency buffer in the figure uses an OpAmp not compensated, with *Amin=10, A0=200'000* and *A=1* at *10MHz*.

a) Choose the values of *R* and *C* to compensate it with a phase margine higher than *45°*.

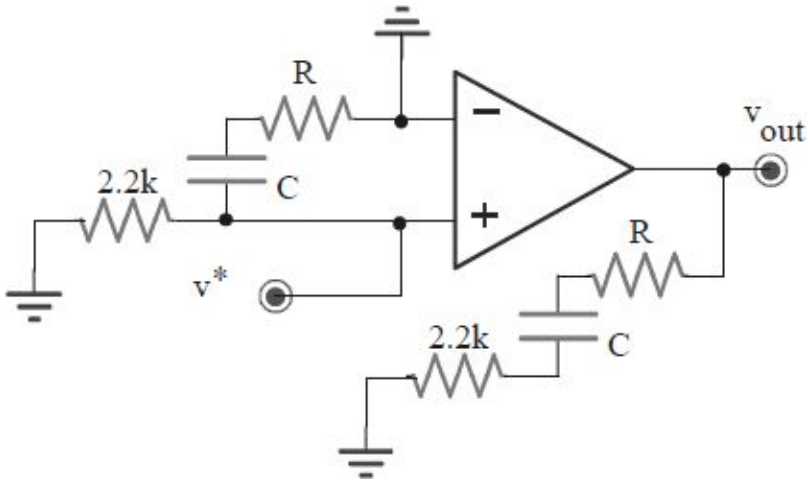
b) Calculate the bandwidth of the buffer and the input impedance at *10Hz*.



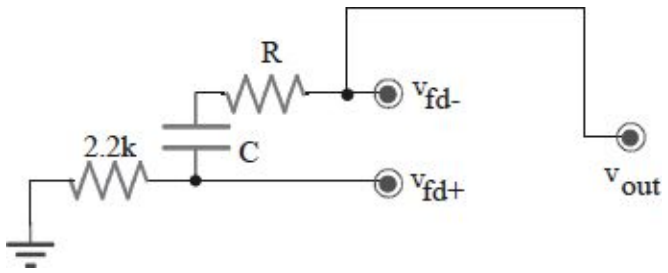
5.5.a

The indicated stage is a voltage buffer. At low frequency the capacitor is an open circuit and the *RC* network is therefore of no influence. At high frequency the capacitor is a short-circuit and the resistance *R* with the input *2.2kΩ* introduce a little positive feedback that, however, don't affect the closed loop gain equal to $A_V=+1$.

Since the OpAmp is not compensated, the stability of the stage is not assured. To study it is useful to investigate the loop gain of the circuit in order to analyze its phase margin when $|G_{loop}|=1$. To do it, it is convenient to break the loop gain $G_{loop}(s)$ into the forward gain of the OpAmp, $A(s)$, and the feedback block $\beta(s)$. The following figure shows the equivalent circuit for the computation of the forward gain. On the output, the closed loop load network has been reconstructed. As you can observe, the gain $v_{out}(s)/v^*(s)$ coincides with the OpAmp gain $A(s)$.



For the feedback block we need to study the circuit here reported. As you can see, the feedback is double: one is negative, equal to $\beta_-=v_{fb-}/v_{out}=+1$; the other is positive $\beta_+(s)=v_{fb+}(s)/v_{out}(s)$, a function of the frequency.



At low frequency the positive feedback is $\beta_+(0)=0$, while at infinite frequency it tends to $\beta_+(\infty)=2.2k/(2.2k+R)$. The overall feedback is equal to:

$$\beta(s)=(\beta_- -\beta_+)=\frac{1+s\cdot R\cdot C}{1+s\cdot(2.2k+R)\cdot C}$$

With a pole and a zero at frequencies:

$$R=\frac{2.2k\Omega}{A_{min}-1}=244\Omega \quad \text{and}$$

$$polo_\beta=\frac{-1}{(2.2k+R)\cdot C} \quad \text{and} \quad zero_\beta=\frac{-1}{R\cdot C}$$

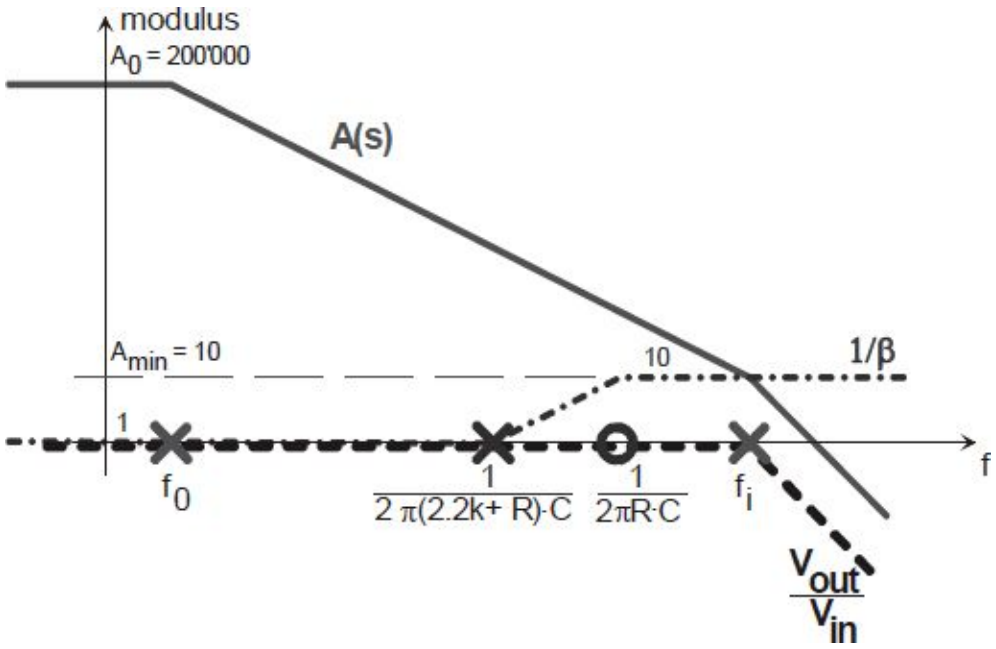
The Bode diagrams of the two gains, forward and feedback, allows to display the pattern of the module of $|G_{loop}(s)|$, equal to the distance between the two patterns. The following figure shows these patterns, together with the total closed loop gain. Without R and C, the intersection between $\beta=1$ and $A(s)$ would occur in the point in which $|A(s)|=1$, with a *–40dB/dec* slope, which would determine instability.

Instead, tank to the pole and the zero introduced by the RC network, the gain $\beta(s)$ is increased. Placing the singularities in order to make the intersection occur just in correspondence of the changing in slope of $A(s)$ from *–20dB/dec* to *–40dB/dec* (that is to say in correspondence of the OpAmp pole), the phase margin will be just of *45°*.

The OpAmp has a $A_{min}=10$ and the intersection with the $A=1$ axes occurs at *10MHz*: it is easily derivable that $f_i=10M/\sqrt{10}=3.16MHz$. The zero of $\beta(s)$ must be sufficiently far from the second pole of the OpAmp (that is also a pole of $G_{loop}(s)$), in order to guarantee the complete recovery of *+90°*, typical of a zero. For this reason it's necessary to place the zero at least a decade before the f_i, that is to say at about $3.16M/10=316kHz$. Since the value of $\beta(s)$ must pass

from *1* (at low frequency) to $A_{min}=10$ (at high frequency), it is necessary that: $\beta(\infty)=\beta_{-}(\infty)-\beta_{+}(\infty)=1-2.2k/(2.2k+R)=1/A_{min}$, from which:

$$R = \frac{2.2k\Omega}{A_{\min} - 1} = 244\Omega$$



Since we have to choose a commercial value, it's better to choose *220Ω* instead of *330Ω*, in order to make the phase margin a little better. At the end, the value of the capacity is derived from the desired value for the zero: $C = \frac{-1}{zero \cdot R} = \frac{1}{2\pi \cdot 316k \cdot 220} = 2.3nF$.

5.5.b

As you can observe from the diagram in the previous page, the closed loop bandwidth of the buffer coincides with the second pole of the OpAmp, that is to say f_i=*3.16MHz*.

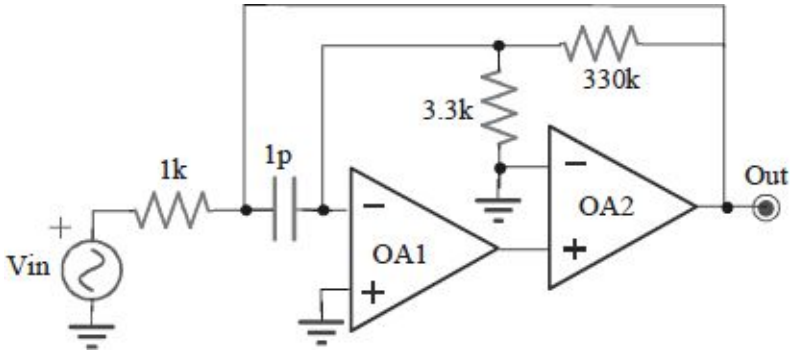
The input impedance would be infinite if there were not the RC network. Yet, also introducing these components between the input terminals of the OpAmp, the series of the resistance and the capacity is affected from the bootstrap effect due to the follower effect of the OpAmp itself: any voltage applied to the non-inverting terminal is transferred to the inverting one; any difference of potential across the two components is present and, therefore, they is no signal current in them. This is more and more right at *10Hz*, when the negative feedback is very high (with $|G_{loop}(10Hz)|$ equal to about *200'000*). In the end, the input impedance is kept extremely high and is practically limited from the input impedance of the non-inverting terminal of the OpAmp.

5.6.

The OpAmp is characterized by *GBWP=5MHz* and R_{out}*=5kΩ*.

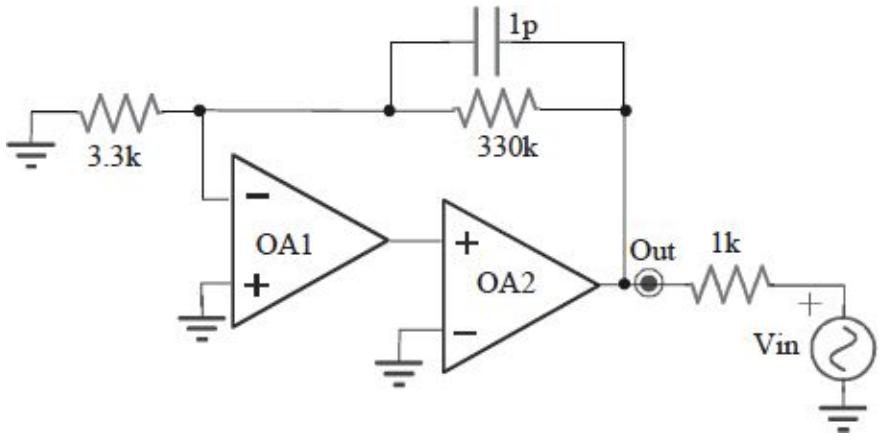
a) Redraw the circuit for ease the analysis and evaluate the stability and the phase margin of the stage.

b) Compute the impedance at *50kHz* experienced by V_{in}.



5.6.a

The circuit becomes more understandable if you notice that the negative terminal of OA2 is held to virtual ground and that the 1pF capacitor is in parallel to the *330kΩ* resistance. Including also the R_{out} the circuit can be redrawn as follows:



To evaluate the stability it is convenient to consider the two OpAmp as connected in series to form one OpAmp with a gain equal to $A^2(s)$, and at this point to compute the simpler β obtained. The output resistance of the OpAmp makes the computation a little harder, this is why it is easier to proceed in a more concise way. At low and at high frequency it results:

$$\beta(0) \cong \frac{1k\Omega}{5k\Omega + 1k\Omega} \cdot \frac{3.3k\Omega}{3.3k\Omega + 330k\Omega} = \frac{1}{606}$$

$$\beta(\infty) \cong \frac{1k\Omega \,//\, 3.3k\Omega}{5k\Omega + 1k\Omega \,//\, 3.3k\Omega} = \frac{1}{7.5}$$

In addition, you can see that the feedback network introduces a zero at the frequency $f_z = 1/2 \cdot \pi \cdot 1pF \cdot 330k\Omega = 482kHz$. Remembering that the gain-bandwidth ratio is constant, the pole turns out to be at the frequency $f_p = 39MHz$. Basically, the Bode diagrams representing $1/\beta$ and $A^2(s)$ intersect with an angle of *40dB/dec*, so it follows that the configuration in exam is unstable. To compute the phase margin, we need to calculate the frequency position of the closed loop pole, that is to say the intersection frequency. Keeping in mind that in the sections with a *-40dB/dec* slope the “squared root

gain-bandwidth" product is constant, we obtain a closed loop pole at *203kHz*. The phase margin is therefore:

$$MF = 180° - 2 \cdot \arctan\left(\frac{f_{CL}}{f_0}\right) + \arctan\left(\frac{f_{CL}}{f_z}\right) - 2 \cdot \arctan\left(\frac{f_{CL}}{f_{GBWP}}\right) - \arctan\left(\frac{f_{CL}}{f_p}\right) =$$

$$\cong 180° - 2 \cdot 90° + 22° - 2 \cdot 2.32° - 0.3° \cong 17°$$

In doing the calculation, we assumed the second pole of the formula, that shows however an unimportant effect, to occur at the *GBWP* and the first pole, on the contrary, to contribute with all its phase shift of *-90°*. Notice that the contribute of the two singularities is double because of the series connection between the two OpAmps.

5.6.b

The impedance evaluated at *50kHz* from the *1kΩ* terminal is equal to the open loop one divided by the G_{loop}, calculated in the new configuration obtained by removing the *1kΩ*. Since it is required to compute the result at *50kHz*, it is convenient to replace the capacity with its impedance at this frequency, equal to *1/2πfC=3.2MΩ*. $G_{loop}(50kHz)$ turns out to be:

$$G_{loop}(50kHz) = A^2(50kHz) \cdot \frac{3.3k\Omega}{3.3k\Omega + 5k\Omega + 300k\Omega} = 100^2 \cdot 0.01 = 100$$

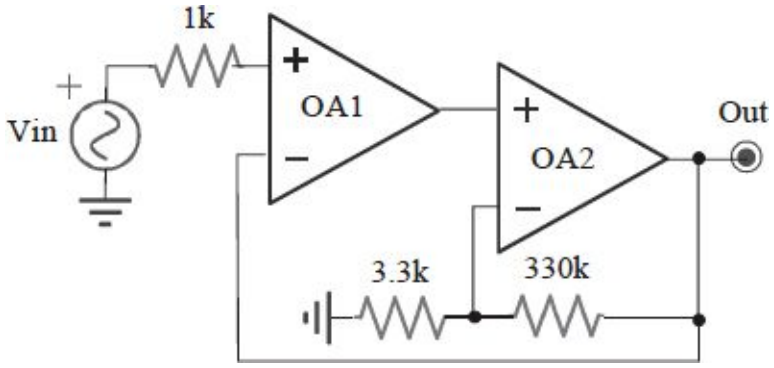
Given the open-loop impedance of about *5kΩ*, the closed-loop impedance will be *5kΩ/100=50Ω*.

5.7.

A_0=100dB, GBWP=50MHz, v_{OS}=5mV, i_B=1μA.

a) Evaluate the stability of the stage and compensate it, if necessary.

b) Determine the total output static error (**no compensation**), and discuss the role of OA2.

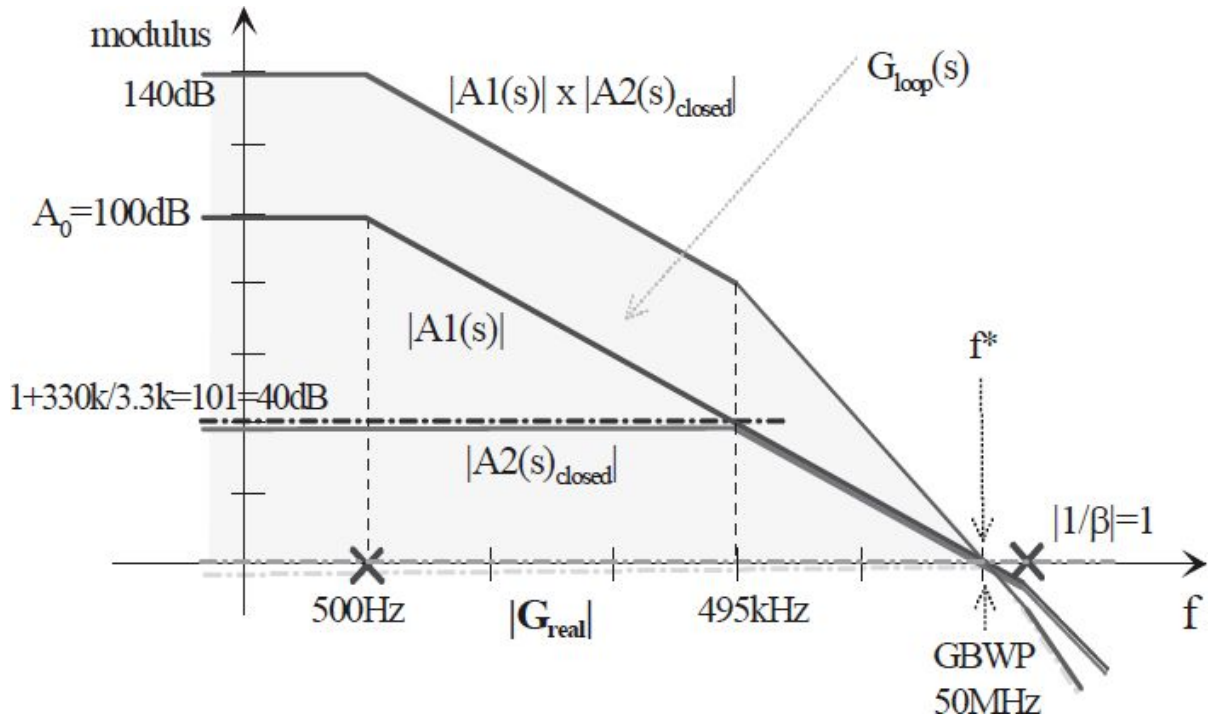


5.7.a

The stage has two OpAmp in a negative feedback configuration. Since OA2 shows its own local feedback, it is possible to examine its stability independently from the rest of the circuit, and then to put its real transfer function in the outer feedback loop.

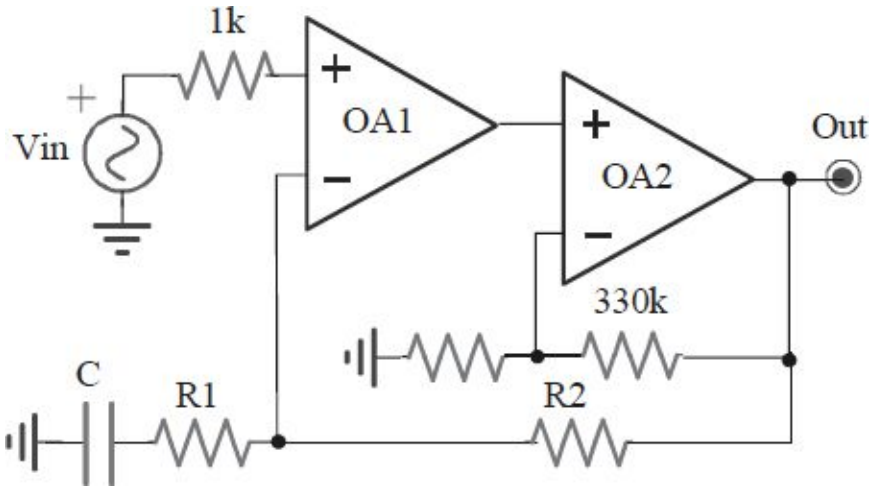
OA2 it's in a well-known non-inverting configuration. We can observe that there are not reactive components and that the OpAmp is è internally compensated, so the stage is stable, for sure. Its transfer function $A2(s)_{close}$ has therefore a gain of 101 from DC till the closed-loop pole, situated at the frequency GBPW/101=495kHz.

Let's now examine the external loop. It could be broke down into a forward block, made up of the OpAmp1 connected in series with the stage just analyzed, and a feedback block β with a gain equal to 1. A(s) is therefore the result of the product $A1(s) \cdot A2(s)_{close}$. The Bode diagrams of $A(s)$and of $1/\beta$ are the following.

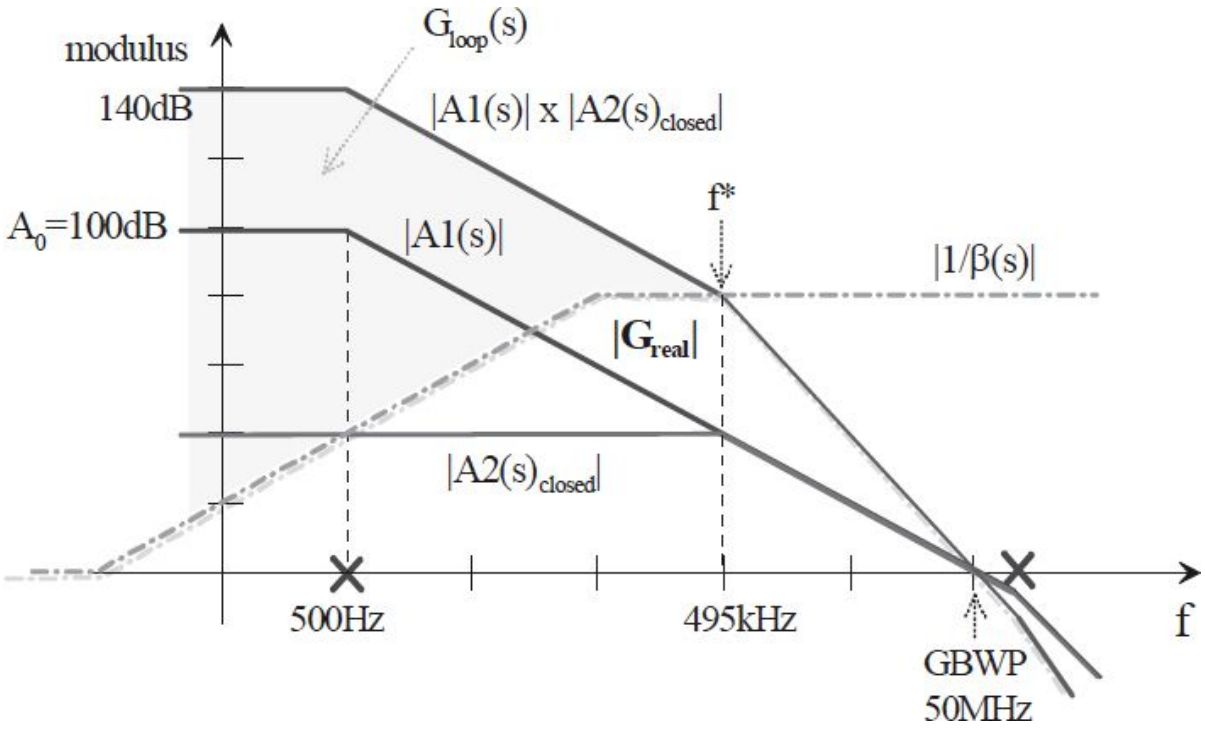


Since the intersection between $A(s)$ and $1/\beta$ presents an angle of 40dB/dec the stage turns out to be unstable.

To compensate the stage it is necessary to increase the value of *1/β* at high frequency so that the intersection will occur with an angle of 20dB/dec. For example it is possible to add a series RC network between the positive terminal of OA1 and ground, while putting in a feedback resistor, as shown in the figure.



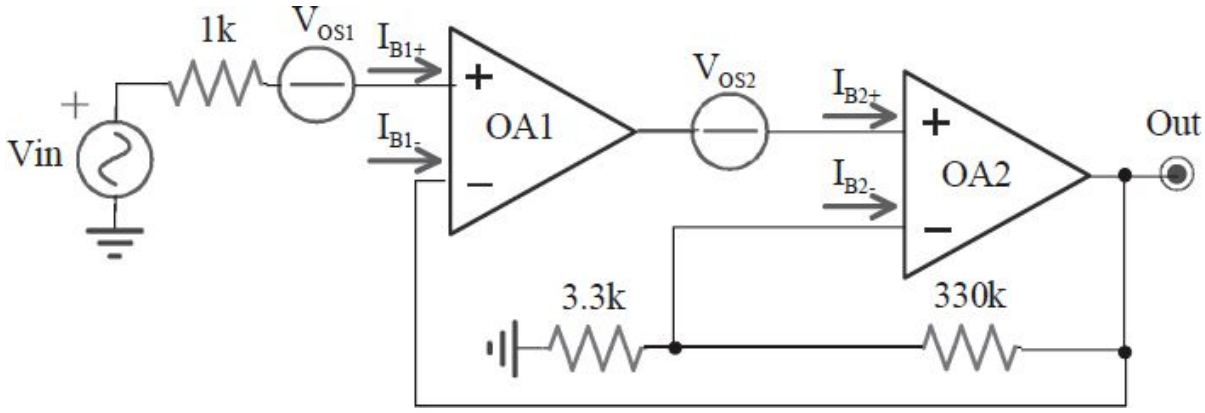
This network introduces in *1/β* a pole at the frequency $f_p = \frac{1}{2\pi C R_1}$ and a zero at the frequency $f_p = \frac{1}{2\pi C(R_1 + R_2)}$, making its high frequency value equal to $(1+R_2/R_1)$. It's quite evident that, in order to provide stability, this value ha sto be higher than 80dB. Moreover, it's appropriate to place the pole of *1/β* at least one decade before the pole of *A(s)*. Setting $\frac{1}{\beta}(\infty) = 10000$ = 10000 and therefore for instance $R_1=1k\Omega$ and $R_2=10k\Omega$, the capacity *C* will consequently have to be chosen higher than *3.2nF*, in order to place the pole at frequencies lower than *49.5kHz*.



However, such a compensation modifies the real gain of the stage, not equal to 1 till the closed loop pole anymore, but just till the zero of *1/β*, that is to say till *495kHz/10/10,000≈5Hz*. This could be not unacceptable for specific applications, for example because it amplifies high frequency disturbances. So, let's find as a useful exercise another compensation configuration able to keep the real gain unchanged.

5.7.b

The stage shows several sources of static errors. They are individuated in the offset voltages of the OpAmps, in their bias currents and in their limited gain. To better evaluate their contribute to the output, let's consider the following scheme:



The offset voltage of the first OpAmp is situated at the input of the stage, so it is transferred to the output as the signal is. The DC gain is 1. The bias current at the positive input terminal of the first OpAmp turns into a voltage across the *1kΩ* resistor, and then it is transferred to the output through the voltage gain of the stage.

The bias current on the negative input terminal of the first OpAmp has no effect on the output, because of the low output impedance of the second OpAmp.

The offset voltage of the second OpAmp is applied to a loop node, so the feedback will nullify its effect on the output. Its contribution is therefore about zero, because the dc G_{loop} is very high.

The bias current on the positive input terminal of the second OpAmp give no contribution to the output, again because of the low output impedance of the first OpAmp.

The bias current on the negative input terminal of the second OpAmp turns into voltage on the *330kΩ* resistor. Yet, in this case too the global feedback will have the tendency to nullify the effect on the output proportionally to the value of $G_{loop}(0)$. The output contribute will therefore be negligible.

The limited gain of the first OpAmp causes the two input terminal not to be exactly at the same potential, because an even little error signal has to exist. This error depends upon the value of V_{out} and its evaluable as $V_{out}/G_{loop}(0)$. So, it's contribute is insignificant. A similar reasoning could be carried on for the error that is introduced by the limited gain of the second OpAmp.

In the end, the total static error on the output is given by:

$$\varepsilon_{out} = V_{os1} \cdot 1 + I_{B1+} \cdot 1k\Omega \cdot 1 + \frac{V_{os2} \cdot 101}{G_{loop}(0)} + \frac{I_{B2-} \cdot 330k\Omega}{G_{loop}(0)} + \frac{V_{out}}{G_{loop}(0)} + \frac{V_{out} \cdot (1 + 330k\Omega / 3.3k\Omega) / A_{O2}}{G_{loop}(0)} \cong$$

$$\cong \pm 5mV \pm 1mV + 0 + 0 + 0 + 0 \cong \pm 6mV$$

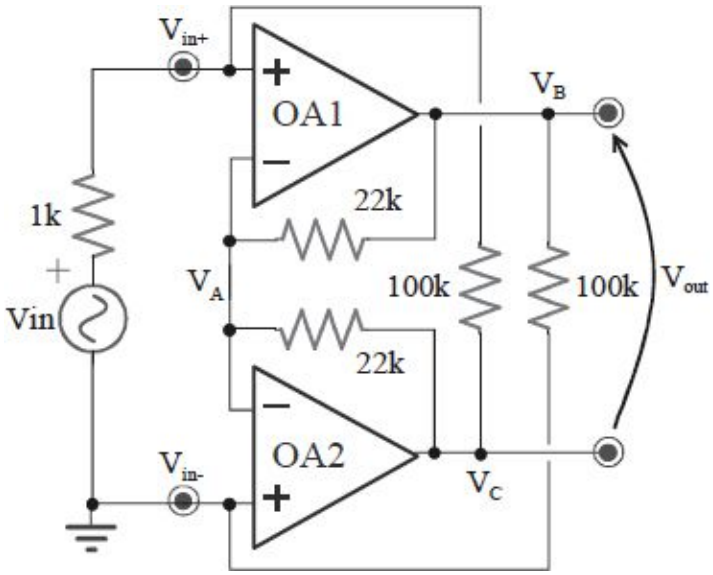
5.8.

The OpAmps are characterized by $I_B=10nA$, $V_{OS}=5mV$, $A_0=100dB$ and $GBWP=30MHz$. Notice the connection of v_{in-} to ground.

a) Determine voltages and currents when $v_{in}=+20mV$ and compute the gain v_{out}/v_{in}.

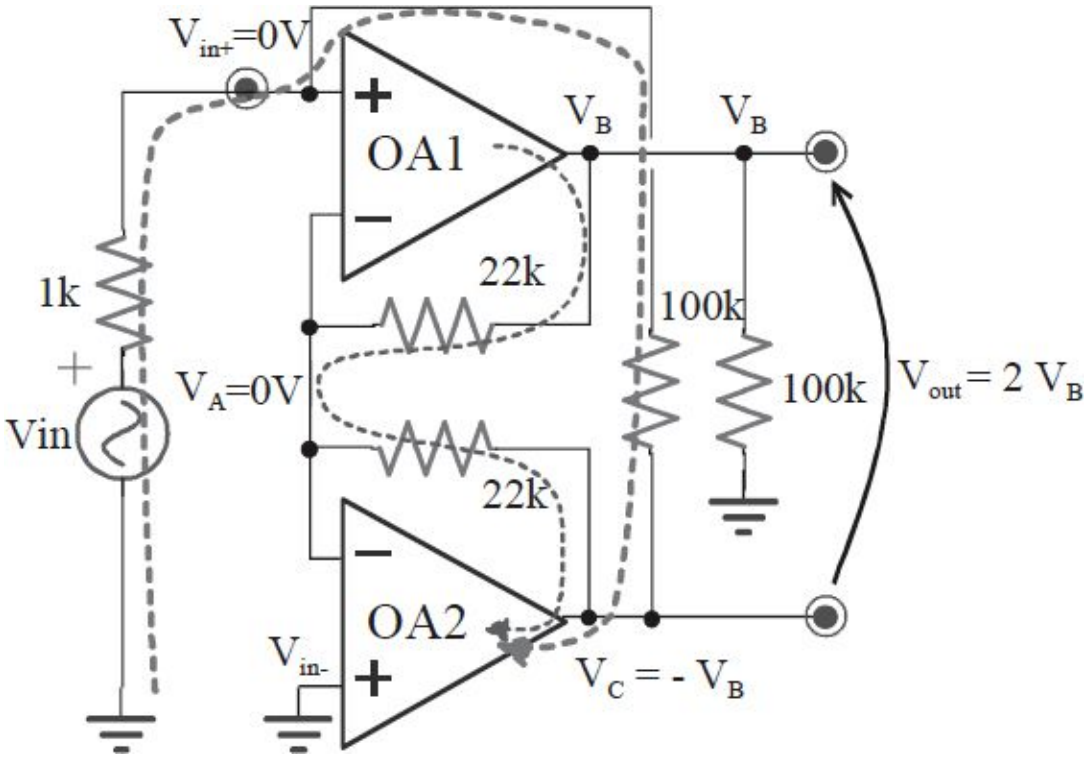
b) Drow the G_{loop} diagram of OA1 assuming OA2 to be perfectly ideal.

c) Determine the static output error v_{out} when $v_{in}=0$.



5.8.a

To easily examine the stage, it can be useful to redraw it as it follows:



Let's now proceed to evaluate the gain. Considering currents and voltages as indicated in the figure, and observing that all the input terminals of the OpAmps are connected to because of the feedback, it's easy to verify that the following relations are valid:

$$I_+ = \frac{V_{IN}}{1k\Omega} = 20\mu A$$

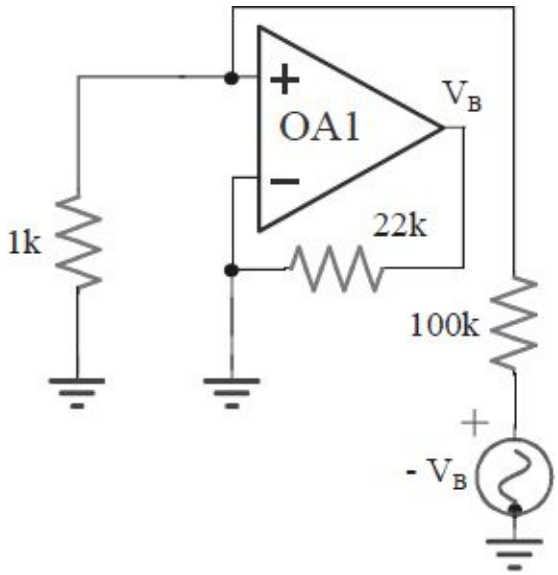
$$V_{O2} = I_+ \cdot 100k\Omega = -V_{IN} \cdot \frac{100k\Omega}{1k\Omega} = -2V$$

$$I_{F2} = \frac{V_{O2}}{22k\Omega} = 90\mu A = I_{F1} = -\frac{V_{O1}}{22k\Omega} \rightarrow V_{O1} = -V_{O2}$$

da cui: $$V_{OUT} = V_{O1} - V_{O2} = -2V_{O2} = 2V_{IN} \cdot \frac{100k\Omega}{1k\Omega} = 200V_{IN} = 4V \cdot$$

5.8.b

To draw the G_{loop} diagram of OA1 it is necessary to evaluate its β. Since OA2 is supposed to be ideal, it is convenient to refer to the following scheme:



To evaluate the β one needs to test the feedback network with a signal v_{test}, and to compute its effect on the OpAmp input terminals. Since the second OpAmp is ideal, it forced its inverting terminal to ground, and consequently also the negative terminal of the first OpAmp. This is to say that *β_0*, because the inverting terminal is fixed to ground independently from the test signal. The situation is different for the positive terminal. In this case the test signal gives birth to a voltage equal to $-v_{test} \cdot 22k\Omega/22k\Omega$ at the second OpAmp output, just like the inverting configuration. This voltage is transferred to the OA1 positive terminal proportionally to the partition between $100k\Omega$ and $1k\Omega$. In the end:

$$v_- = 0$$

$$v_+ = -v_{test} \frac{1k\Omega}{1k\Omega + 100k\Omega}$$

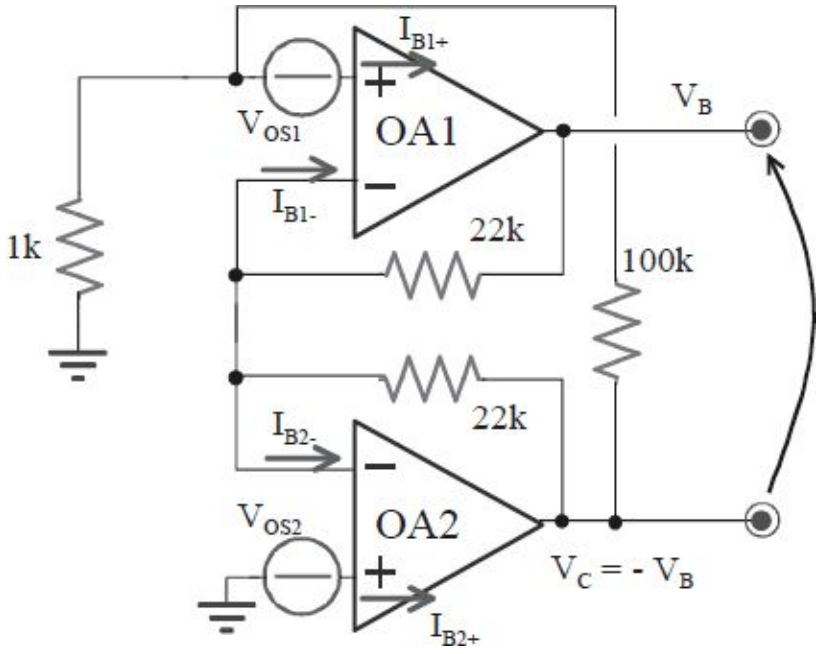
$$\beta = \beta_+ - \beta_- \cong -1/100$$

The G_{loop} can be redesign on the Bode diagram as the difference between *A(s)* and *1/β*. Since the intersection occurs with a *20dB/dec* angle the configuration is stable.

5.8.c

Determine the static error v_{out} on the output when $v_{in}=0$ is equivalent to evaluate the effect of the offset voltages and of the bias currents. To compute the calculations it is appropriate to remember that, assuming an ideal feedback, all the four input terminals of the OpAmps are at the same potential. The bias current on the positive terminal of the first OpAmp follows entirely the path of the feedback resistor, given that all the input terminals are held to ground by the feedback. This causes a voltage on the OA2 output equal to $I_{B1+}\cdot 100k\Omega$. Following a reasoning similar to the point a) one, one can obtain that $V_{out2}=-V_{out1}$, so that the total output error due to the bias current is equal to $2I_{B1+}\cdot 100k\Omega=2mV$.

The bias current on the positive terminal of the second OpAmp flows towards ground, so it does not cause an output contribute. The two bias currents on the negative terminals are connected in parallel and are transferred in the same way. Even if it could seem that they are equally distributed between the two $22k\Omega$ resistors, in reality in the lower one any current can flow. In fact, if some current could flow into it, the OA2 output should be at a potential different from zero, so demanding a current through the $100k\Omega$.feedback resistor. Yet, the required current could not come from OA1, nor from the input $1k\Omega$, because if feedback is well operating it has $0V$ across its terminals. Therefore the bias current flows through the resistor in feedback to OA1, determining $V_{o1}=2I_B\cdot 22k\Omega$. Since in this way $V_{o2}=0$, it results $V_{out}=2I_B\cdot 100k\Omega=2mV$.



The offset voltage of OA1, that we assume applied to its positive terminal for convenience, determines a current flowing through the input resistor that follows entirely the path of the resistor towards the OA2 output. So $V_{o2}=V_{os}+V_{os}\cdot 100k\Omega/1k\Omega$. For the already noticed symmetry, $V_{o2}=-V_{o1}$, so $V_o=V_{o1}-V_{o2}=-2V_{os}-2V_{os}\cdot 100k\Omega/1k\Omega=\pm 1.01V$.

The offset voltage of OA2, that we assume applied to its positive terminal for convenience, is applied to all the other inputs of the OpAmps because of the feedback. So as before $V_{o2}=V_{os}+V_{os}\cdot 100k\Omega/1k\Omega$. But in this case the negative terminals are at V_{os}, not at ground, so the outputs are symmetric with respect to V_{os}. Therefore $V_{o1}=V_{os}-V_{os}\cdot 100k\Omega/1k\Omega$, so $V_o=V_{o1}-V_{o2}=-2V_{os}\cdot 100k\Omega/1k\Omega=\pm 1V$.

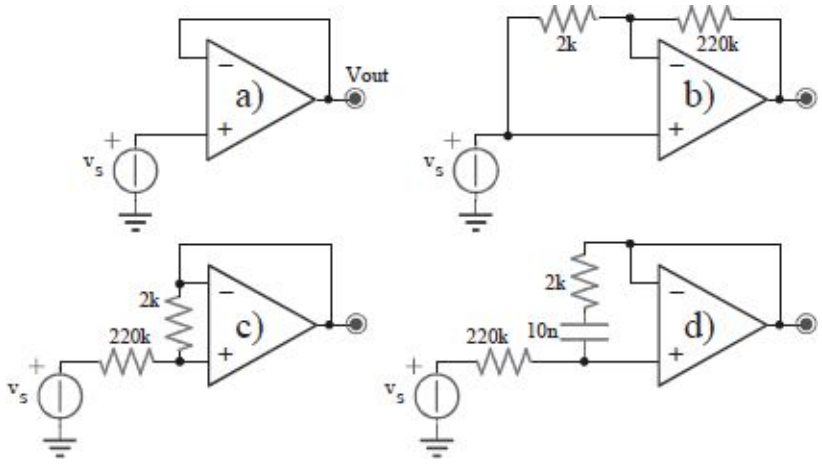
5.9.

The circuits in the figure use the same compensated OpAmp, with $A_0=100dB$, $GBWP=100MHz$ and $v_{noiseOpAmp}=20nV/\sqrt{Hz}$.

a) For each circuit, draw the Bode diagram of $v_{out}/v_{noiseOpAmp}$.

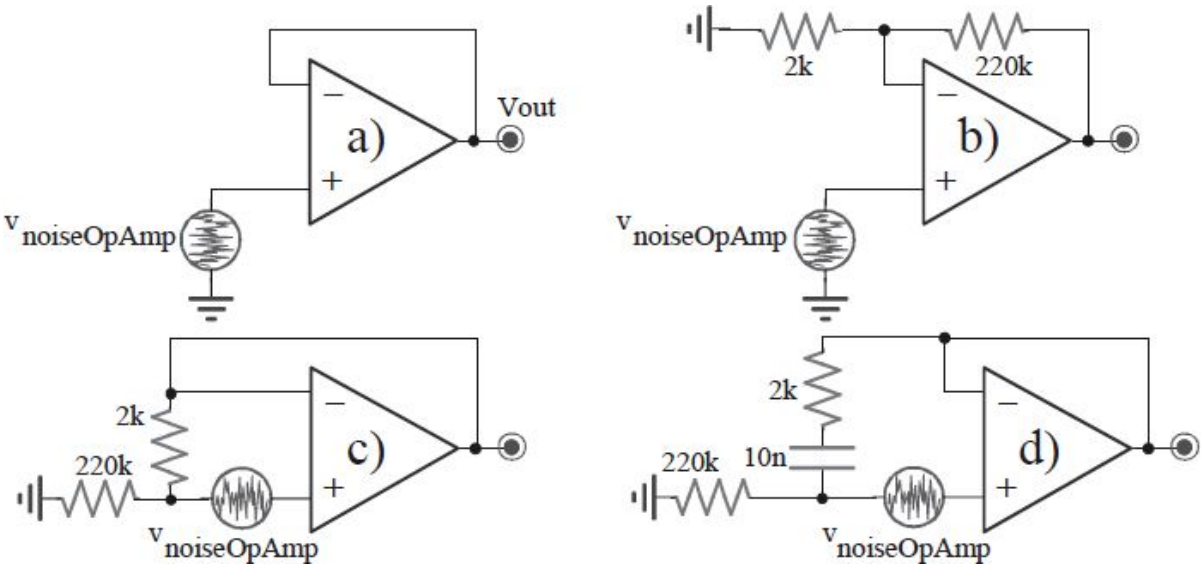
b) For each circuit, compute the effective value of the output noise due to the single OpAmp.

c) For each circuit, compute the **ideal** gain v_{out}/v_s and the **real** impedance from generator v_s.

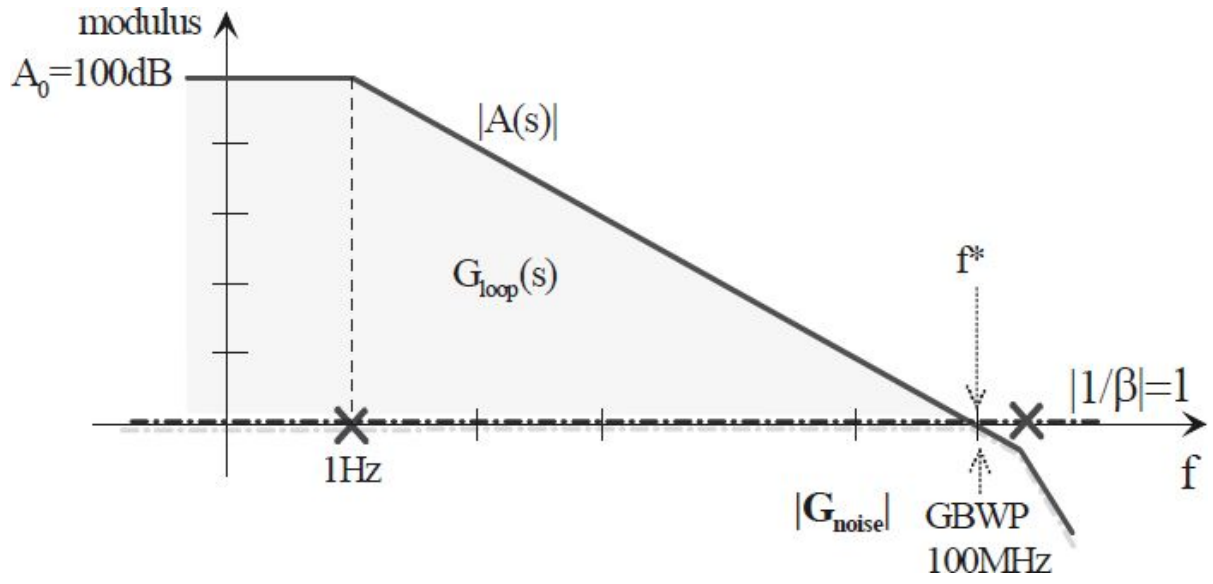


5.9.a

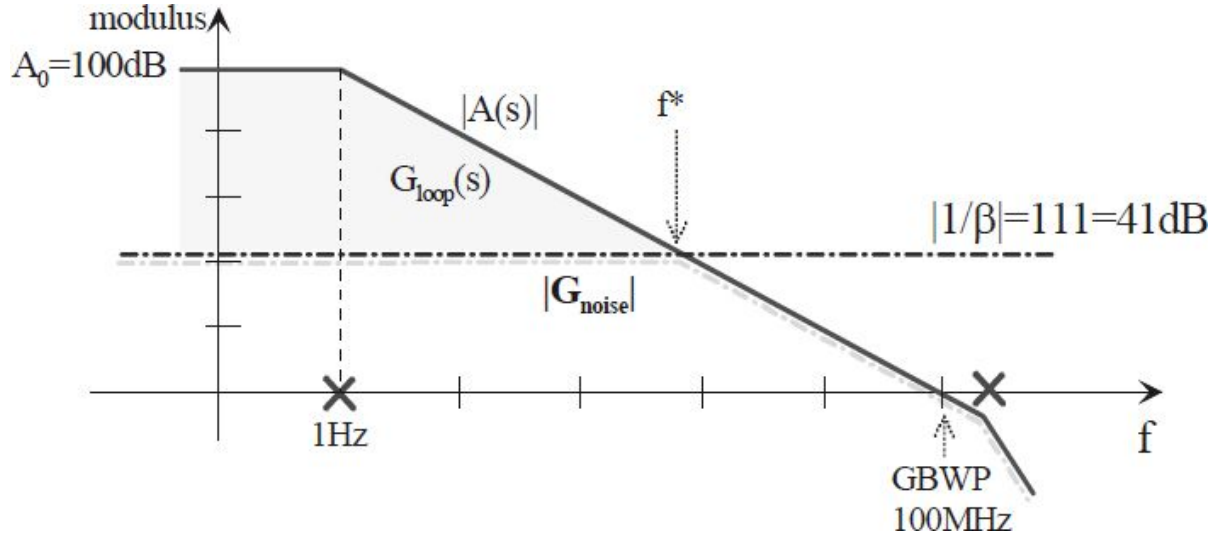
The four circuits under consideration show the same ideal gain for what concerns the signal, because in the network between the input terminals, assuming the feedback to be ideal, any current can flow. So, on signal, they are all buffers. The transfer of the voltage noise generator of the OpAmp, on the contrary, is different in the four cases, as you can see from the following circuits. In fact, the feedback forced the voltages at the input terminals of the OpAmp to be equal. Since the noise source is located between the OpAmp terminal and the node to which the additional networks merge, in this case we can have current flowing in them.



The *a)* circuit is the basic one. Since there is no network between the input terminals of the OpAmp the voltage noise transfer is equal to 1 as the signal one.

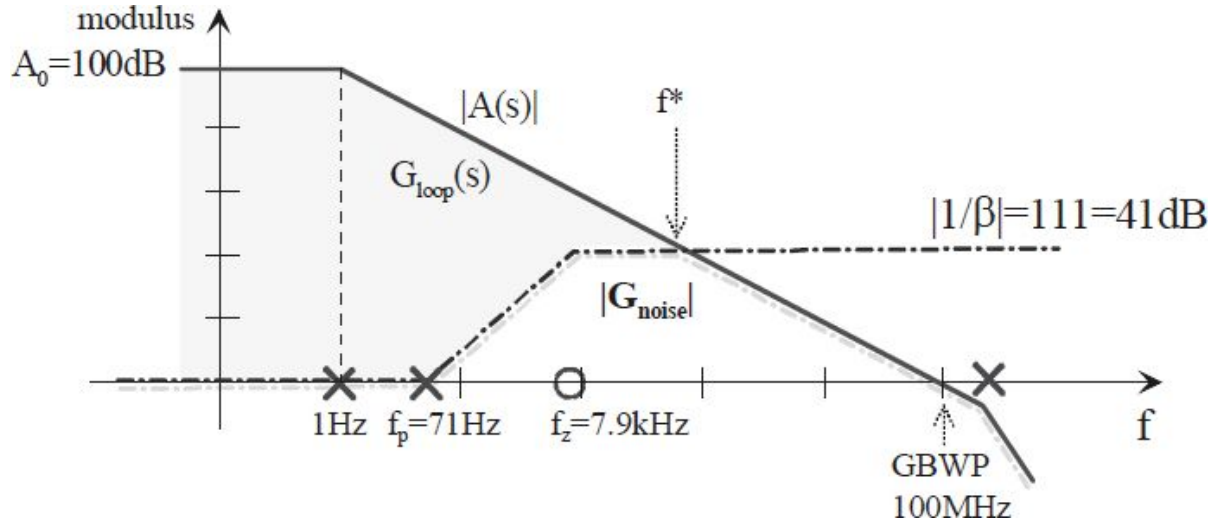


The *b)* circuit is in reality an non-iverting configuration for what concerns the noise, with a gain equal to *1+220kΩ/2kΩ=111*. In fact, with the signal generator disabled, both the noise generator and the *2kΩ* resistor are connected to ground.



The *c)* and *d)* circuits are similar, with the only difference constituted by the network applied between the OpAmp terminals. In this network with impedance *Z(s)* a current of value $v_n/Z(s)$ flows, determining a voltage across the *220kΩ* equal to $220k\Omega \cdot v_n/Z(s)$. On the output, therefore, we have a voltage $v_O=v_n(1+220k\Omega/Z(s))$. The *c)* circuit has consequently a constant gain of 111. The Bode of the gain for what regards the noise is the same of the *b)* case.

Instead, the *d)* circuit has a gain equal to 1 at low frequency and equal to 111 at high frequency, with a pole at the frequency *1/(2π·10nF·2kΩ)=7.9kHz* and a zero at *1/[2π·10nF·(2kΩ+220kΩ)]=71Hz.* Its Bode diagram is shown in the following.



5.9.b

To compute the effective output noise, we need to know not only the ideal transfer function, obtained in the previous section, but also the closed loop bandwidth of the circuit.

In the *a)* case the bandwidth is simply the *GBWP*, so that the effective output noise is equal to $v_{out}=\sqrt{(20nV/\sqrt{Hz})^2 \cdot 100MHz \cdot \pi/2}=250\mu V$.

In the cases *b), c), d)* the feedback network has not a gain equal to 1, so we need to evaluate the $1/\beta$ to find the position in frequency of the closed loop pole.

In the *b)* case, $1/\beta$ is simply $1+220k\Omega/2k\Omega=111$, so the closed loop pole is situated at $GBPW/111=900kHz$. Therefore the effective output noise is equal to $v_{out}=\sqrt{(20nV/\sqrt{Hz})^2 \cdot (111)^2 \cdot 900kHz \cdot \pi/2}=2.6mV$.

The cases *c)* and *d)* can be approximately considered equivalent, given that in the *d)* configuration the ideal gain singularities occur at low frequencies. Therefore we can assume that to evaluate the closed loop pole the capacity is a short circuit. Analyzing separately β^+ and β^- and calculating $\beta= \beta^+-\beta^-$, we obtain $1/\beta=1+220k\Omega/2k\Omega=111$. Also in these two cases the closed loop pole is therefore at *900kHz*, and the effective output noise equal to $v_{out}=\sqrt{(20nV/\sqrt{Hz})^2 \cdot (111)^2 \cdot 900kHz \cdot \pi/2}=2.6mV$. Remember that, however, the *d)* configuration will actually shows a slightly lower noise because the low frequency OpAmp noise will be transferred with a gain equal to 1.

5.9.c

As mention in the *a)* section, the ideal gain for what concerns the signal for the four configurations is the same and equal to 1. In fact with an ideal feedback in the networks between the OpAmp input terminals any current flows, so they have no influence. What is different among the four configurations is the closed loop pole.

For what concerns the impedances from v_s, in the case of ideal feedback they are infinite in all the four cases, because as already specified the networks between the input terminals are insignificant. Actually, in the case of ideal feedback, the impedance is strictly infinite only in the *a)* case. In the other cases the impedance is the open loop one increased proportionally to the G_{loop}. In fact, notice that the network between the input terminals has a *bootstrap* effect. Let's consider for example the *b)* configuration. As a consequence of a positive signal step, a current would flow into the $2k\Omega$ resistor. Yet, as a consequence of the step the OpAmp output increases, and in the same way the non-inverting terminal, so decreasing the current demanding to the generator. Therefore, the effect is to increase the impedance of the node. The impedance effectively computed is equal to $(220k\Omega+2k\Omega)\cdot(1-G_{loop})=(220k\Omega+2k\Omega)\cdot(1+A_0/111)=200M\Omega$.

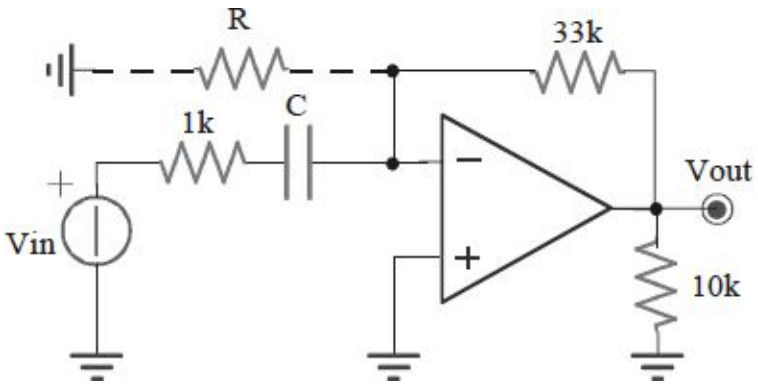
In the cases *c)* and *d)* we need to remember that the resistance is equal to the $220k\Omega$ in series to the resistance from the non-inverting terminal. The last one is equal to the open loop resistance multiplied for the G_{loop}, *evaluated in the new configuration without the* $220k\Omega$.. In the end we obtain $220k\Omega+2k\Omega\cdot(1-G_{loop,\ new})=220k\Omega+2k\Omega\cdot(1+A_0)=200.22M\Omega$. In the *d)* case the presence of the capacity makes the low frequency impedance effectively infinite, while the high frequency value is that just calculated.

5.10.

The OpAmp in the circuit is not-compensated. (A_0=*100dB*, second pole at *20MHz*, A_{min}=*40dB*). Do not consider *R* initially.

a) Starting with *C=160pF* draw the Bode diagram for the t.f. $v_{out}(f)/v_{in}(f)$. Is the stage stable?

b) Connecting the resistor *R=33Ω*, specify the changing in the t.f. $v_{out}(f)/v_{in}(f)$. Is now the stage stable?

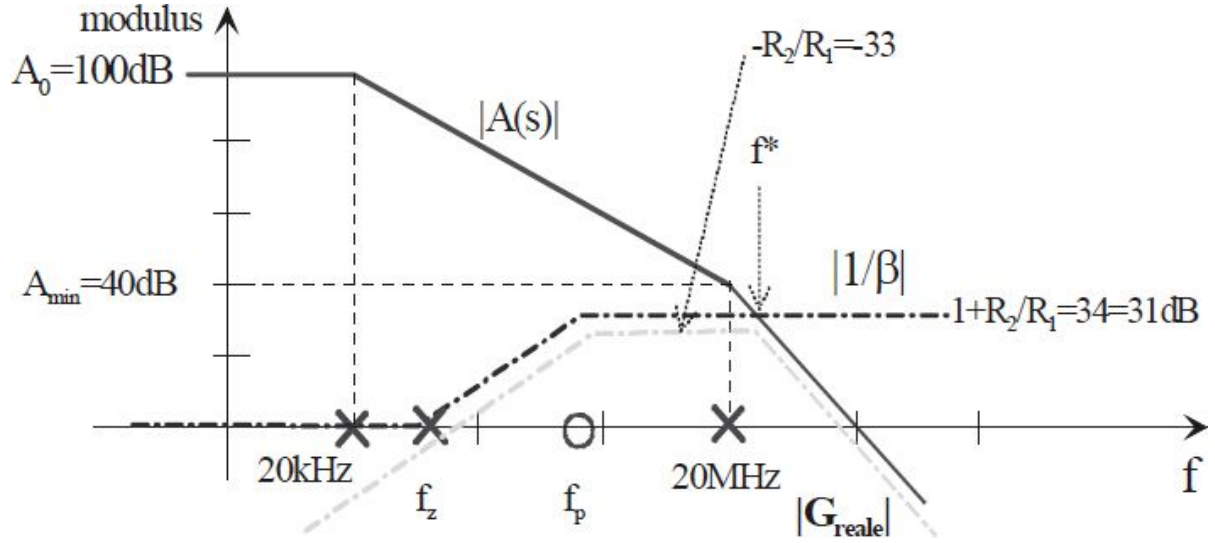


5.10.a

The proposed stage is a real differentiator. Since the used OpAmp is not internal compensated, you have to verify if the $1/\beta$ at high frequency is high enough to cross the A(s) with an angle of *40dB/dec*.

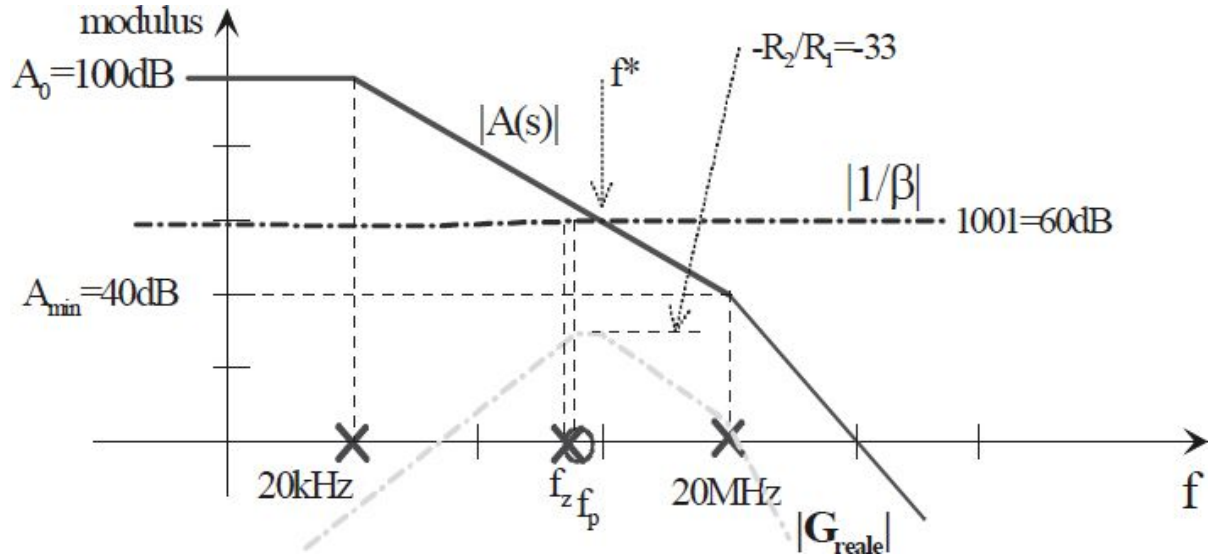
To evaluate the $1/\beta$ instead of the analytics computation it is possible to resort to some synthetic reasoning. At low frequency (*C* open) $1/\beta$ is equal to 1, while at high frequency (*C* closed) is equal to $(1k\Omega+33k\Omega)/1k\Omega=34$. Besides the $1/\beta$ pole, that is the zero of the β, is at $f_p=1/2\pi C\cdot 1k\Omega=995kHz$ (remember: when the signal is interrupted by a stage you have a zero, therefore RC series towards ground or RC parallel along the signal path). Thanks to the constant Gain BandWidth Product concept, you can deduce the $1/\beta$ zero frequency, (ie the β pole) equal to $f_z=995kHz/34=29kHz$. It is possible to compute directly the β pole, equal to $1/2\pi C\cdot(1k\Omega+33k\Omega)=29kHz$.

The A(s) and $1/\beta$ Bode plots are shown in figure. Since $A_{min}=40dB$, the crossing with $1/\beta$ occurs with an angle of *40dB/dec*, and the stage is therefore unstable. In the same graph is shown the ideal transfer $v_{out}(f)/v_{in}(f)$, for the typical inverting real differentiator. One can notice the presence of the same $1/\beta$ pole and the absence of the zero. This is not surprising because the $1/\beta$ is the not-inverting gain and the poles, unlike the zeros, do not depend on the input of the stage but are a property of the circuit.



5.10.b

The introduction of the *R* resistance does not change the ideal transfer, because the new resistance is between ground and a virtual ground node. The most impressive changing is in the alteration of *1/β* that in the new configuration worth *1+33kΩ/33Ω=1001* at low frequency and *1+33kΩ/(33Ω∥1kΩ)≅1001* at high frequency. The pole does not change in frequency, but now the zero is very close: substantially *1/β* is constant with a value of 1001. In the new configuration the intersection between *1/β* and A(s) occurs with an angle of 20dB/dec, making the configuration stable, with the drawback of a f* frequency much lower than before. Therefore the actual gain is affected and does not have a flat response.

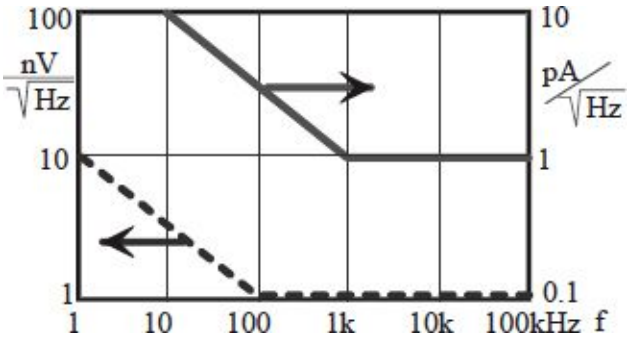


5.11.

The OpAmp has a *GBWP=10MHz*. Using R_1*=100kΩ* and $R_{feedback}$*=200kΩ*, introduce *C* for low-pass filtering at *100kHz*.

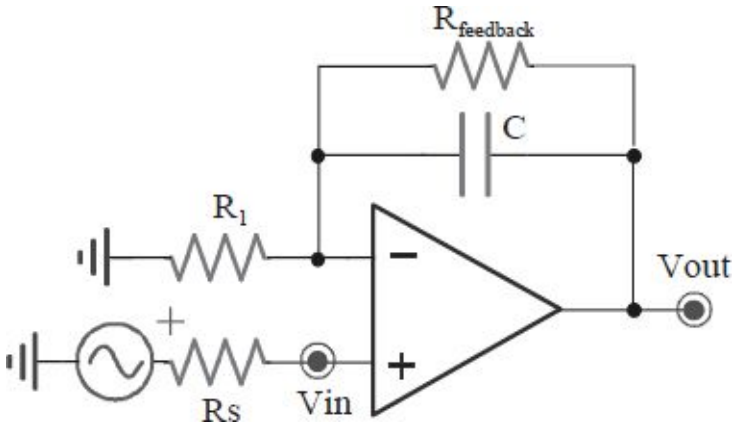
a) Taking into account **all** the noise contribution and a *1kΩ* source connected to **not-inverting** input, compute the output total noise.

b) Starting from **all** the noise contribution, calculate the stage *Noise Figure* at *1kHz*.



5.11.a

The stage is the following. Because the pole introduced by the feedback capacitance has a frequency of $f_p=1/2\pi R_F C_F$, to put arbitrarily the pole one decade before 100kHz one will obtain $C_F=80pF$.



Consider now the transfer function for the various noise source. The noise for the resistance R_1 gets multiplied by the inverting gain, therefore $v^2_{R1}=4kTR_1\cdot(R_F/R_1)^2$. The noise for the resistance R_F is already on the output and then $v^2_{RF}=4kTR_F$. The noise for R_S gets multiplied by the not-inverting gain, i.e. $v^2_{RS}=4kTR_S\cdot(1+R_F/R_1)^2$. Regarding the OpAmp noise, the input contribution is multiplied by the not-inverting gain before the output, obtaining $v^2_{OPA,e}=e^2_n\cdot(1+R_F/R_1)^2$. The current noise for the OpAmp negative input gets multiplied by the feedback resistance, thus $v^2_{OPA,i-}=i^2_n\cdot R_F{}^2$. It is possible to convert the positive input current noise in voltage noise multiplying it by the R_S resistance using the not-invertin gain to compute the output noise, $v^2_{OPA,i+}=i^2_n\cdot R_S{}^2\cdot(1+R_F/R_1)^2$.

Obviously, to compute the overall noise you have to integrate the noise density on the bandwidth of interest, but the transfer functions are different for every density. The transfer function for the noise source connected to the inverting input and for the feedback resistance is a single pole t.f. (with a frequency $f_{pn}=f_{ps}\cdot\pi/2=15.7kHz$), while the noise sources applied to the not-inverting input have a zero and a pole; thus the ideal gain after the pole is unit

up to the OpAmp GBWP. Furthermore the OpAmp noise sources have a *1/f* component.

We can consider now different contribution. The R_1 noise undergoes only the pole of the stage, then $noise_{R1}=4kTR_1\cdot(R_F/R_1)^2\cdot f_{pn}=(10\mu V)^2$. The same is for the R_F noise, thus $noise_{RF}=4kTR_F\cdot f_{pn}=(7\mu V)^2$. On the other hand the R_S noise undergoes both pole and zero. However the gain is low (with a value of 3) and it is possible to overestimate the noise assuming the stage gain equal to 3 up to the GBWP, therefore $noise_{RS}=4kTR_S\cdot(1+R_F/R_1)^2\cdot GBWP=(38\mu V)^2$. The OpAmp voltage noise undergoes both zero and pole, too. Assuming the reasoning made before we obtain $noise_{OPA,e}=e^2_n\cdot(1+R_F/R_1)^2\cdot GBWP=(9\mu V)^2$. The *1/f* noise contribution is equal to:

$$noise_{OPA,e,1/f}=k\cdot\left(1+\frac{R_F}{R_1}\right)^2\int_1^{GBWP}\frac{1}{f}df=\left(10\frac{nV}{\sqrt{Hz}}\right)^2\cdot 3^2\ln\frac{10MHz}{1Hz}=(120nV)^2$$

Consider now the current noise; the OpAmp inverting input noise undergoes the sole stage pole, thus $noise_{OPA,i-}=i^2_n\cdot R_F^2\cdot f_{pn}=(25\mu V)^2$. You must add the *1/f* noise contribution, that is:

$$noise_{OPA,i-,1/f}=k\cdot R_F^2\int_1^{15.7kHz}\frac{1}{f}df=\left(32\frac{pA}{\sqrt{Hz}}\right)^2\cdot(200k\Omega)^2\ln\frac{15.7kHz}{1Hz}=(20\mu V)^2$$

On the other hand, the positive input noise undergoes both the pole and the zero of the stage, therefore $noise_{OPA,i+}=i^2_n\cdot R_S^2\cdot(1+R_F/R_1)^2\cdot GBWP=(9.5\mu V)^2$. This term should also be added to the *1/f* contribution equal to:

$$noise_{OPA,i+,1/f}=k\cdot R_S^2\left(1+\frac{R_F}{R_1}\right)^2\int_1^{GBWP}\frac{1}{f}df=\left(32\frac{pA}{\sqrt{Hz}}\right)^2\cdot(1k\Omega)^2\cdot 3^2\ln\frac{10MHz}{1Hz}=(380nV)^2$$

Essentially, adding all the quadratic contribution, the output noise is equal to $58\mu V_{rms}$.

5.11.b

To compute the Noise Figure at *1kHz* it is more useful to use the definition: $NF = \frac{TotalOutputNoise}{OutputNoiseSource}$. Therefore:

$$NF = 10\log\frac{4kTR_1 \cdot G_{inv}^2 + 4kTR_F + 4kTR_S \cdot G_{noninv}^2 + e_n^2 \cdot G_{noninv}^2 + i_n^2 \cdot R_S^2 \cdot G_{noninv}^2 + i_n^2 \cdot R_F^2}{4kTR_S \cdot G_{noninv}^2} =$$

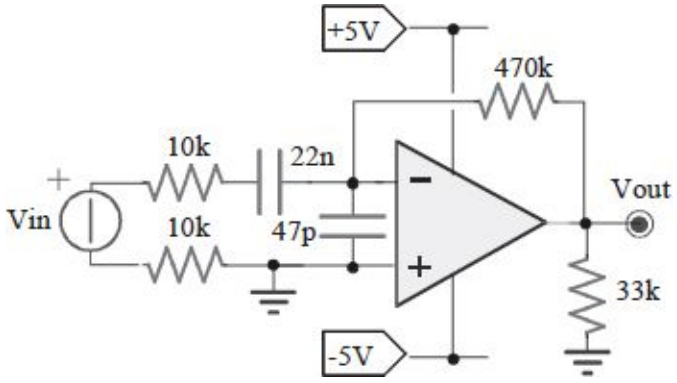
$$= 10\log\frac{(160\frac{nV}{\sqrt{Hz}})^2 + (56\frac{nV}{\sqrt{Hz}})^2 + (12\frac{nV}{\sqrt{Hz}})^2 + (3\frac{nV}{\sqrt{Hz}})^2 + (3\frac{nV}{\sqrt{Hz}})^2 + (200\frac{nV}{\sqrt{Hz}})^2}{(12\frac{nV}{\sqrt{Hz}})^2} = 26dB$$

5.12.

OpAmp with A_0=*100dB*, *GBWP=80MHz*, V_{OS}=*15mV*, I_B=*1µA*, *SR=10V/µs*. The input parasitic capacitance equal to *47pF* is not eliminable.

a) Compensate the stage to assure *MF=90°*.

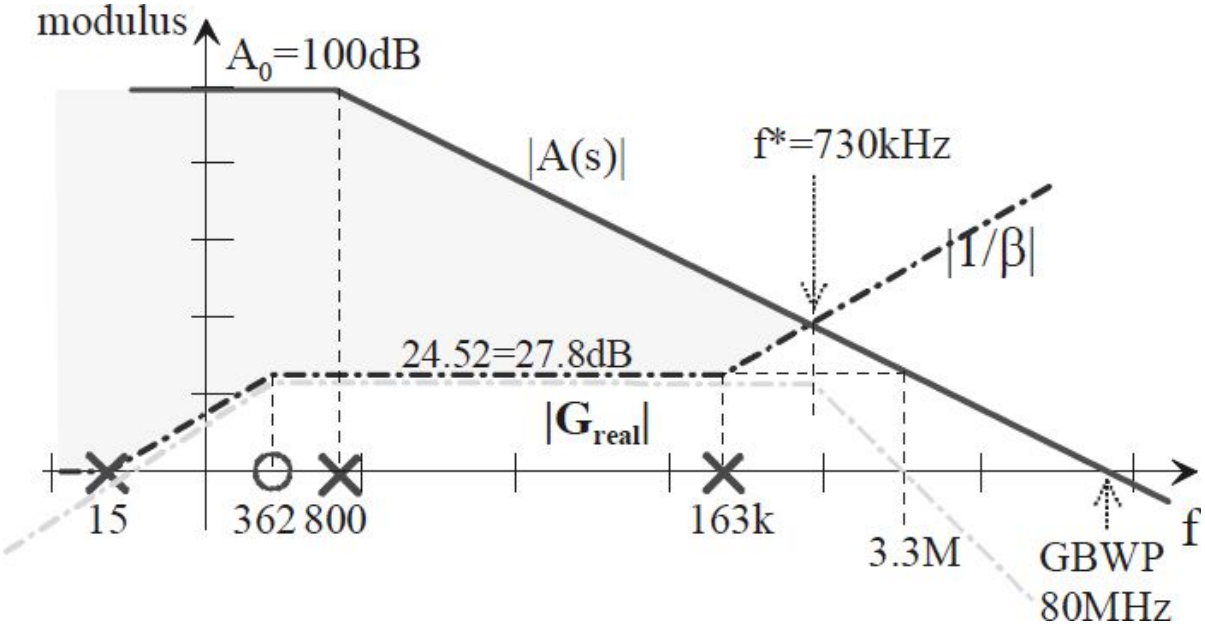
b) Without the *47pF* parasitic capacitance, apply a *+100mV* step input and draw the output voltage quoted trend.



5.12.a

To compensate the stage one has to evaluate the feedback network, taking into account the input capacitance that is not eliminable. Proceeding briefly to study β, and observing that the positive input is connected to ground, it is possible to note that:

- At high frequency β is null.
- At low frequency β is unitary.
- At medium frequency β is equel to $20k\Omega/(20k\Omega+470k\Omega)=0.04$.
- There is a zero with a frequency of $f_z=1/2\pi\cdot22nF\cdot20k\Omega=362Hz$.
- To compute the first pole, it is supposed that the parasitic capacitance is open, because it's smaller than the other and then it will introduce a higher frequency pole. Thus, the first pole is $f_{p1}=1/2\pi\cdot22nF\cdot490k\Omega=15Hz$.
- To compute the second pole, it is supposed that the input capacitance is a short, because it's greater than the other and then it will introduce a lower frequency pole. Thus, the second pole is $f_{p2}=1/2\pi\cdot47pF\cdot470k\Omega//20k\Omega=176kHz$.

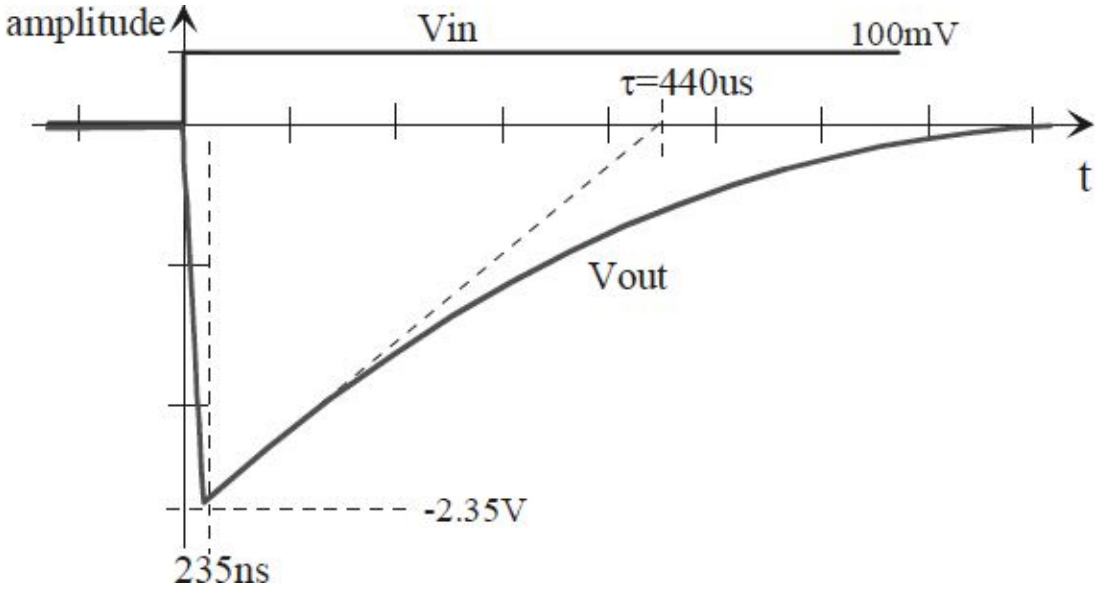


The Bode diagram for $1/\beta$ and $A(s)$ is shown in figure. It is observed that the intersection between the two graphs occurs with an angle of $40dB/dec$, therefore the stage is unstable. To make it stable the second pole of β must be

moved at higher frequency. Assuming that, after the compensation, the $1/\beta$ at medium frequency is equal to *24.5*, it is required that the second pole intervenes almost a decade after $GBWP/24.5=3.3MHz$. Actually, a possible compensation without affecting the ideal gain of the stage consists to insert a parallel resistance in the parasitic capacitance. The new resistance modifies the position for the two poles, moving them at higher frequencies. The $1/\beta$ value at low and medium frequencies changes, too. For example, inserting a $1k\Omega$ resistance it is obtained that $f_{z1}=1/2\pi\cdot 22nF\cdot 21k\Omega=344Hz$, $f_{z2}=1/2\pi\cdot 47pF\cdot 470k\Omega//20k\Omega//1k\Omega=7\cdot 2MHz$, low frequency gain equal to *471*, while the medium frequency gain is *517*. In this way the stage become more stable, with a closed loop pole at a frequency of $f_{CL}=GBWP/517=154kHz$.

5.12.b

The stage is an approximated differentiator. Therefore, the highest frequencies will have a finite gain, while the DC will be stopped. It is evident that providing an input step signal the output will rise suddenly then will go down to zero exponentially. The edge of the step signal behaves as an *"infinite frequency"* and then it will be amplified, while the DC regime is a signal at f=0Hz that will not pass. Obviously, the transition from the peak to ground will occur with the differentiator time constant, equal to $\tau=20k\Omega \cdot 22nF=440\mu s$ (i.e. the pole with a frequency of 362Hz). The peak amplitude will be (consider *C* a short, in virtue of "infinite" bandwidth of the edge) $V_{O,peak}=-100mV \cdot 470k\Omega/20k\Omega=-2 \cdot 35V$. Actually this voltage will not be reach immediately, because of the limit of the OpAmp. The response speed is limited by the slew-rate and by the closed loop pole. Considering the sole slew-rate, the rise time is equal to *235ns*.

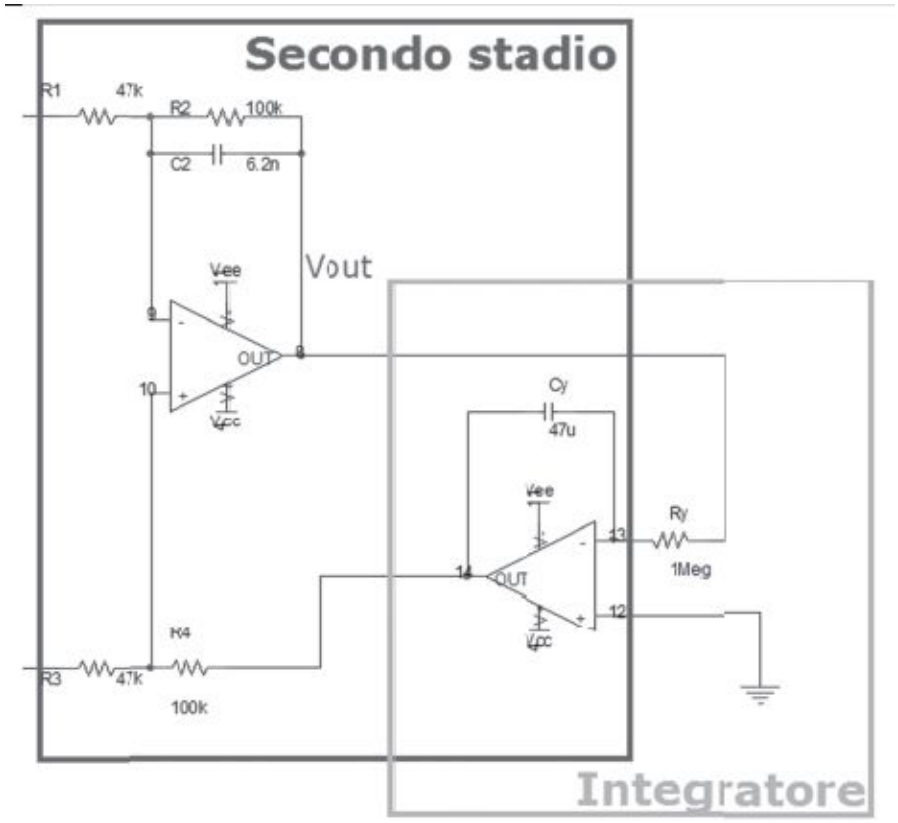


5.13.

The third OpAmp of an INA is followed by another OpAmp connected as an integrator.

a) Compute the gains $v_{out}/v_{DIFF,in}(f)$ e $v_{out}/v_{CM,in}(f)$ without the integrator, with R_4 connected to $+2V$, not to pin 14.

b) Draw the Bode diagram for the gain $v_{out}/v_{DIFF,in}(f)$ when the integrator is connected and comment the purpose.



5.13.a

Without the integrator and with R_4 connected to *+2V*, the circuit is similar to a common differential amplifier, with a DC gain $G_{diff}=R_2/R_1=2.12$ and $G_{CM}=0$, with the only differences of feedback capacitance and R_4 not connected to ground. The latter has no effect on the signal transfer function because the resistance is connected with a low impedance. Obviously, without signal the OpAmp output will not be connected to ground but will have a signal of *+2V*: substantially is the *Ref* input of the INA.

The feedback capacitance determines the filtering on the input signal, with a pole at the frequency of $1/2\pi R_2C_2=256Hz$. However is observed that, for the non-inverting input, this capacitance introduces also a zero at the frequency of 800Hz and at high frequency the gain is unit (you can exploit the property of the gain-bandwidth product to compute the zero frequency). The non-inverting input undergoes only the pole. Since we are interested in the differential and common-mode gain, this zero introduces an asimmetry in the stage, because at high frequency the sole negative input signal will be attenuated. Of course, this difference will cause a deviation of the differential gain in respect to the traditional differential amplifier gain. Moreover the common-mode signal will be carried out. Substantially it is obtained:

$$\frac{V_O}{V_{diff}}=\frac{R_2}{R_1}\cdot\frac{1+sC_2\cdot\frac{R_1 // R_2}{2}}{1+sC_2R_2}$$

$$\frac{V_O}{V_{cm}}=\frac{R_2}{R_1}\cdot\frac{sC_2\cdot R_1 // R_2}{1+sC_2R_2}$$

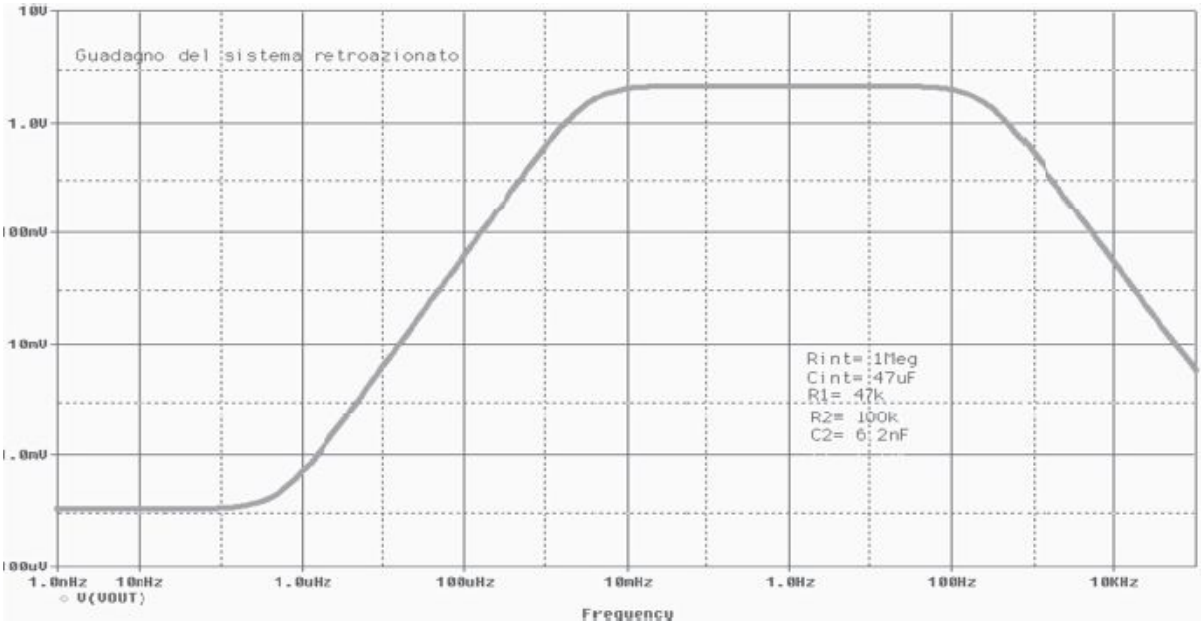
Notice how at low frequency the transfer function is lower than expected, while at high frequency dominates the non-inverting transfer, because the inverting input signal is attenuated. For this reason the common-mode transfer is null in DC, growing with frequency, up to the value of the sole non-inverting input transfer.

5.13.b

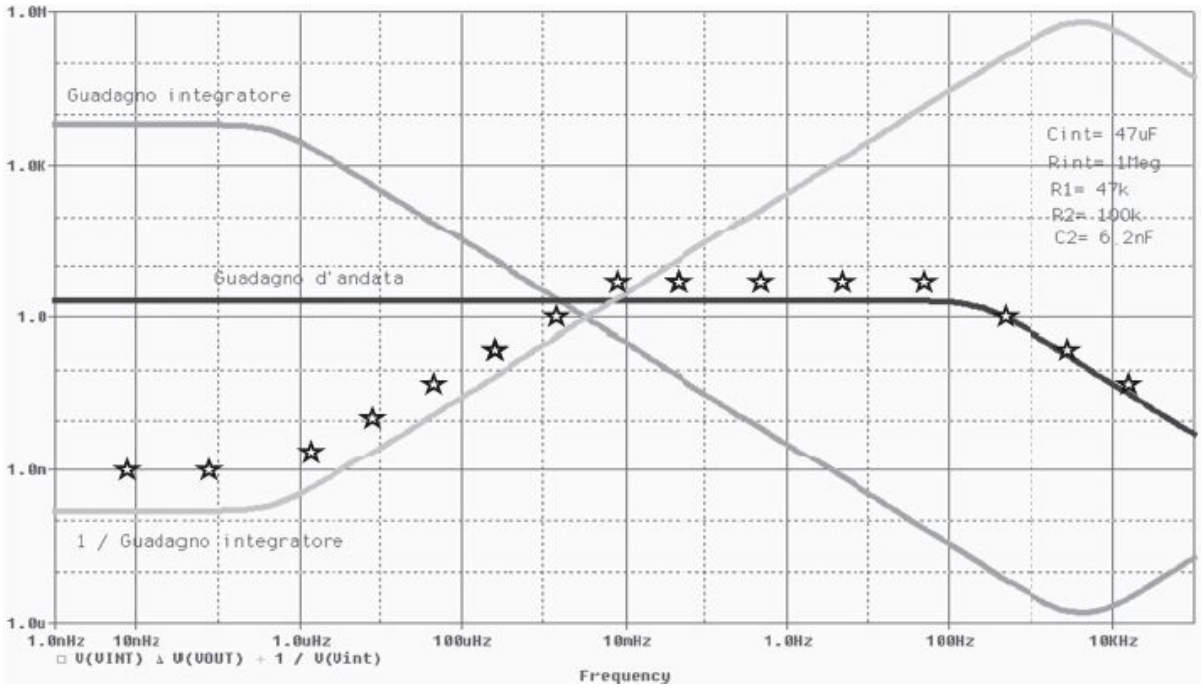
Noting that the *Ref* input is a further input with a gain of -1, it is possible to write the following relationship:

$$V_O = V_{diff} \cdot A_{diff} - \frac{V_O}{sR_yC_y} \quad \rightarrow \quad \frac{V_O}{V_{diff}} = A_{diff} \cdot \frac{sC_yR_y}{1+sC_yR_y}$$

Therefore the output integrator causes a high-pass filtering of the differential input signal. The reason being that the integrator extract the average value of the output signal and then applies such an average to the *Ref* pin, is being subtracted from the output signal, making the high-pass filtering. You can see the following transfer function diagram.



It is possible to apply the known concepts for the feedback: the forward gain is A(s)=1 for the “wire” connected to the OpAmp pin 8 and the feedback block β(s) is due to the integrator (for this reason *1/β* is equal to a rising straight line with a slope of +20dB/dec). The graphs would be the following: the closed-loop gain is indicated by the stars.

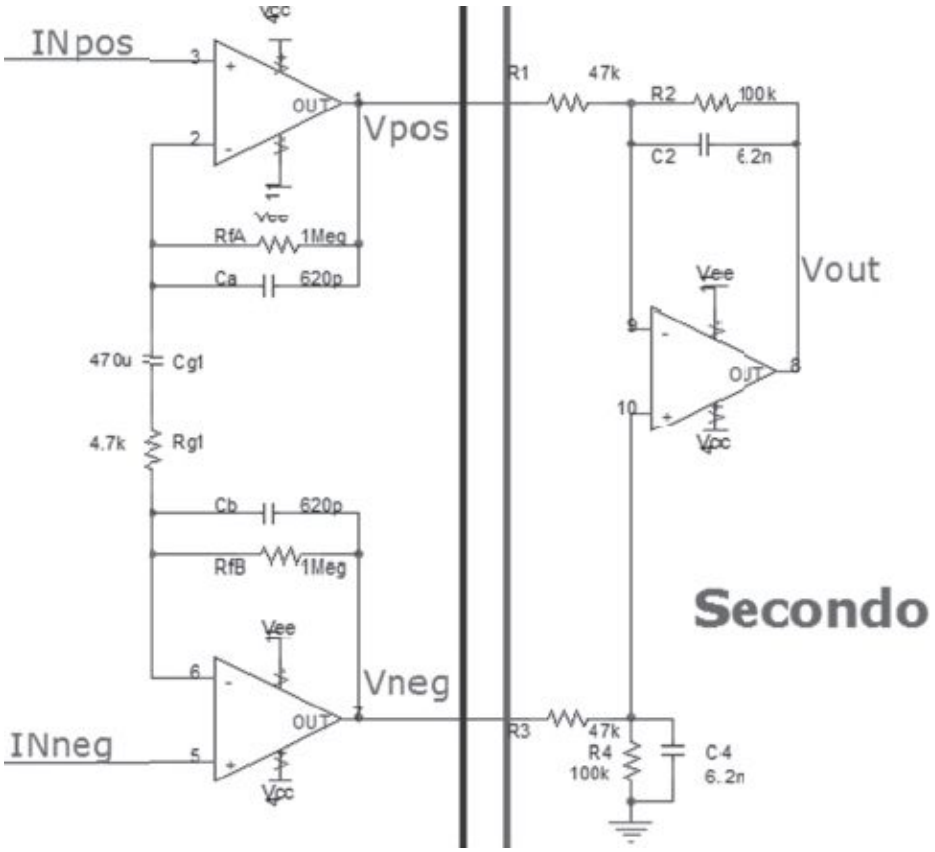

1.0M
1.0K
1.0
1.0n
1.0u
Guadagno integratore
Guadagno d'andata
1 / Guadagno integratore
Cint= 47uF
Rint= 1Meg
R1= 47k
R2= 100k
C2= 6.2nF
1.0nHz
10nHz
1.0uHz
100uHz
10mHz
1.0Hz
100Hz
10KHz
□ V(VINT) △ V(VOUT) ◇ 1 / V(Vint)
Frequency

5.14.

The OpAmps have $A_0=100dB$ and $GBWP=1MHz$.

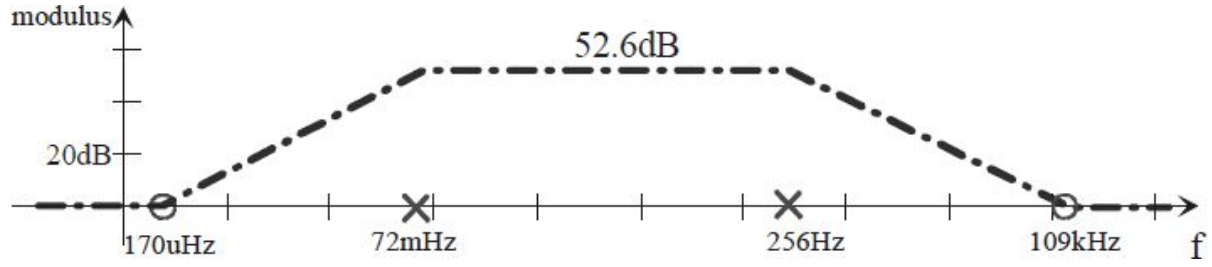
a) Draw the transfer function for the differential-mode gain $(v_{pos}-v_{neg})/(IN_{pos}-IN_{neg})$ for the first stage.

b) **Without C_4,** draw the quoted Bode diagram for v_{out}/v_{pos} and v_{out}/v_{neg} for the second stage, commenting on the need for C_4.



5.14.a

To compute the differential-mode gain $(v_{pos}-v_{neg})/(IN_{pos}-IN_{neg})$, in order to avoid the analytical calculations, it is more practical to proceed in a synthetic way. We can observe that in DC (open capacitances) the differential gain is unit (in fact th two OpAmps are buffers). Also at high frequency (closed capacitance) the transfer is unit, because OpAmps are still buffers. At medium frequency (only C_{g1} closed) the gain is equal to $G_{medium}=1+2R_f/R_g=52.6dB$. The transfer, having three ranges with different gains must have two poles and two zeros. Assuming that the singularities will be far enough, we can evaluate the first pole considering R_f open. The pole intervenes with a frequency $f_{p1}=1/2\pi R_g C_g=72mHz$. In the same way to estimate the pole at high frequency we can assume R_g closed, obtaining $f_{p2}=1/2\pi R_f C_f=256Hz$. Using the gain-bandwidth product and ratio properties zeros' frequencies are extracted, $f_{z1}=170\mu Hz$ and $f_{z2}=109kHz$. The corresponding Bode diagram is shown in figure.



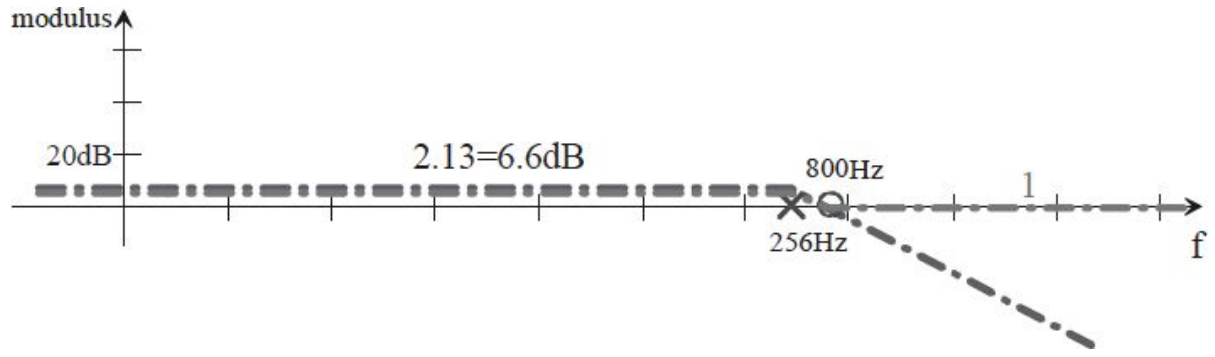
5.14.b

Without C_4, the second stage feedback capacitance determines a filtering on the input signal, with a pole frequency of $1/2\pi R_2C_2=256Hz$. The V_{pos} input undergoes the sole pole and has an inverting gain of -2.13. Indeed on the V_{neg} input, this capacitance introduces a zero and at high frequency the gain v_{out}/v_{neg} is unit, while at low frequency is equal to $100k/147k\cdot(1+100k/47k)=+2.13$. We can obtain graphically the value for this zero ($800Hz$) using the gain-bandwidth product. This zero introduces an asymmetry in the stage, because at high frequency the sole V_{pos} input signal will be attenuated. Because of this asymmetry, the differential gain will be different in respect to the normal differential amplifier. Substantially we have:

$$\frac{V_O}{V_{neg}}=\frac{R_2}{R_1}\cdot\frac{1+sC_2\cdot R_1//R_2}{1+sC_2R_2}$$

$$\frac{V_O}{V_{pos}}=-\frac{R_2}{R_1}\cdot\frac{1}{1+sC_2R_2}$$

The Bode diagrams are shown in figure.



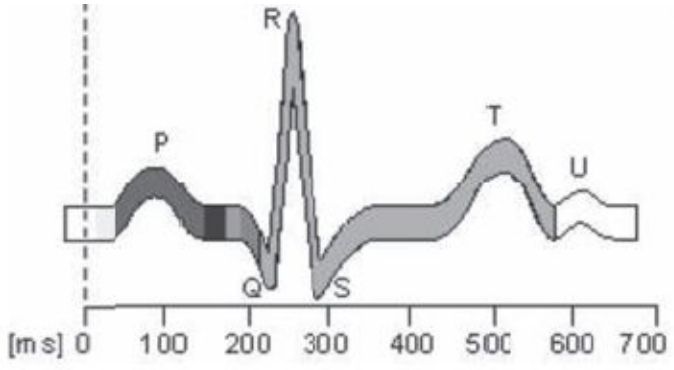
The C_4 capacitance would introduce an further pole on the V_{neg} transfer function, exactly at the same frequency of the zero. In this way the gains v_{out}/v_{pos} and v_{out}/v_{neg} for the second stage would be the same (with the exception of the sign), restoring the correct behavior at all frequencies.

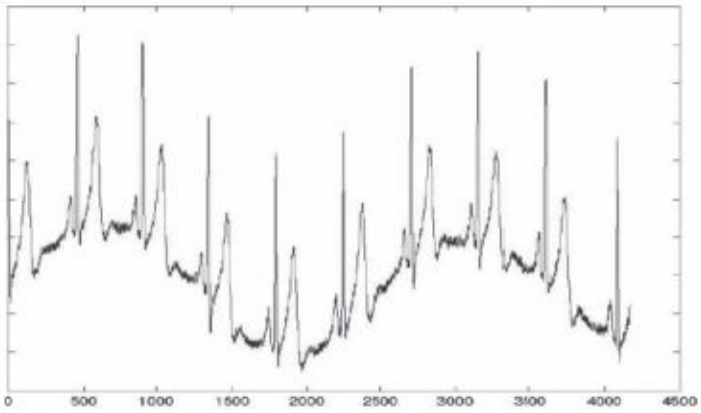
5.15.

Starting from the amplified ECG signal measure the peak amplitude of the ventricular repolarization (T peak with an amplitude of *100÷800mV*), masked by the massive left ventricular (R of *1÷3V*).

a) Design the circuit that, with the base line centered around the *0V*, is synchronized on the P peak (with an amplitude of about *20÷200mV*), and masks for *300ms* the signal and then extends the T peak.

b) Because the base line is not around *0V*, but moves slowly between ±*2V* as shown in figure, design the circuit that subtracts the base line signal (time variant) and centers around *2V*.





5.15.a

It is possible to splitt the circuit into two sections, the first generating a signal synchronized on the peak P arrival, which, starting from this *sync* signal extends the peak *T*.

The first section may be formed by a peak stretcher followed by a comparator with a *100mV* threshold, in order to detect the arrival of the P peak. The time constant of the extension must be long enough to cover the entire burst, in order to avoid that subsequent *R*, *T* and *U* peaks latch the comparator.

The second section will, at rest, mask completely the signal and only after an input edge, must propagate the signal form the peak stretcher after *300ms* delay. Substantially, a *MOSFET* switch will allow the signal transit to the peak stretcher. The *MOS* will be controlled by a delayed repetition of the comparator output signal. The simplest way to obtain this replica consists to use a *RC* network followed by an appropriate discriminator, for example an inverter. Estimating conveniently the time constant and the discriminator threshold it is possible to obtain the required delay.

5.15.b

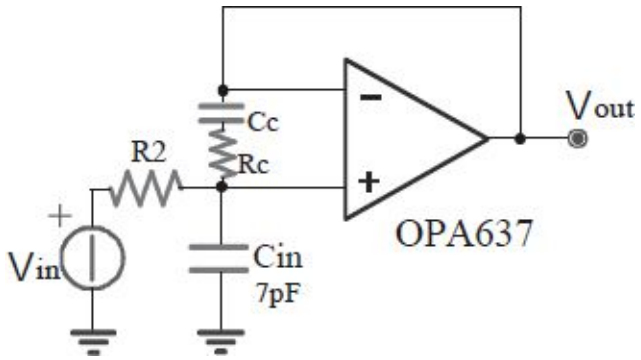
Starting from the waveform it can be noticed that on the signal is superimposed a variable base line with a very low frequency. It is enough to implement a high-pass filter on the signal and then add a $+2V$ constant baseline. It could simply be done with an OpAmp. The period it's around $4s$, and the high-pass filter pole is at a frequency of $1Hz$.

5.16.

The source has a series resistance $R_2=100k\Omega$, the OpAmp is not compensates ($A_0=120dB$, $A_{min}=60dB$, $f_2=1MHz$) and has a noise equal to $10nV/\sqrt{Hz}$.

a) Compensate the stage to ensure $MF=45°$.

b) Compute the output *RMS* noise due to the OpAmp.



5.16.a

The circuit uses an OpAmp not compensated. To estabilish if the stage is stable or not we must evaluate the β. Since two feedback are present, you must compute separately β_+ and β_-, and then calculate $\beta=\beta_+-\beta_-$. It is immediately noticed that $\beta_-=1$, while it is easier to estimate briefly β_+ instead to calcilate it analytically. Assuming that the effect of C_{in} is present at higher frequency then C_C, it is possible to operate roughly as follow:

- β_+ at low frequency is zero (so there is a zero in DC);
- β_+ at high frequency is zero;
- β_+ at medium frequency (C_C closed and C_{in} open) is $R_2/(R_C+R_2)$;
- the pole due to C_C has a frequency equal to $f_{p1}=1/2\pi\cdot C_C\cdot(R_C+R_2)$;
- the pole due to C_{in} has a frequency of $f_{p2}=1/2\pi\cdot C_{in}\cdot(R_C//R_2)$.

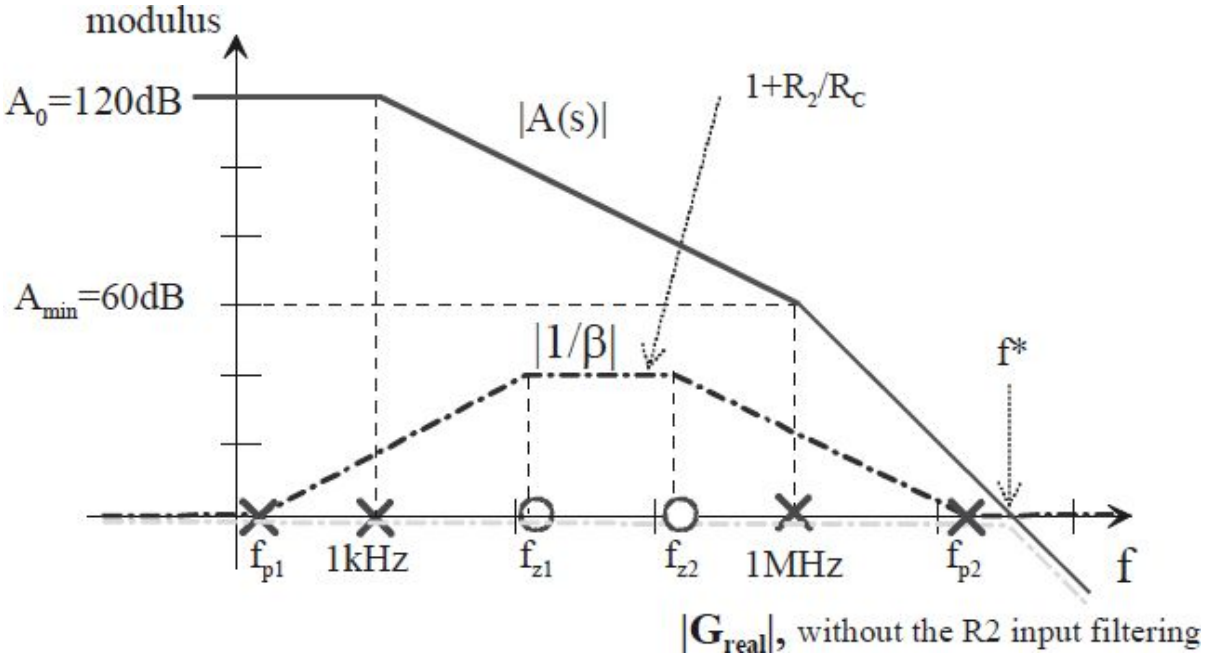
The β_+ has the following expression:

$$\beta_+ = \frac{sR_2C_C}{\left[1+sC_C(R_2+R_C)\right]\cdot\left[1+sC_{in}(R_2 // R_C)\right]}$$

The $\beta=\beta_+-\beta_-$ is equal to:

$$\beta = \frac{sR_2C_C}{\left[1+sC_C(R_2+R_C)\right]\cdot\left[1+sC_{in}(R_2 // R_C)\right]} - 1 = -\frac{s^2R_2C_CR_CC_{in} + s\left[C_{in}(R_2 // R_C)+C_CR_C\right]+1}{\left[1+sC_C(R_2+R_C)\right]\cdot\left[1+sC_{in}(R_2 // R_C)\right]}$$

The zeros cannot be easily simplified, therefore you can observe that β is -1 at both low and high frequency, while at medium frequency (C_C closed and C_{in} open) is $R_C/(R_C+R_2)$. The Bode diagrams for $1/\beta$ and $A(s)$ are shown in figure:



To have at least *MF=45°* the intersection should be with and angle of *40-20dB/dec*. Substabtially the second pole for *1/β* should coincide with the second pole of *A(s)* both in frequency and amplitude. Therefore *1/β* at medium frequency must be equal to *1000 (60dB)*, while the second pole must be at *1MHz*. To satisfy the first condition it will be $R_C=R_2/999=100\Omega$, while the *1/β* second pole is due to C_{in} and has a fixed frequency equal to *225kHz* (you must use the Gain-BandWidth Product to obtain the result). Basically the phase margin is *PM>>45°*, regardless the *Cc* value (as long as the singularities have a frequency less than *225kHz*). The presence of *Cc* is only for increasing the G_{loop} at low frequency.

5.16.b

The OpAmp voltage noise source is applied to the positive pin. With ideal feedback, it determines a voltage drop v_n across the series R_C-C_C. To evalute the output transfer it is better to proceed in a syntetical way. It is noticed that the gain is unitary at low and high frequency, while at medium frequency it is $1+R_2/R_C=1000$. The poles are the same for $1/\beta$, because they don't rely on selected input. For this reason also the zeros are the same: substantially the ideal transfer on the noise is $1/\beta$. The real tranfer will have an further pole with higher frequency, caused by the closed-loop pole. To calculate the *RMS* noise, because the pole-zero pair has a low frequency, we can overestimate the noise approximating the transfer with a single pole with a noise frequency equal to $225kHz \cdot \pi/2=353kHz$ and gain 1000. The overestimation is required because we don't take in account that the trasfer at low frequency is unitary and beyond the closed-loop pole the slope is *-40dB/dec*. We obtain that:

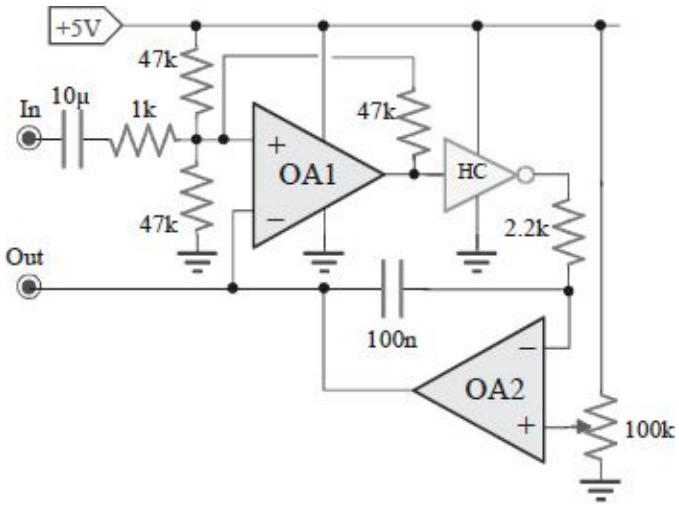
$$v_n = \sqrt{e_n^2 \cdot 1000^2 \cdot 353kHz} = 6mV_{rms}$$

5.17.

Consider the circuit.

a) With floating input and halfway trimmer, draw the quoted waveforms for all nodes and explain the task of the circuit.

b) Explain what will happen when the input would be connected with a sinusoidal source with a peak amplitude of *2V* and a frequency of *1Hz*.



5.17.a

The circuit is an inverting Schmitt trigger, feedback with an integrator. Substantially, until the trigger output is high, the integrator output voltage will rise with a ramp slope. When this waveform reach the high threshold the trigger will latch, bring the output to ground. Now the integrator will produce a descending ramp output, until cross the low threshold or bring the output to the high level. Therefore the circuit is an oscillator, with a triangular waveform as integrator output and square waveform as trigger output.

The oscillator half-period depends on the amplitude that the integrator output has to cover (equal to the trigger hysteresis). Because the trimmet is halfway, the current flowing upon the capacitance in the two half-periods is equal to $\pm 2.5V/2.2k\Omega = \pm 1.1mA$, thus the integrator output will change with a slope of $\pm 11V/ms$. The trigger thresholds are instead equal to $+5V \cdot 0.33 = 1.65V$ and $+5V \cdot 0.66 = 3.3V$, with an hysteresis of $1.65V$. The half-period is therefore $1.65V / 11V/ms = 150\mu s$.

5.17.b

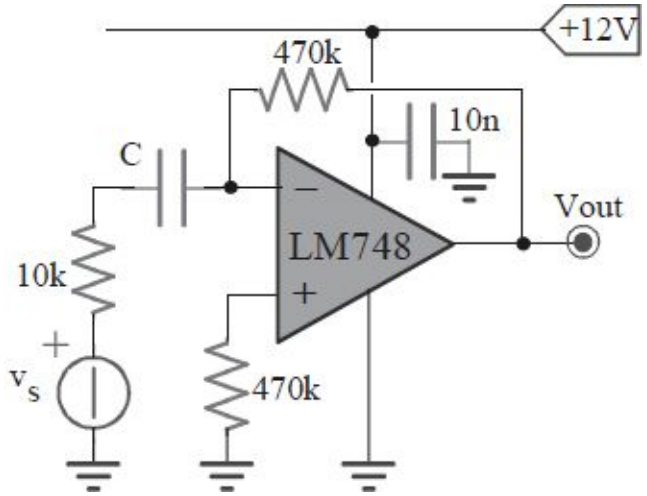
The sinusoid is very slow in respect to the oscillator period. The decoupling capacitance stops the *DC*, but *1Hz* the impedance due to this capacitance is around *16kΩ*. The effect on the input is therefore to shift slowly the trigger threshold, but with no change in hysteresis and in oscillating period. A slow drift in the output peak values will be noticeable.

5.18.

The OpAmp is uncompensated (*A0=100dB*, A_{min}*=40dB*, f_0*=10kHz*).

a) Compensate the stage to ensure *PM=90°*.

b) Substitute the OpAmp with a Norton having A_i*=1* and determine the voltage gain.



5.18.a

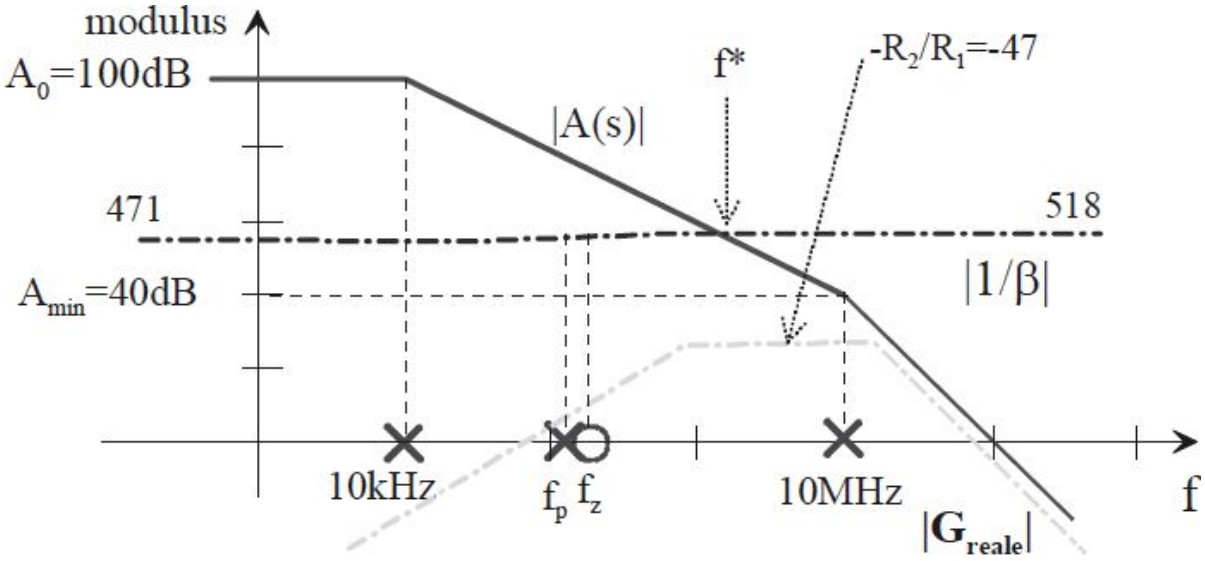
To evaluate the stage stability we need to calculate the β, to verify next the intersectio angel between *A(s)* and *1/β*. Applying a test signal to the amplifier output we can see the corresponding signal on the negative pin equal to:

$$\beta = \frac{1 + sC \cdot 10k\Omega}{1 + sC \cdot (10k\Omega + 470k\Omega)}$$

It is noticeable that setting the singolarities at low frequency, *1/β* maximum is 48. Then the intersection with *A(s)*, that has A_{min}*=40dB*, will happen always with an angle of *40dB/dec*. To stabilize the stage we need to increase the *1/β* value at high frequency over *40dB*, and then choose a value for *C* that move the singularities at low enough frequency.

We can act on the *470kΩ* resistance; however, in this way, the ideal transfer will be modified. A simpler way without modification is to insert a network that works on the virtual ground node and ground, or between the virtual ground node and the positive input pin. In fact, in both cases this new network will play no role in the ideal gain.

For example, with a resistance of *1kΩ* between ground and virtual ground, *1/β* at high frequency is *(10kΩ*//1k*Ω*+470k*Ω*(10k*Ω*//1k*Ω)≅518*, while at low frequency is *(1kΩ+470kΩ)/1kΩ≅471*· The singularities will be very close, with the zero moving near the pole, almost nulling it. By placing the pole a decade before the intersection between the new *1/β* and *A(s)*, i.e. at $A_0 \cdot f_0/518 = 1.9MHz$, we can be sure that the *PM* will be greater than 90°. We obtain *C=1/(2π·190kHz·10kΩ)=83pF*. The Bode diagrams are shown in figure.



It is noticeable that this kind of compensation flatten the G_{loop} also at low frequency. Propose a compensation that raises the *1/β* at high frequency, preserving high G_{loop} at low frequency in the uncompensated configuration.

5.18.b

The Norton amplifier is a differential current amplifier. When the feedback is good and with ideal input pin, ally the following properties:

- low impedance inputs (because they are current reader);
- the input pins have the same voltage (low impedance towards a common reference);
- if the current gain will be infinite, the incoming current would be the same, because the feedback would cancel the current error signal.

These relationships are dual in respect to those for VOA. However the weak current gain (in this case equal to 1) makes the third relationship practically not applicable and we have to apply the current balance to the nodes:

$$\begin{cases} I_+ = 0 \\ I_- = I_{OUT} + \dfrac{V_{IN}}{Z_{in}(s)} \\ I_{OUT} = (I_+ - I_-) \cdot A_i \\ V_{OUT} = I_{OUT} \cdot 470k\Omega \end{cases}$$

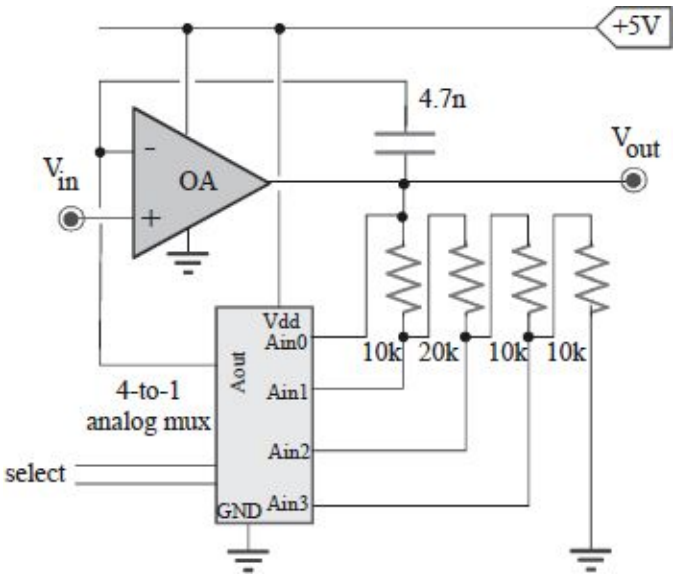
Eventually we obtain $\dfrac{V_{OUT}}{V_{IN}} = -\dfrac{A_i}{1+A_i} \cdot \dfrac{470k\Omega}{Z_{in}(s)} = -\dfrac{A_i}{1+A_i} \cdot \dfrac{sC \cdot 470k\Omega}{1+sC \cdot 10k\Omega}$.

5.19.

OpAmp with A_0=100dB, GBWP=100MHz, 20nV/√Hz e 5pA/√Hz. Analog mux with R_{on}≤100Ω, R_{off}≥5MΩ, I_{leak}<200nA.

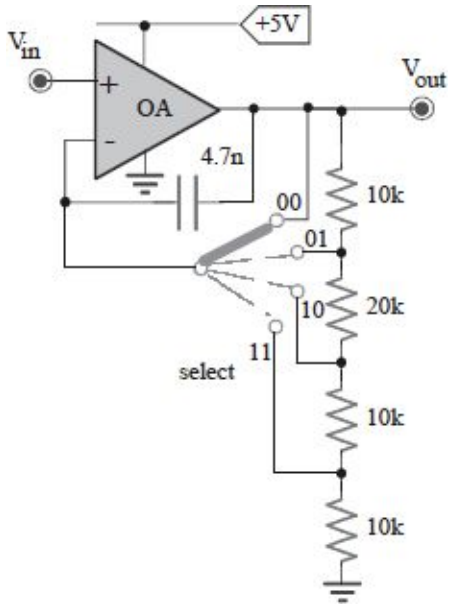
a) Explain the behavior of the stage and calculate the bandwidth varying the selection.

b) Calculate the output RMS noise with selection equal to *00* and *11*, taking into account all the noise sources.



5.19.a

The proposed stage is a non-inverting amplifier with gain and singularities depending on the mux selection. Thanks to the mux it is possible to change the resistance in parallel to the *4.7nF* capacitance; in fact when a mux input is selected a resistance is in parallel to the others. Substantially the stage is equivalent to the following:



It is noticeable that regardless the resistances values the high frequency gain will be unitary. The poles and gains in the 4 configurations will be:

$$00 \rightarrow \begin{cases} G = 1 + \dfrac{R_2}{R_1} = 1 + \dfrac{0}{50k\Omega} = 1 \\ f_p = \dfrac{1}{2\pi R_2 C} = \dfrac{1}{2\pi \cdot 0 \cdot C} = \infty \end{cases}$$

$$10 \rightarrow \begin{cases} G = 1 + \dfrac{R_2}{R_1} = 1 + \dfrac{30k\Omega}{20k\Omega} = 2.5 \\ f_p = \dfrac{1}{2\pi R_2 C} = \dfrac{1}{2\pi \cdot 30k\Omega \cdot C} = 1.13kHz \end{cases}$$

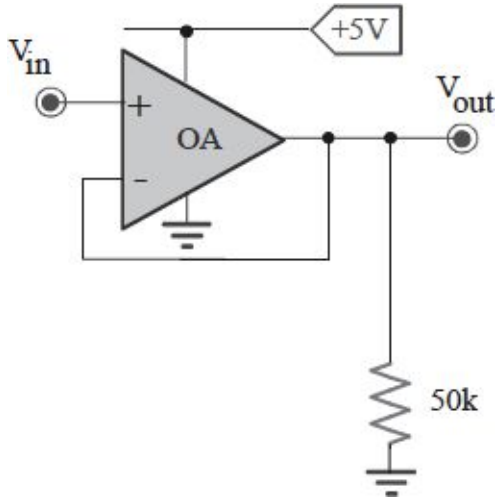
$$01 \rightarrow \begin{cases} G = 1 + \dfrac{R_2}{R_1} = 1 + \dfrac{10k\Omega}{40k\Omega} = 1.25 \\ f_p = \dfrac{1}{2\pi R_2 C} = \dfrac{1}{2\pi \cdot 10k\Omega \cdot C} = 3.38kHz \end{cases}$$

$$11 \rightarrow \begin{cases} G = 1 + \dfrac{R_2}{R_1} = 1 + \dfrac{40k\Omega}{10k\Omega} = 5 \\ f_p = \dfrac{1}{2\pi R_2 C} = \dfrac{1}{2\pi \cdot 40k\Omega \cdot C} = 846Hz \end{cases}$$

The zero frequency is easily obtainable using the gain bandwidth product between pole and zero, beacuse the high frequency gain is unitary. To simplify the calculations the mux resistances R_{on} and R_{off} were neglected.

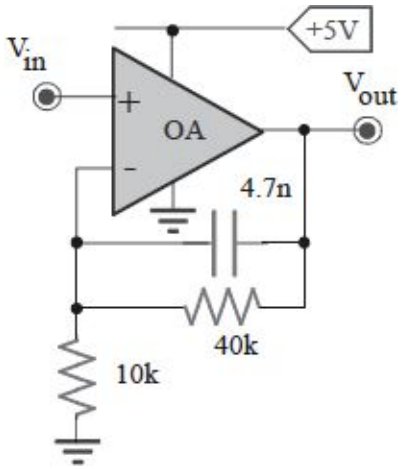
5.19.b

For the 00 configuration, the stage is the following:



Because the feedback is a short circuit, the transfer through the output due to the current source and R_1 resistance are zero. Indeed the noise voltage source is unitary, and the pole is the OpAmp GBWP. Therefore the output RMS noise is $v_{out}=\sqrt{e_n^2}\cdot GBWP\cdot\pi/2=\sqrt{(20nV/\sqrt{Hz})^2}\cdot 100MHz\cdot\pi/2=250\mu V$.

For the 11 configuration, the stage is:



In this case all noise sources will cause a response, apart from the current generator of the positive input, that is short circuited to ground by V_{in}. The current noise source for the negative input pin will be transferred through the

feedback resistance with the configuration pole. Thus the RMS value is $v_i^2=i_n^2\cdot(40k\Omega)^2\cdot f_p\cdot\pi/2=(7.2\mu V)^2$. The R_1 noise will be transferred with an inverting approximated integrator with a frequency pole equal to f_p. Thus the RMS value is $v_{R1}^2=4kTR_1\cdot(40k\Omega/10k\Omega)^2\cdot f_p\cdot\pi/2=(1.8\mu V)^2$. The R_2 noise is on the output with the f_p pole and its RMS value is $v_{R2}^2=4kTR_2\cdot f_p\cdot\pi/2=(0.9\mu V)^2$.

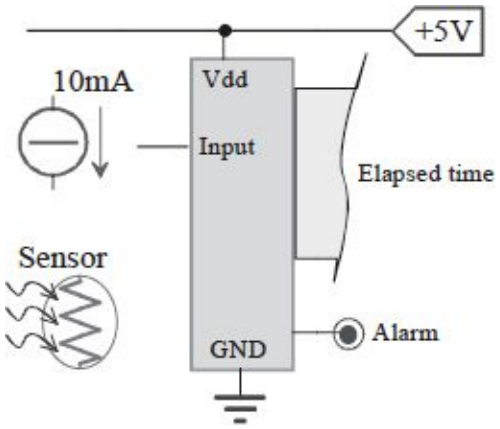
The voltage noise source is multiplied by the non-inverting tranfer with gain equal to 5 at low frequency and 1 till the GBWP (in fact through the positive pin there is a zero). Because the pole has a frequency lower than the GBWP, we can round down the output noise assuming a constant gain till the GBWP neglecting the frequency range in which the gain is 5. We obtain $v_e^2=e_n^2\cdot GBWP\cdot\pi/2=(250\mu V)^2$.

5.20.

The circuit measures the elapsed time between the threshold crossing over *50°* and below *25°*. The thermo-resistance is *50Ω+1Ω/°C*.

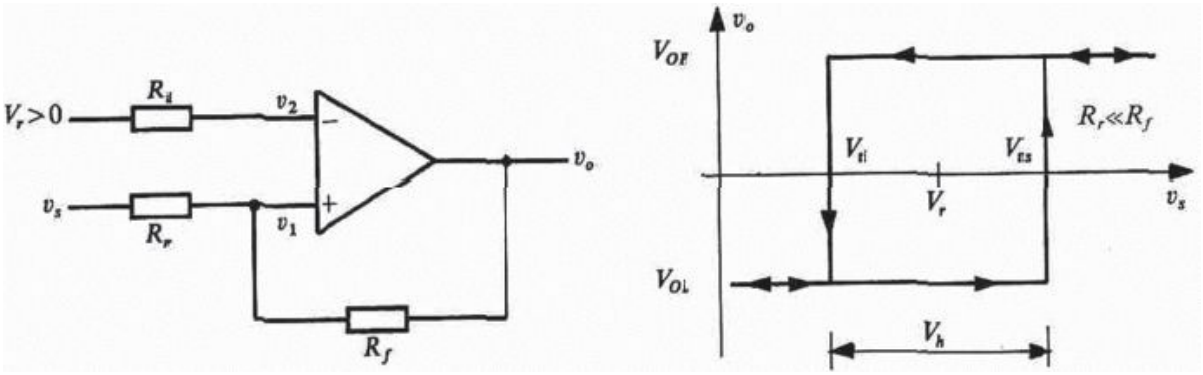
a) Design the analog stage to detect the threshold crossing.

b) Design the digital circuit able to count the elapsed time, up to *15minuti* with a risolution of *1s*. when the time is over a quarter of hour a signal must be high.



5.20.a

Since the circuit must assert the output high when the temperature exceed 50°C and return low when the temperature drops below 25°C, it is possible to use a non-inverting Schmitt trigger. Threshold voltages are $V_{TL}=10mA\cdot(50\Omega+1\Omega/°C\cdot25°C)=750mV$ and $V_{TH}=10mA\cdot(50\Omega+1\Omega/°C\cdot50°C)=1V$. Here is the static characteristic of the trigger.



The two upper and lower switching threshold are given by the expressions:

$$V_{TH} = \frac{R_f + R_r}{R_f} V_r - \frac{R_r}{R_f} V_{OL} \qquad V_{TL} = \frac{R_f + R_r}{R_f} V_r - \frac{R_r}{R_f} V_{OH}$$

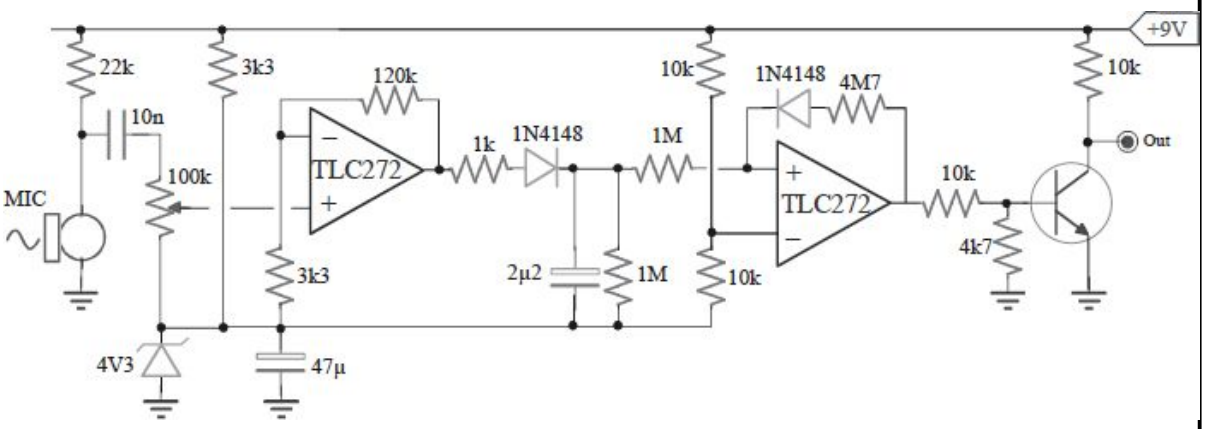
To proceeding to the sizing, please note that the difference between the second terms determine the hysteresis. In this case $V_{OL}=0$, $V_{OH}=5V$, $V_{isteresi}=250mV$, so we can choose $R_f=200k\Omega$ and $R_r=10k\Omega$. Since the highest threshold is $1V$, the reference voltage must be $V_r=950mV$.

5.20.b

The required circuit is simply a clock with a counter. The clock is enabled when the output trigger is high, and can be achieved using a relaxation oscillator with a period of *1s*. the counter will be reset on the trigger positive edge, through a high-pass network, and counts the clock pulses. Since it must count up to 15 minutes (i.e. 900 seconds), it must have at least *10bit*. Through a digital network it is possible to verify when the counts reach *900*. This signal can be used to set a flip-flop and trigger the alarm.

5.21.

a) Explain the circuit behavior (note OpAmps are rail-to-rail).
b) With the trimmer halfway and an input voltage with a frequency of 5kHz, determine the voltage level which saturates the transistor.



5.21.a

The stage is an amplifier, followed by a peak stretcher and a strange Schmitt trigger. The input signal, after the AC decoupling, is amplified by the first stage with a gain of *37/2=18.5*. the peak stretcher with a long time constant “stores” the signal peak and provides it to the trigger. The latter presents a threshold for the rising signal equal to *(9V-4.3V)/2+4.3V=6.65V*, because the diode is disabled with the low output, while the threshold for decreasing signal is lower, since with high output the diode is on. Assuming the peak stretcher be a voltage source, the lowest threshold is given by:

$$V_{TL} = V_R \cdot \left(1 + \frac{R_s}{R_f}\right) - \frac{R_s}{R_f} \cdot V_{OH} = 6.65V \cdot \left(1 + \frac{1M\Omega}{4.7M\Omega}\right) - \frac{1M\Omega}{4.7M\Omega} \cdot 9V = 6.15V$$

Substantially, if the amplified peak exceeds the high threshold, the BJT is turned on. The circuit remains in this condition until the capacitance of the peak stretcher has not been discharged below the lowest threshold. At this point, the BJT is turned off.

5.21.b

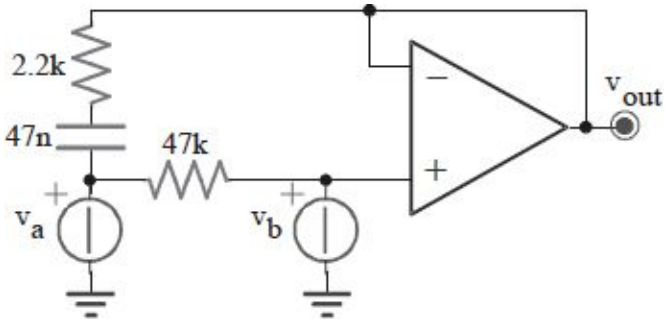
Starting from the output, to saturate the transistor the trigger output must be high, than the voltage on the output of the peak stretcher must be over *6.65V*. Since the first stage has a gain of 37, and all is referred to *4.3V* (zener diode), the signal on the positive OpAmp pin must have a peak amplitude of *(6.65V-4.3V)/37=64mV.* At *5kHz* the impedance of the decoupling capacitance is negligible compared to the potentiometer. If it is halfway, the required input signal must have a peak of *64mV·2=128mV.*

5.22.

The OpAmp is **non** compensated and has A_0=*100dB*, f_{high}=*5MHz*, A_{min}=*20dB*.

a) Calculate the ideal gain and the input impedance from v_a and v_b, when the other is removed.

b) Estimate the stability for both inputs and compensate, if needed.



5.22.a

Notice that the exercise requires to study the circuit when the other voltage source is *removed*, not *disabled*. Thus, we won't use the super position principle and the circuit topology in the two cases is different.

Start by removing v_b and calculate the ideal transfer for v_a. With ideal feedback and since the *47kΩ* does not conduct current, the two OpAmp inputs have the same voltage. Thus in the network between them there is no current flow and it is irrelevant for the gain. The circuit is then a buffer. Regarding the input impedance, for the bootstrap effect due to the network between the input pins, it is equal to the open-loop impedance *multipled* for the G_{loop}, ie $Z_{in}= (1/s\cdot 47nF+2.2k\Omega)\cdot G_{loop}(s)$.

Remove now the v_a and calcualte the ideal transfer for v_b. In this case, with ideal feedback, the OpAmp inputs have the same voltage. Thus in the network bettween the two inputs there is no current so it is irrelevant for the gain. The circuit is a bufferas before. Regarding the input imedance, because of the bootstrap effect of the netwrok between the input pins, it is equal to the open-loop one *multiplied* for G_{loop}, ie $Z_{in}=(1/s\cdot 47nF+2.2k\Omega+47k\Omega)\cdot G_{loop}(s)$.

5.22.b

Since the topology is different for the two cases, the singularities are different. In the first case it is easy to verify that $\beta=1$ for all frequencies, in fact *removing* v_b e *disabling* v_a the $47k\Omega$ input resistance is connected to ground. Thus the feedback is purely negative, because the OpAmp positive pin is fixed to ground independently to the applied signal. Because the intersection between $1/\beta$ and $A(s)$ has an angle of *40dB/dec* the stage is unstable. Also in the second case it is easy to verify that always $\beta=1$, in fact *removing* v_a and *disabling* v_b the positive input is connect to ground. Therefore the feedback is purely negative. Because the intersection between $1/\beta$ and $A(s)$ has an angle of *40dB/dec* the circuit is unstable.

The two circuits are topologically differet, but practically equal, because of the strange connections, the same fot both cases (for example the $47k\Omega$ resistance is unuseful).

To compensate the stage you must increase $1/\beta$ at high frequency, for examole moving the $47k\Omega$, interponating it between the signal source and the positive input pin. In this way the positive pin is not connected to ground but can change its voltage. This is useful because the $\beta=1$ at low frequency, while at high frequency $\beta=2.2k\Omega/(2.2k\Omega+47k\Omega)=0.045$, with a pole with a frequency $f_p=1/(2\pi\cdot47nF\cdot49.2k\Omega)=68Hz$ and zero with a frequency $f_z=1/(2\pi\cdot47nF\cdot2.2k\Omega)=1.5kHz$. Because $1/\beta(\infty)>A_{min}$ and the singularities of $1/\beta$ have a very low frequency, the intersection occurs with an angle of *20dB/dec*, and the stage is stable.

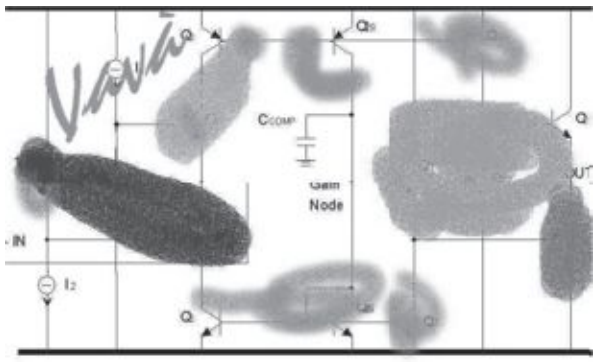
5.23.

Valentina bungled a CFA with $m=5$ pencils.

a) redraw the **corrected** internal scheme for the CFA, simplifying with blocks and calculating the expression for $G_{loop}(s)$, $G_{loop}(0)$ and for the closed loop pole, for **low gain**.

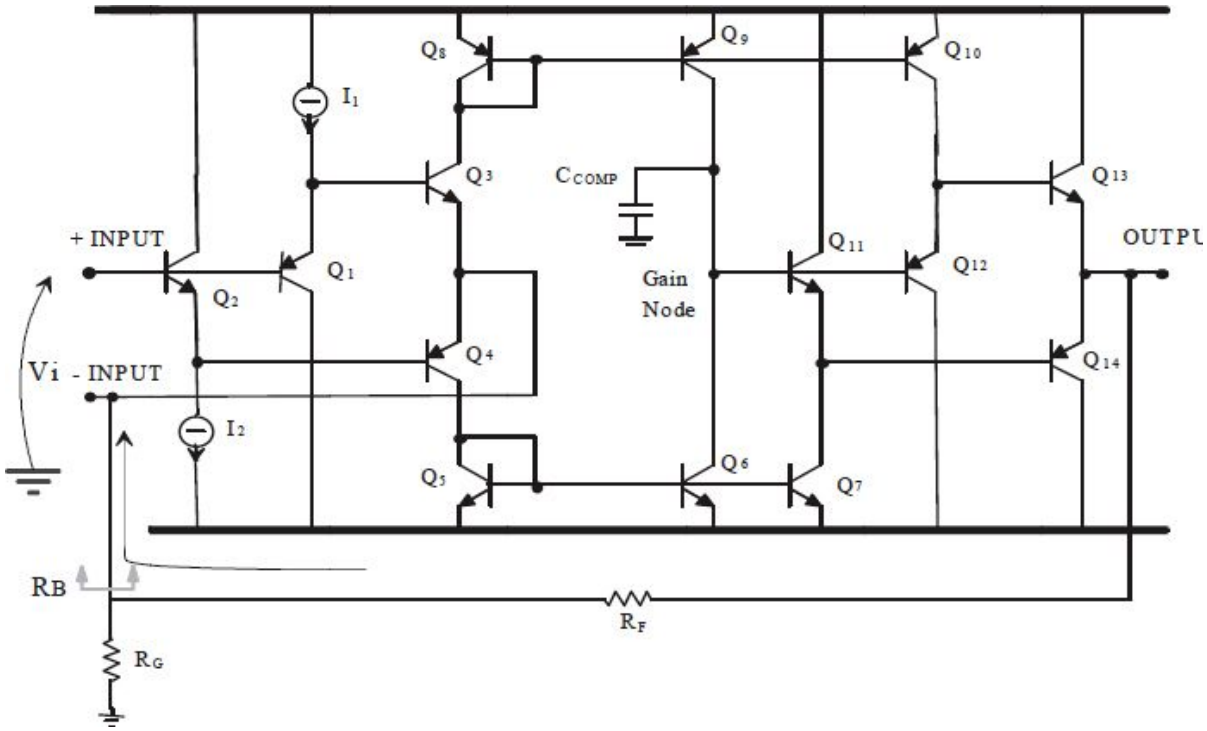
b) Design one stage with a gain of $+1$ and one with a gain of -1, usign **CFA**.

c) Design a stage with a gain of +1, using a **Norton Amplifier** with $A_i=m$.

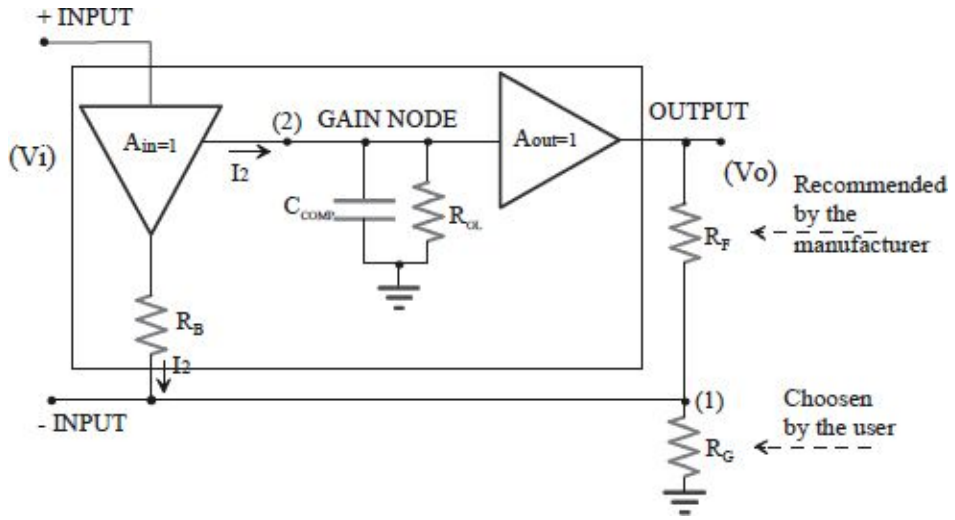


5.23.a

The correct internal scheme for the CFA is the following.



To compute the required analytical expressions it is easier use the simplified block diagram. In principle, the CFA consists of a voltage buffer (that performs the comparison of current), a current mirror (on a high-impedance node) and an output buffer.



To calculate the G_{loop} you need to break off the loop, for example between the mirror and the high-impedance node, and inject a test signal i_{test}. It is easy to verify that G_{loop} is:

$$G_{loop} = \frac{i_{feedback}}{i_{test}} = -\left(R_O \,//\, \frac{1}{sC_{comp}} \right) \cdot \frac{1}{R_F + (R_G \,//\, R_B)} \cdot \frac{R_G}{R_G + R_B}$$

For low gain is $R_G >> R_B$ and the feedback current enters into the CFA negative pin. The G_{loop} is independent in the stage gain (determined by R_G), and has constant bandwidth. In this case:

$$G_{loop} \cong -\frac{R_O}{1 + sR_OC_{comp}} \cdot \frac{1}{R_F + R_B}$$

Because $R_F >> R_B$ we have:

$$G_{loop} \cong -\frac{R_O}{1 + sR_OC_{comp}} \cdot \frac{1}{R_F} = -\frac{R_O / R_F}{1 + sR_OC_{comp}}$$

It is possible to compute the closed loop pole imposing $G_{loop}(s)=1$, and:

$$polo_{chiuso} \cong polo_{aperto} \cdot (1 - G_{loop}(0)) \cong \frac{1}{2\pi R_O C_{comp}} \cdot \left. \frac{R_O \big/ R_F}{1 + sR_O C_{comp}} \right|_{s=0} = \frac{1}{2\pi R_F C_{comp}}$$

As noticed, the closed-loop pole is independent of the stage gain, provided you only use the R_G to change its value.

5.23.b

To size the required stages, you should bear in mind the operation of a CFA. When a CFA has a good feedback and the inputs are ideal, we have the following properties:

- the non-inverting input has high impedance (because is a buffer input);
- the inputs have the same voltage (because they are input and output of the same buffer);
- the signal current flowing in the inverting input is zero, because it is equal to the error signal the feedback try to cancel.

Note that these input properties are the same for a VOA with a good feedback. Therefore it is possible to use the same procedure to obtain the ideal gain. In a non-inverting stage it will be $G_{ID}=1+R_F/R_G$, while in a inverting stage the gain will be $G_{ID}=-R_F/R_G$. However, note that, though formally identical, they are based on assumptions quite different in the two cases. For example, the negative terminal of a VOA has the same voltage of the positive same only in the presence of well-functioning feedback, because it is precisely the reaction that requires such equality, while in a CFA that is always true, because it is a characteristic of the amplifier .

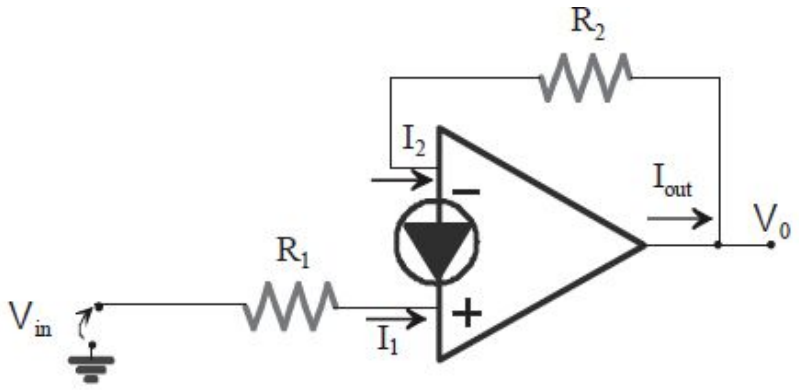
The non-inverting buffer reaction must have a resistance and not just a short: it was usually necessary to avoid instability, given the particular topology of the internal CFA. An appropriate value is given in the datasheet.

5.23.c

The Norton amplifier is a differential current amplifier. When has a good feedback and inputs are ideals, apply the following properties:

- the inputs are low impedance (they are current readers);
- the two inputs are at the same voltage (because the inputs are low impedance to a common reference);
- the inward current in the two terminals tend to be the same, because it is the error signal that the reaction tends to cancel if the current gain was infinite.

These properties are the dual compared to VOA. However, the low current gain (in this case equal to 5) makes the third report not applicable in practice and must then use the balance of power at the nodes. For example in the figure for the following stage we have:



$$\left(\frac{V_{IN}}{R_1}-\frac{V_{OUT}}{R_2}\right)\cdot A_i = i_{out} = \frac{V_{OUT}}{R_2} \qquad \text{from which:} \qquad \frac{V_{OUT}}{V_{IN}} = \frac{R_2}{R_1}\left(\frac{A_i}{1+A_i}\right) = \frac{R_2}{R_1}\cdot \frac{5}{6}$$

To have a unity gain is then simply choose $R_2/R_1=6/5$.

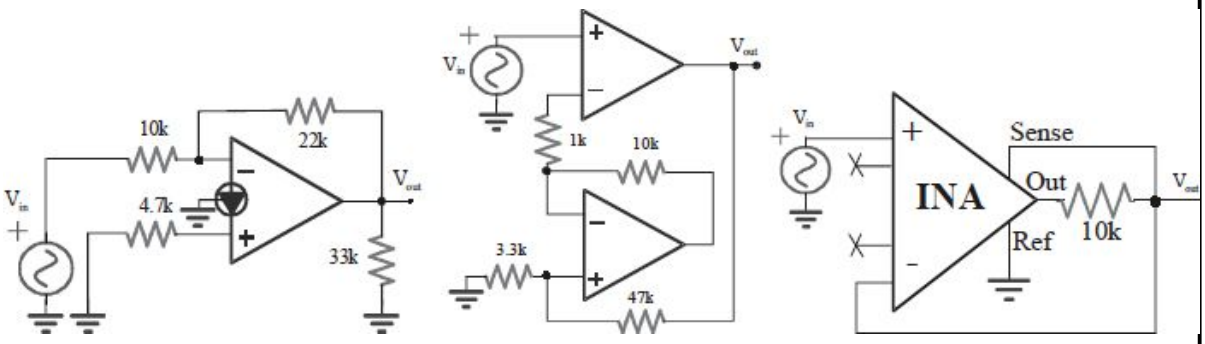
5.24.

Given the following blocks, employing INA, CFA and amplifiers Norton:

a) Calculate the gain V_{out}/V_{in} for INA noting the absence of the R_G and evaluate roughly the output impedance.

b) Calculate the gain V_{out}/V_{in} for Norton stage with Ai = 4 and evaluate the input impedance.

c) Calculate the gain V_{out}/V_{in} for CFA stage.



5.24.a

To calculate the gain of the INA is sufficient to apply its input-output relationship $V_{out}=(V_+-V_-)\cdot G$, with $G = 1$ since it was R_G. Note that the $10k\Omega$ resistor is irrelevant because the INA *Sense* signal is picked up after it. Then:

$$\begin{cases} V_{out} = (V_{in} - V_-)\cdot 1 \\ V_{out} = V_- \end{cases}$$

where $V_{out}=V_{in}/2$.

To evaluate the output impedance should be noted that it is a node of the loop, so its impedance is reduced by the reaction: would be to say that it tends to be zero. In fact in this particular circuit G_{loop} is small, equal to -1 (try to cut the ring to the negative terminal of the INA). So the open-loop impedance is indeed reduced, but only by a factor of two. It is therefore $5k\Omega$.

5.24.b

The Norton amplifier is a differential current amplifier. When has a good feedback and inputs are ideals, apply the following properties:

- the inputs are low impedance (because are current readers);
- the two inputs are at the same voltage (because the inputs are low impedance to a common reference);
- the inward current in the two terminals tend to be the same, because it is the error signal that the reaction tends to cancel if the current gain was infinite.

These properties are the dual compared to VOA. However, the low current gain (in this case equal to 5) makes the third report not applicable in practice and must then use the balance of power at the nodes:

$$\begin{cases} I_+ = 0 \\ I_- = \dfrac{V_{IN}}{10k\Omega} + I_{OUT} - \dfrac{V_{OUT}}{33k\Omega} \\ I_{OUT} = (I_+ - I_-) \cdot A_i \\ V_{OUT} = I_{OUT} \cdot 33k\Omega \,//\, 22k\Omega \end{cases}$$

Doing the calculations we have:

$$\frac{V_{OUT}}{V_{IN}} = -\frac{A_i}{10k\Omega} \cdot \frac{33k\Omega \,//\, 22k\Omega}{1 + A_i\left(1 - \dfrac{22k\Omega}{33k\Omega + 22k\Omega}\right)} = -1.55$$

If we had considered the relationship ($A_i = +\infty$), we would obtain $V_{OUT}/V_{IN} = -22k\Omega/10k\Omega = -2.2$.

The input impedance is equal to $10k\Omega$ independently on the feedback, because of the low impedance input for the Norton amplifier.

5.24.c

To evaluate the gain of the stage, keep in mind the particular behavior of a CFA. When a CFA has a good feedback and inputs are ideal, apply the following properties:

- the non-inverting input has high impedance (because the buffer input)
- the two inputs has the same voltage (because they are the input and output of the same buffer)
- the current in the negative terminal of the input signal is zero, because it is the error signal that the reaction tends to cancel.

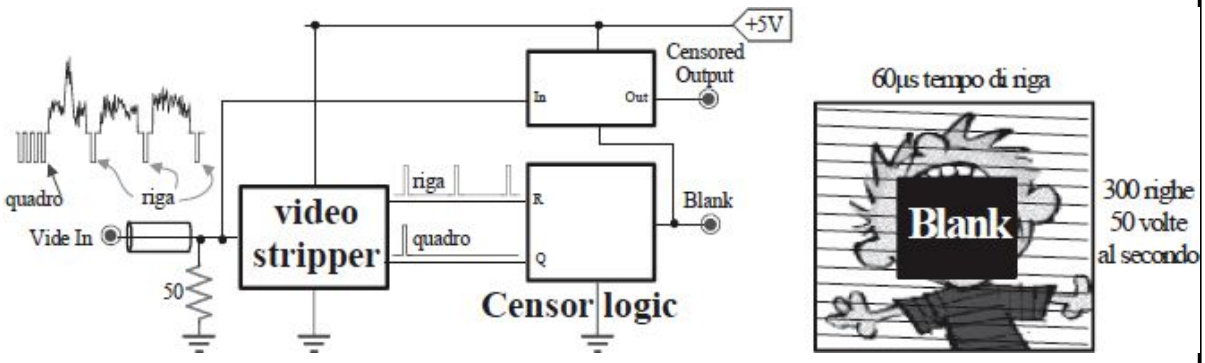
Applying these properties it is concluded that the input terminals of the two CFA have all the same potential (in fact in the $1k\Omega$ resistor current can not flow). So the stage composed of two CFA is equivalent to a single VOA (in terms of the calculations to be done to assess the gain). The equivalent VOA shows a non-inverting configuration, so the gain is $1+47k\Omega/3.3k\Omega=15.24$.

5.25.

The Electronic Censurer should obscure the central part (*1/9* of the entire screen) of a PAL video signal (ranging from white, *0V*, and black, *0.6V*). The frame sync (*50Hz*) and row (approximately every 60µs) are extracted form an integrate VideoStripper, which recognizes them in the signal (as equal to *–0.6V*).

a) Design the circuit that should create the digital signal BLANK at the area to be obscured.

b) Design the analog circuit which maintains the output to a fixed V_{ref} when there is the BLANK command. The level V_{ref} will be settable with a trimmer, choosing the desired brightness between white and black (*0÷0.6V*).



5.25.a

We want to make a circuit that, starting from the sync signals (row and frame) given by the *video stripper*, generates an obscuring logic signal for the 1/9 central part of the image. Substantially the first 20ms/60µs/3=100 rows can pass, the central part for the second 100 rows will be obscured, and then again the last 100 rows can pass.

Using the provided sync, it is easy to use the frame sync to reset the whole circuit, and the row one to count the elapsed rows with a simple binary counter. An appropriate combinational network identifies the present row and decides if the row should pass or be obscured. Where the row will be partially obscured, a latch activated by the sync signal produces the 20µs obscuring signal with a delay of 20µs.

A set-reset flip-flop will be set between the rows 100 and 200. Only when it is set two latches can work, triggered on the row sync positive edge. The AND among the latches output signals gives a high signal exactly at the central part of the 100 central rows. The incoming frame sync acts on the reset asynchronous from the flip-flop and counters, bringing back the circuit to the initial state.

5.25.b

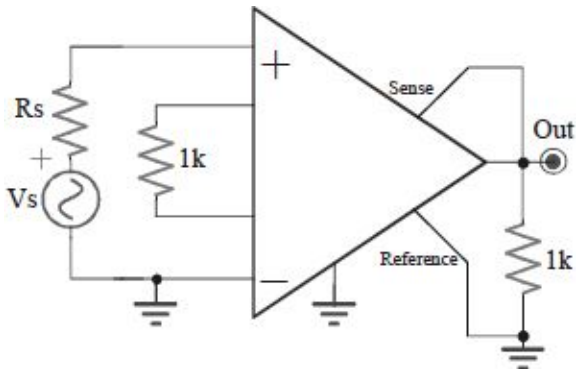
To obscure the video signal we need to give out the same video signal or a fixed voltage level. This is possible interposing between input and video output a MOS which works as switch controlled by the *Blank* signal. When the *Blank* signal is high the MOS is open. At the same time another MOS is closed, connecting the video output to a voltage buffer output. The voltage level can be set from $0V$ up to $0.6V$ partitioning the voltage drop across a diode.

5.26.

The INA has internal resistances of 10kΩ and OpAmps with noise sources of *5pA/√Hz* and *3nV/√Hz*, *GBWP=5MHz* and A_0*=100dB*. The source has *Rs=3kΩ* and amplitude of *100mVpp*.

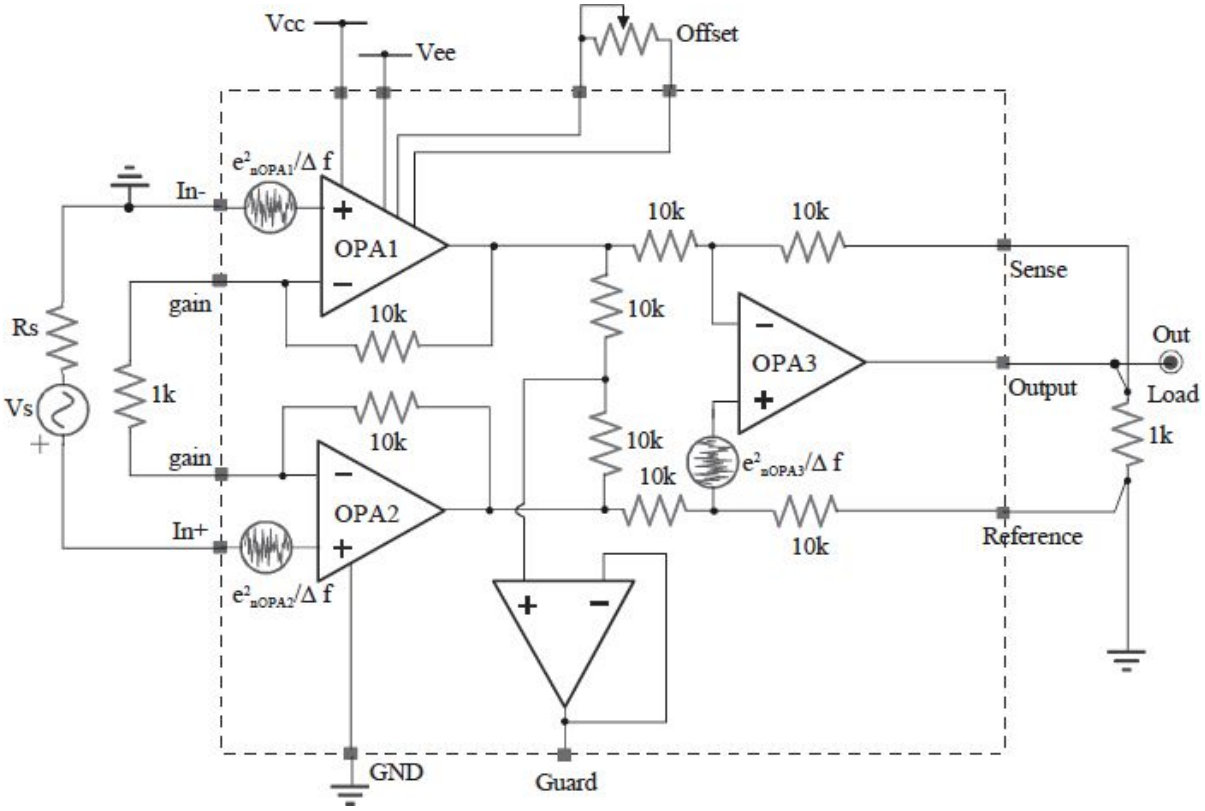
a) Taking in account **only voltage noises** of OpAmps, calculate the equivalent input voltage and the corresponding *NF*.

b) Taking in account **only current noises**, compute SNR_{out}.



5.26.a

Starting from the only voltage noises of INA OpAmps, it is useful to refer to the following schematic.



To calculate the voltage equivalent noise, it is enough to refer the OpAmp output noise to the input divided by the INA voltage gain:

$$G_{INA} = 1 + \frac{2R}{R_G} = 1 + \frac{2 \cdot 10k\Omega}{1k\Omega} = 21$$

Then analyze the noise contributions of the three OpAmps. The two input OpAmps have the own equivalent voltage noise already referred to input, and it is unuseful to bring them to output and then to input. The noise of the third OpAmp (conveniently connected to the non inverting terminal) is referred to output with a non-inverting stage with gain *1+R/R=2*.

The overall voltage input noise (and then the equivalent one) is equal to:

$$e_{n,IN}^2 = e_{n,OPA1}^2 + e_{n,OPA2}^2 + e_{n,OPA3}^2 \cdot (2/21)^2 =$$
$$= (3nV/\sqrt{Hz})^2 + (3nV/\sqrt{Hz})^2 + (3nV/\sqrt{Hz})^2 \cdot (2/21)^2 \cong$$
$$\cong 18nV^2/Hz \cong (4.24nV/\sqrt{Hz})^2$$

As expected the third OpAmp noise is highly negligible compared to the input one. To calculate the NF it is sufficient to apply the definition:

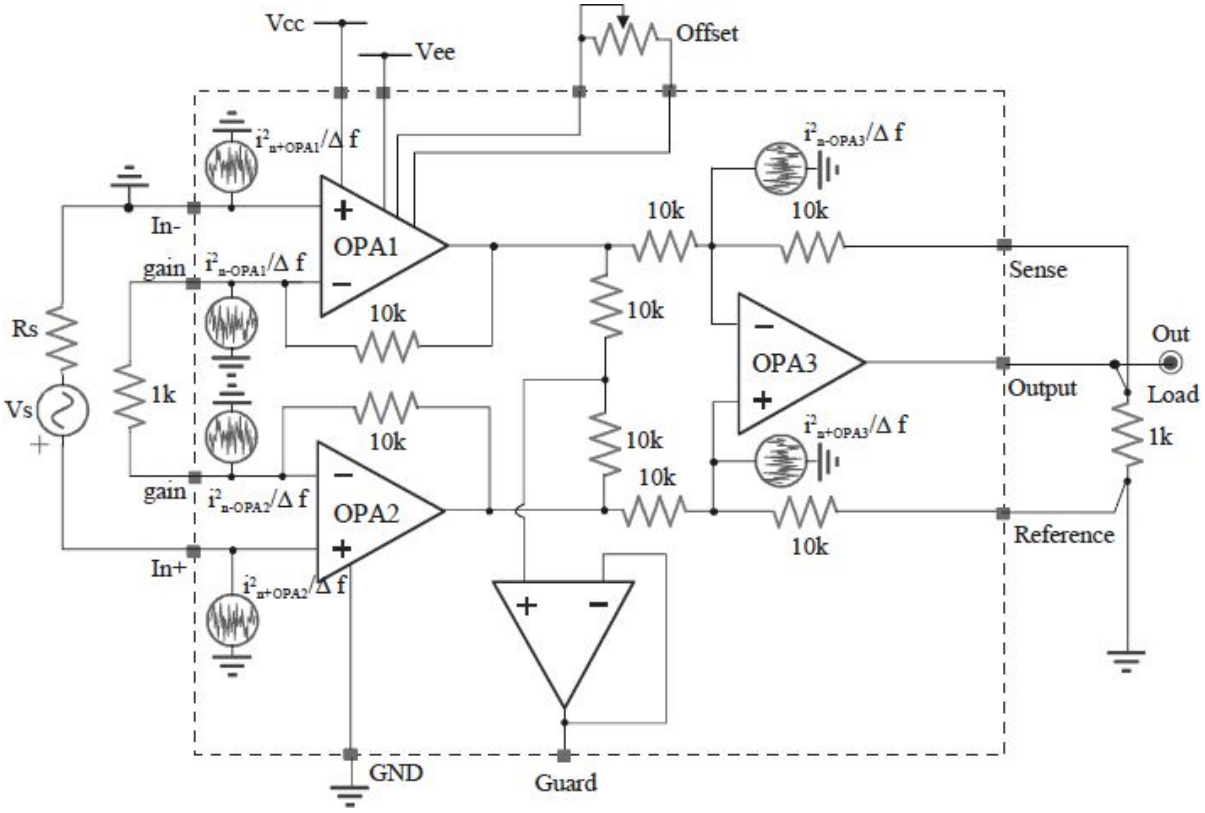
$$NF = \frac{TotalNoiseInput}{NoiseSource}$$

then:

$$NF = 10\log\left(\frac{e_{n,IN}^2 + 4kTR_S}{4kTR_S}\right.$$

5.26.b

When considering only the noise power is useful to consider the following circuit.



We can value the output voltage transfer for the various noise sources, because they may be compared with the output signal to compute the SNR_{out}.

The OpAmp2 contributes with its two current sources. The one connected to the positive terminal is multiplied by R_S and then is referred to the output multiplied by the INA voltage gain. The one connected to the negative terminal is multiplied by R (in fact through R_G there is no current flow). This noise is multiplied by the unitary gain of the differential amplifier to the INA output. Therefore:

$$e_{n,OUT2}^2 = i_n^2 \cdot R_S^2 \cdot G_{INA}^2 + i_n^2 \cdot R^2 \cdot 1$$

The OpAmp1 contributes with its two current sources. Nevertheless, the positive terminal one is shorted to ground and doesn't give noise to output. The other one (on the negative terminal) gives to OpAmp1 output a noise multiplied by R (in fact through R_G there is no current) and on the INA output is multiplied by the unitary gain of the differential amplifier. Therefore:

$$e_{n,OUT1}^2 = i_n^2 \cdot R^2 \cdot 1$$

The OpAmp3 contributes with all the noise current source. The source connected to the positive terminal is multiplied by $R//R=5k\Omega$ and then to the output by the non- inverting voltage gain of the OpAmp3 equal to 2. The source on the negative terminal is reffered to the OpAmp3 output multiplied by the feedback R. Therefore:

$$e_{n,OUT3}^2 = i_n^2 \cdot (R/2 \cdot 2)^2 + i_n^2 \cdot R^2$$

The output overall noise spectral density is equal to:

$$e_{n,OUT}^2 = i_n^2 \cdot R_S^2 \cdot G_{INA}^2 + i_n^2 \cdot R^2 \cdot 1 + i_n^2 \cdot R^2 \cdot 1 + i_n^2 \cdot (R/2 \cdot 2)^2 + i_n^2 \cdot R^2 + 4kTR_S \cdot G_{INA}^2 =$$
$$= (330nV/\sqrt{Hz})^2$$

To compute the RMS noise and then the SNR you need to know the noise equivalent bandwidth. Without reactive components the bandwidth will be limited by the GBWP of the OpAmp. On first approximation we can consider the system as a single pole, which is determined by the OpAmp that has the highest gain, then lowest bandwidth. Is is immediately to verify that both the input OpAmps have the closed-loop pole intervening before the pole of the third OpAmp. The signal bandwidth is equal to $GBPW \cdot \beta = 5MHz \cdot 500\Omega/(10k\Omega + 500\Omega) = 238kHz$, while the noise bandwidth is $f_n = \frac{\pi}{2} f_s = 373kHz$.

The SNR is:

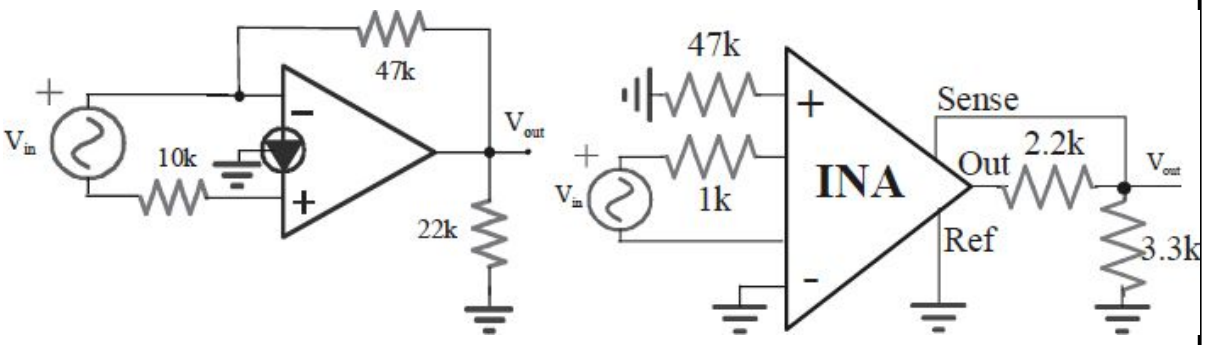
$$SNR_{out} = \frac{(V_p/\sqrt{2}) \cdot G_{INA}}{n_{rms}} = \frac{(50mV/\sqrt{2}) \cdot 21}{\sqrt{e_{n,OUT}^2 \cdot f_n}} \cong 71dB$$

5.27.

Given the two stages shown in figure:

a) Compute the V_{out}/V_{in} gain and the input impedance in the Norton configuration ($A_i=4$).

b) Compute the V_{out}/V_{in} and the impedance experienced by V_{in} in the INA stage (internal resistances of $20k\Omega$).



5.27.a

The Norton amplifier is a differential current amplifier. When has a good feedback and with ideal input apply the following properties:

- inputs have low impedance (because thy are current reader);
- the two inputs have the same potential (the inputs have low impedance through a common reference);
- the input terminals currents are the equals, because it is the error signal then the feedback leads to zero.

These relationships are dual in respect to those for VOA. However the weak current gain (in this case equal to 4) makes the third relationship practically not applicable and we have to apply the current balance to the nodes:

$$\begin{cases} I_+ = -\dfrac{V_{IN}}{10k\Omega} \\ I_- + I_+ = +I_{OUT} - \dfrac{V_{OUT}}{22k\Omega} \\ I_{OUT} = (I_+ - I_-)\cdot A_i \\ V_{OUT} = I_{OUT}\cdot 47k\Omega // 22k\Omega \end{cases}$$

Performing calculations we have:

$$\frac{V_{OUT}}{V_{IN}} = -\frac{2A_i}{10k\Omega}\cdot\frac{47k\Omega // 22k\Omega}{1+A_i\left(1-\dfrac{47k\Omega // 22k\Omega}{22k\Omega}\right)} = -5.28$$

If it had been considered the ideal reaction ($A_i=+\infty$), would have been obtained $V_{OUT}/V_{IN}=-2\cdot 47k\Omega/10k\Omega=-9.4$.

The input impedance of the stage is equal to $10k\Omega$ regardless of reaction, just because of the low input impedance of the Norton amplifier.

5.27.b

To compute the Ina stage gain, we need to keep in mind its internal structure. The INA is made of an input stage and a differntial amplifier. The latter has unitary gain, even in the presence of the OpAmp output resistance, thanks to the *sense* signal taken after tge resistance. To evaluate the gain of the input stage, we can only consider that the negative terminal for both the OpAmps are virtual ground nodes. Therefore in the input resistance the current $i_{in}=v_{in}/1k\Omega$ flows. Such current flows through the two feedback resistances, but with opposite directions in two different branches. Therefore the differential voltage read at the input of the differential amplifier is equal to $v_{dif}=-2i_{in}\cdot 20k\Omega = -2v_{in}/1k\Omega\cdot 20k\Omega = -40v_{in}$, and then $V_{out}/V_{in}=-40$.

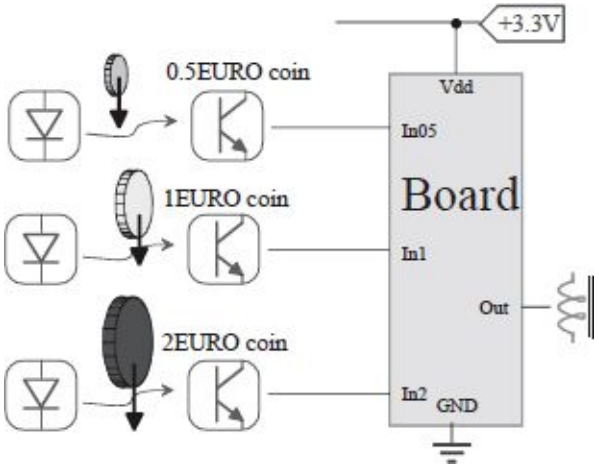
Because the OpAmp negative terminals are virtual ground nodes, the input impedance is simply equal to $1k\Omega$.

5.28.

Upon reaching *10€*, the circuit should enable the relay for *1sec* and then wait for other coins. Different coins can be inserted at the same time (there are three separate holes), while similar ones come in sequence.

a) Design the **fully analog** circuit that, starting from the phototransistor which prompt the insertion of single coins, closes the relay.

b) Reset the circuit if from the insertion of the first coin are elapsed more than *120s*.



5.28.a

One way to design the analog circuit is to use a capacity as a sort of container of money: for each insertion of a coin, a charge is injected in the capacity proportionally to the value of the coin, until reach a predetermined value corresponding to 10€. In essence, since coins of different denominations can be inserted at the same time, we need three current generators that can be made with a typical OpAmp configuration. In this way the generators can be turned on or off simply varying the voltage at the positive terminal of the OpAmp. If we assume that at the insertion of each coin each generator is turned on for the same time, the three generators must have currents in the ratio of 1,2,4. The ignition timing, the beginning of which is determined by the shutdown of the corresponding phototransistor, may be obtained by peak stretcher or high-pass filters, depending on the time during which coin obscures the phototransistor is less or greater of desired. Finally, a comparator detects when the voltage on the capacity has reached the threshold, triggering the reset of the capacity and the switching of the relay. Again the timing of *1s* can be obtained by peak stretcher.

5.28.b

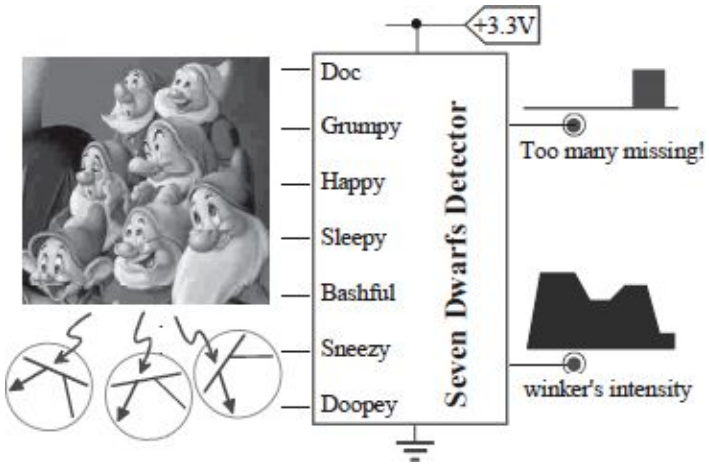
Again you may use the charging of a capacity with a constant current. In this case, however, the charge must continue without intermission from the inserting of the first coin until to the voltage indicates the maximum time of *120s*. The current source can be activated by monitoring the *"bank"* capacitance of the previous point; in fact as soon as you insert the first coin that capacity ceases to be discharged, and then has a voltage different from zero. That voltage, detected by a comparator, activates the current source of the *"timer"* capacitance. Another comparator can wait *120s* by simply reading the voltage across this capacitance, causing the reset of both the capacitances through MOSs that discharge them completely.

5.29.

Monitor the presence of the Seven Dwarfs by phototransistors.

a) Design the analog network that gives a high output signal for at least *20s* when three or more Dwarfs are missing.

b) Design the network "The Winker Takes All" which provides an output voltage proportional to the only dwarf who "winks" at higher frequency (variable between *0÷1Hz*).



5.29.a

Since every dwarf is associated with a phototransistor, it is reasonable to think of adding the collector currents of the seven phototransistors to generate a voltage proportional to the number of these Dwarfs. At this point you can use a comparator to determine if there are at least three dwarf missing. You can for example use a trans-impedance stage, where all the collectors of phototransistors are merge into the same virtual ground node. The output voltage will be equal $n \cdot I \cdot R$, where n is the number of these Dwarfs.

It is now necessary to generate a pulse of *20s* in response to the lack of at least three Dwarfs. Assuming that the lack of Dwarfs is longer than 20s, this can be done using a high-pass filter (with a very low frequency pole) after the output of the comparator. After this stage we can use another comparator. In fact, if you consider the step response of a high-pass in the time domain: it is clear that by using a time constant sufficiently long and the filter followed by a comparator with appropriate threshold, the circuit will respond with a pulse duration required to each rising edge of the first comparator.

The falling edge instead, as it will determine a negative pulse output of the filter, will not cause any switching of the comparator. However, if the lack of Dwarfs is less than *20s* the previous circuit doesn’t work properly. In fact the output of the filter doesn’t drop below the comparator threshold thanks to the exponential discharge of the RC network, but thanks to the sudden switching of the input filter. In essence, the circuit reduces the input pulse duration to *20s* if it is longer, but otherwise is not able to lengthen it. In this case it is necessary to use a proper monostable, for example made with a Schmitt trigger.

5.29.b

Assuming that the Dwarfs will wink in a sine wave and all with the same width, you can understand who winks at higher frequency by passing their signals through a high-pass filter with a pole frequency of *1Hz*. In this way, signals with lower frequencies will be progressively attenuated, and the output of the 7 filters will be sinusoids with amplitude proportional to frequency. At this point, using peak stretcher with these output signal, you get almost constant voltage signals, also with amplitude proportional to the frequency of blinking.

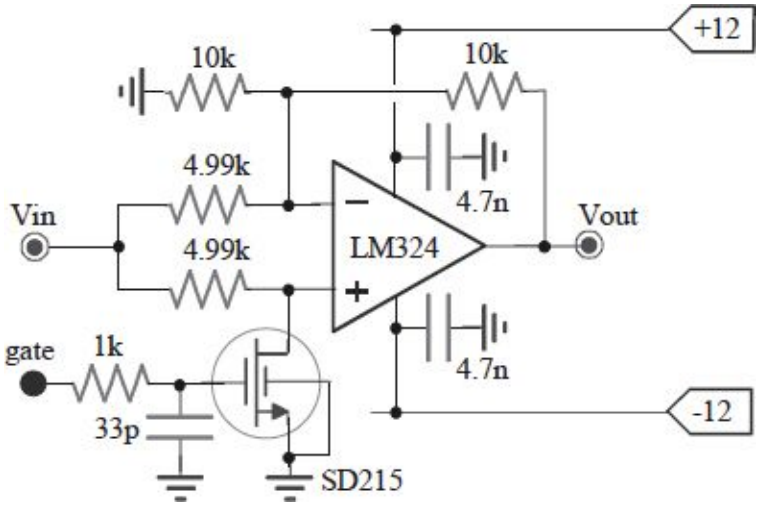
To determine which of the 7-levels voltage is greater it is sufficient read a high impedance (to avoid to load the peak stretchers) and then connect them to the same load resistance through seven diode. In this way only the greater signal will activate the diode and may impose its voltage on the load, while others are blocked by the diodes.

5.30.

In the following circuit, the command signal came from a HCMOS gate.

a) Explain the behavior of the circuit with sinusoidal input of *2Vpp*, computing the gain v_{out}/v_{in} and the input impedance, for both the *gate* level.

b) During the operation, the MOS will be damaged; explain the reason, propos a solution and foresee any other problems.

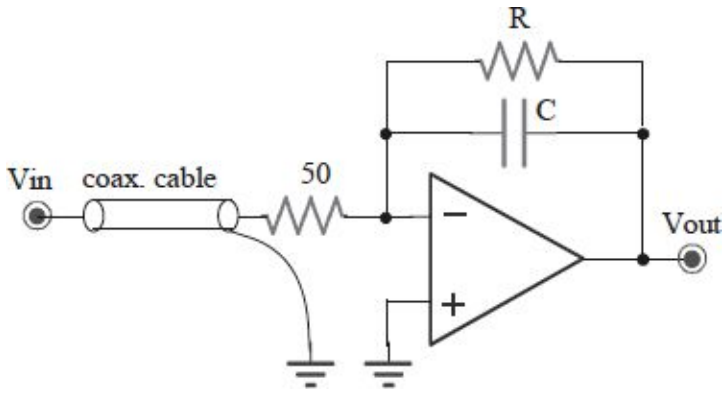


5.31.

The OpAmp is **non-compensated** and has $A_0=120dB$, $A_{min}=40dB$, $f_0=100Hz$, $I_B=3nA$, $SR=10V/\mu s$, $I_{outmax}=\pm 40mA$.

a) Given an input step, computes the maximum amplitude and size C in order that can be reached the OpAmp limit.

b) Calculate the min and max values for R for integrating from $f_{min}=1kHz$ with $C=100nF$.

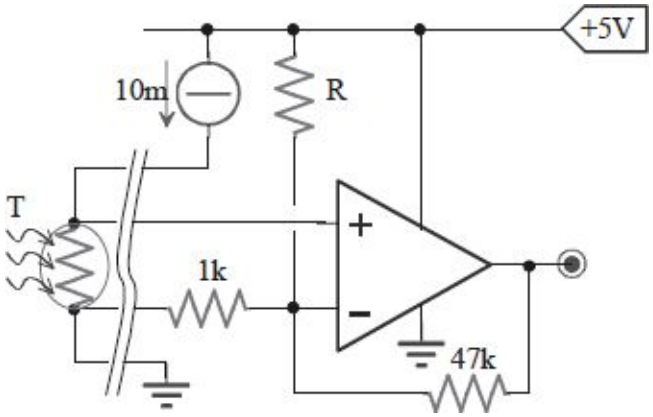


5.32.

The OpAmp has $A_0=100dB$, $CMRR=90dB$, $GBWP=5MHz$ and $I_B=500nA$. For $T=0\div100°C$ we want $V_{Out}=0\div4.7V$. The thermoresistance is $200\Omega+1\Omega/°C$.

a) Size R with commercial values, to have $V_{Out}=0$ @ $T=0°C$. For this value, compute the residual error for V_{Out}.

b) starting from the chosen value for R, compute V_{Out} for $T=0°C$ and $T=100°C$, explaining the reasons of the deviation of the ideal straight line with slope $4.7V/100°C=47mV/°C$ and suggest the possible solutions.

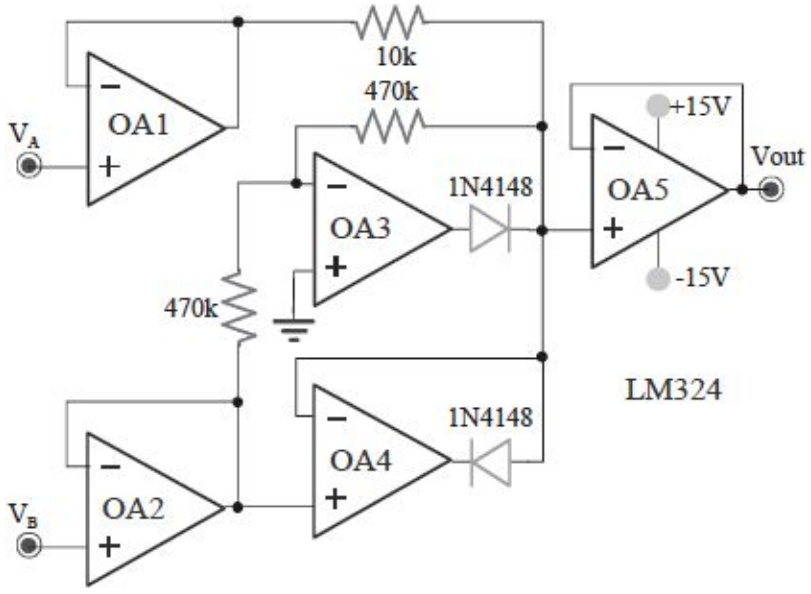


5.33.

The circuit use two quad-OpAmps LM324 (*GBWP=1MHz*, V_{OS}*=3mV*, I_B*=200nA*, I_{out}*=20mA*).

a) Explain the behavior when V_A is a sinusoidal input between ±*5V* and V_B is fixed and positive.

b) Determine the static error for V_{out} when V_A*=+2V* and V_B*=+5V*.

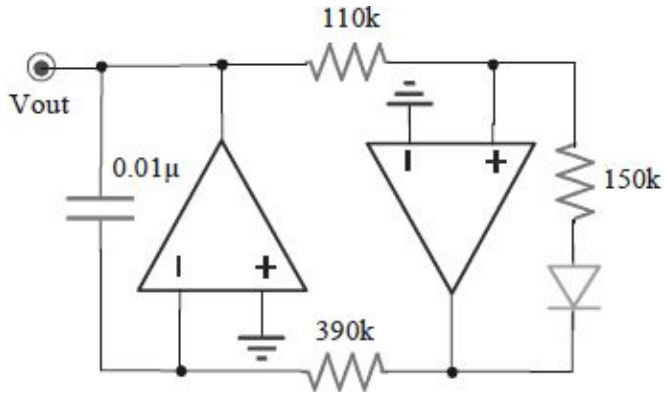


5.34.

The two LM323 are biased with ±5*V*.

a) Reckon the behavior of the circuit, drawing and quoting amplitudes and times of the output.

b) Explain how the behavior changes inverting the diode.

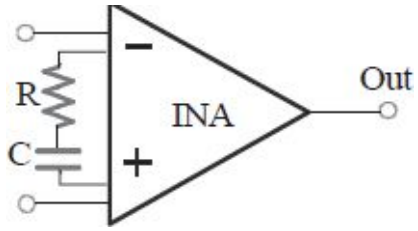


5.35.

The INA has all the internal resistance of *20kΩ*. Moreover *R=200Ω* and *C=47μF*.

a) Draw the internal structure and the quoted Bode diagram (with the analytical expression of poles and zeros introduced by R and C and for the DC gains, at medium and high frequency) of the modulus of the gain $V_{Out}(s)/V_{diff,\,In}(s)$.

b) Determines the minimum *GBWP* that the OpAmps of the INA must have, in order to have the cut-off frequency over *100kHz*.

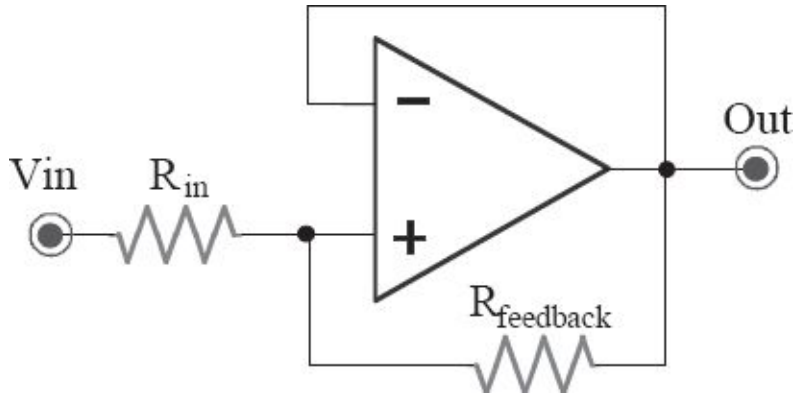


5.36.

The OpAmp is for video frequencies and is non-compensated (A_0=*110dB*, A_{min}=*30dB*, f_0=*1kHz*) and has input parasitic capacitances ($C_{in,Diff}$=*6pF* and $C_{in,CM}$=*4pF*).

a) Size the positive compensation, to assure a bandwidth of *30MHz* with *45°* of phase margin.

b) Because this compensation reduces the G_{loop} at low frequency, propose a simple change to avoid this effect.

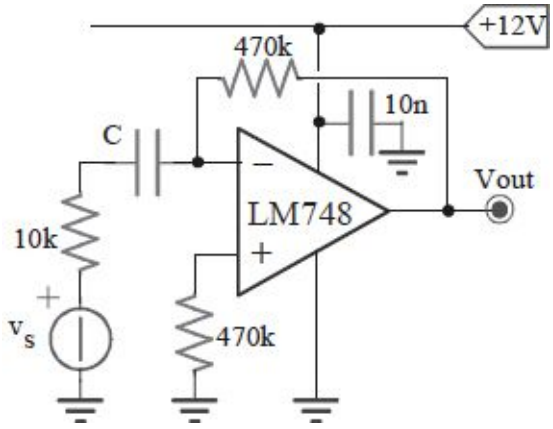


5.37.

The **non-compensated** OpAmp has A_0=100dB, f_{high}=1MHz, A_{min}=20dB.

a) Draw the Bode diagram that assures the stage stability, draw the frequency response and extract the closed-loop poles.

b) Size *C* to allow the passage of signals with f_{in}>10Hz.

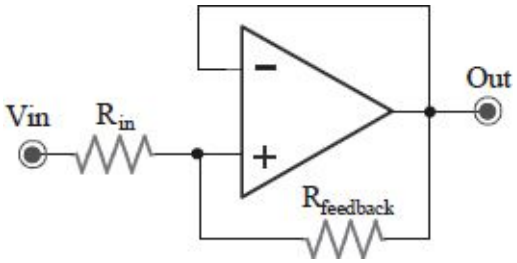


5.38.

The OpAmp is non-compensated (A_0=*110dB*, A_{min}=*30dB*, f_0=*1kHz*) and has input parasitic capacitances ($C_{in,Diff}$=*6pF* and $C_{in,CM}$=*4pF*).

a) Prove that the configuration is unstable without $R_{feedback}$.

b) Size $R_{feedback}$ to have a bandwidth of *30MHz* and *45°* of *PM*.

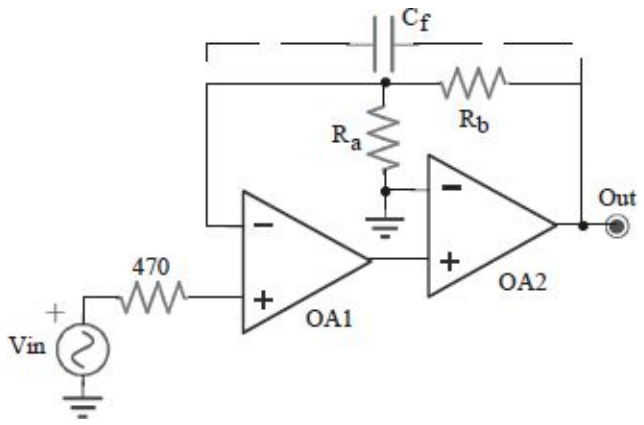


5.39.

Study the double OpAmp stage; every OpAmp has A_0=*200'000*, *GBWP=2MHz*, V_{OS}=*5mV* and I_B=*500nA*.

a) Size the component parts to reach a gain of *1,000;* computes the effect of the output offsets.

b) Compensate the network, if needed, to guarantee a phase margin of at least *90°*.

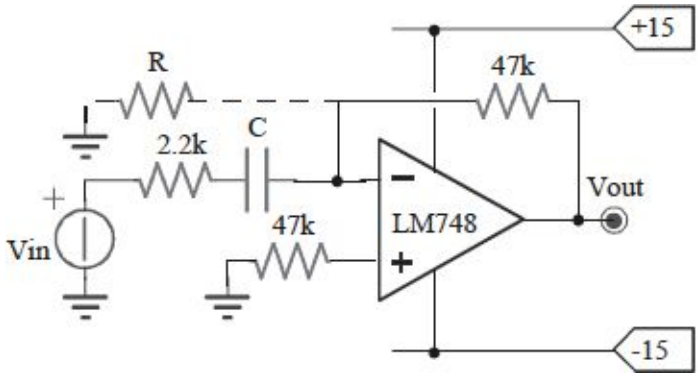


5.40.

Is shown an approximated derivator, which uses a LM748 non-compensated (A_0=*105dB*, second pole at *1MHz*, A_{min}=*20dB*).

a) Establish the value of *C* which guarantees a *PM>45°* and draw the Bode diagram for the gain $v_{out}/v_s(s)$.

b) Explain the changes introducing a *R=2.2kΩ* and how the value should be chosen.

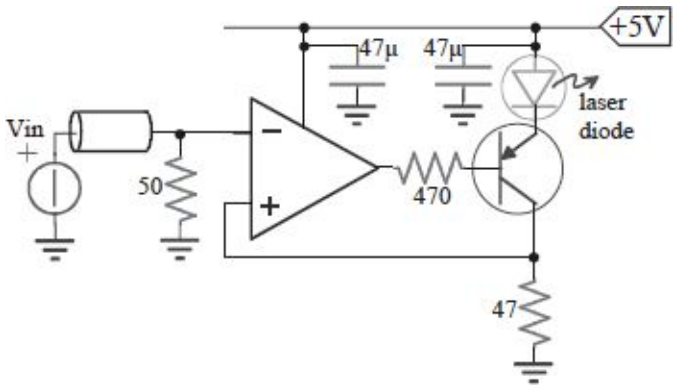


5.41.

The LASER diode ($V_d \approx 1.3V$) is driven by the PNP ($\beta = 40 \div 80$) and by the OpAmp ($A_0 = 110dB$, $GBWP = 100MHz$, *output swing* $= 0 \div [Vdd - 0.8]$).

a) Calculate the C_{bc} value that assures a phase margin of at least $45°$, when is applied an input of $V_{in} = +1V$.

b) Determine the maximum power that the transistor should dissipate in the worst case.

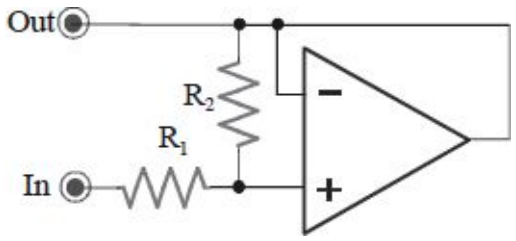


5.42.

The OpAmp is non-compensated ($A_0=120dB$, $A_{min}=40dB$, $f_0=10kHz$) and has input parasitic capacitance ($C_{in,Diff}=6pF$ and $C_{in,CM}=5pF$).

a) Verify if the stage is stable without $R_{feedback}$.

b) Size R_2 to have bandwidth of $30MHz$ and $45°$ of PM.

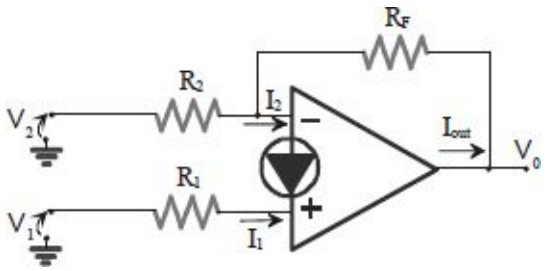


5.43.

The circuit uses a *Current Mode Amplifier* with $A_i=5$.

a) Computes the voltage gains V_0/V_1 and V_0/V_2.

b) Design a stage with buffer *Current Mode Amplifier* which has a gain of *+1*.



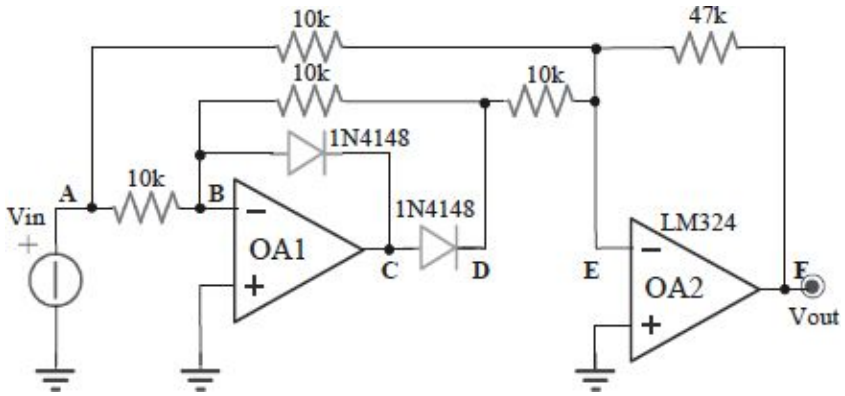
5.44.

The NPN OpAmps have I_B=*100nA*, V_{OS}=*5mV*, *output swing=(–* V_{EE}*+2V)÷(+*V_{CC}*-1V)*, A_{min}=*10dB* and f_{high}=*10MHz*.

a) Study the behavior of the circuit, with a sinusoidal input with a peak of *100mV*, drawing the waveforms for the indicated nodes.

b) Determine the output voltage when *Vin=0V*, due to OpAmps.

c) Determine the worst phase margin.

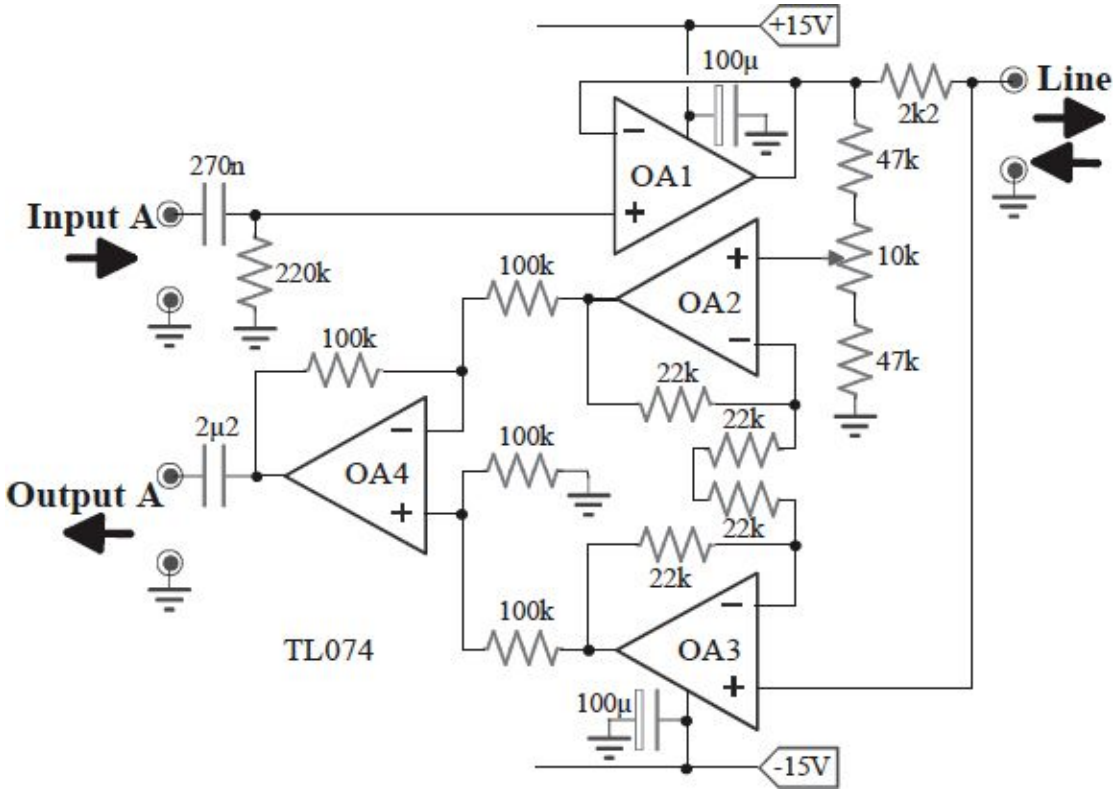


5.45.

The circuit is a "*Duplex Audio Link*", capable of use the *Line* for both transmit *InputA* signal (from left to right) and receive signal from remote user (from right to left) sending it to the *OutputA* terminal. The remote circuit is identical to this (notice the impedance matching).

a) Extract all the relationships input/output, i.e. *Line/InputA*, *OutputA/InputA* and *OutputA/Line*, when the trimmer is correctly calibrated.

b) Compute the residual crosstalk, in *dB*, on the *OutputA* due to the signal *InputA*, caused by the resistive tolerances of *1%*.

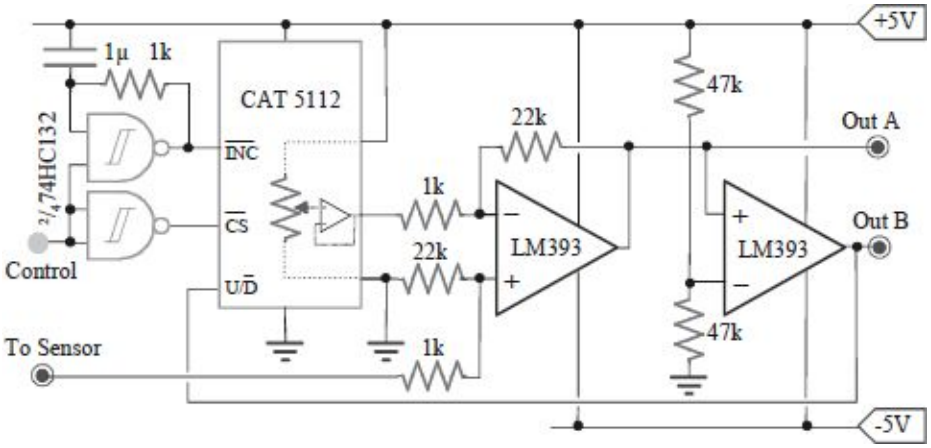


5.46.

The DigPot has *64taps* and connects the output to GND when disabled. The sensor signal can change between *0÷5V*.

a) Draw the output waveforms when *Control=High* and the sensor gives a sinusoid with a frequency of *1Hz*, with maximum amplitude.

b) Discuss what would happen if all the resistances were *1kΩ*.

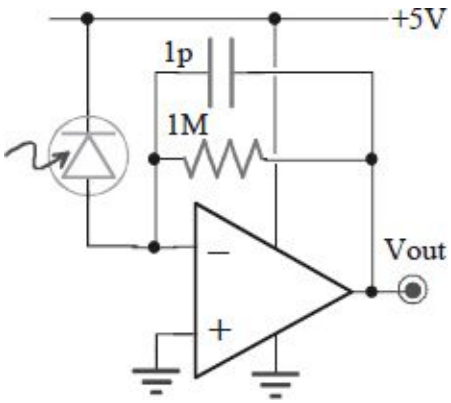


5.47.

The OpAmp is **non-compensated** with $A_0=100dB$, $f_{high}=10MHz$, $A_{min}=20dB$. The photodiode has $C_d=120pF$ and sensitivity of $5mA/W$.

a) Analyze the stability and draw the frequency response $V_{out}/I_{light}(f)$.

b) Explain why the circuit doesn't work and modify to assure an output range of $2V$ for light signals of $400mW$, ensuring the stability.

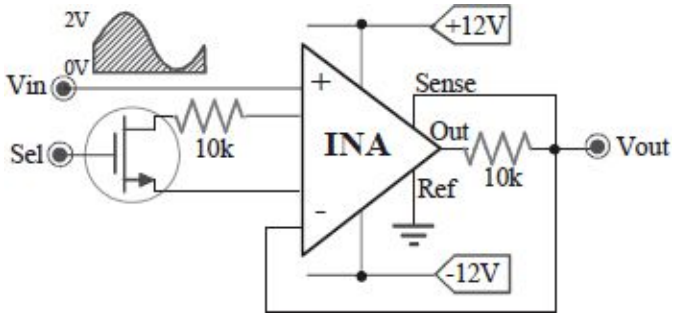


5.48.

The circuit uses an INA (internal resistance of *10kΩ*, and internal OpAmps with V_{os}*=1mV*, I_B*=10nA*, I_{os}*=1nA*).

a) Determine V_{out}/V_{in} when **Sel**=*low*.

b) Size the *high* level for **Sel** that guarantee a good closure of the MOS and with **Sel**=*high* determine the residual output DC offset.

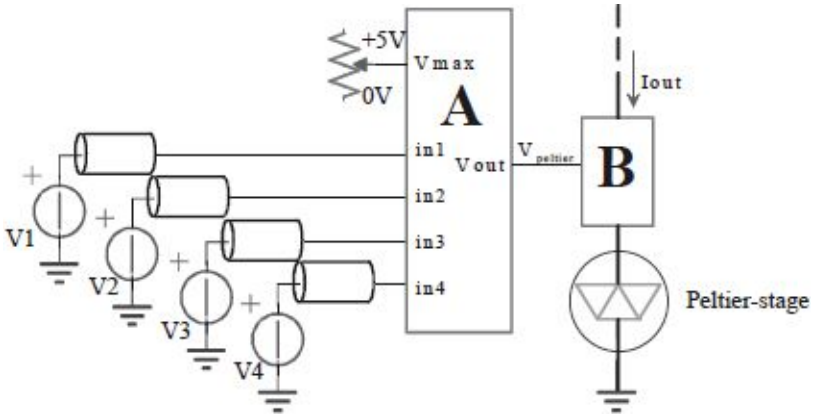


5.49.

The A block provides to the output the maximum signal from the four applied to the inputs (with a range of *0÷5V*) but, however, never greater than the value set by the trimmer, with an accuracy better than *20mV*. The B block provides to the output the current flowing in the generic dipole (in figure the Peltier-stage) when supplied to a $V_{peltier}$ voltage.

a) Design the A network.

b) Design the B network.

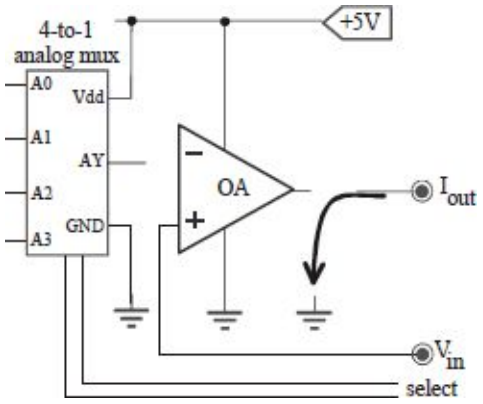


5.50.

Using the parts of the Ex.5.49 we want make a Voltage-Current converter.

a) Design the circuit to have an output variable in the ranges *0÷1mA*, *0÷10mA*, *0÷100mA* and *0÷1A* when the input varies between *0÷5V*.

b) Observe the stage stability and propose mothods to guarantee it in the worst condition.

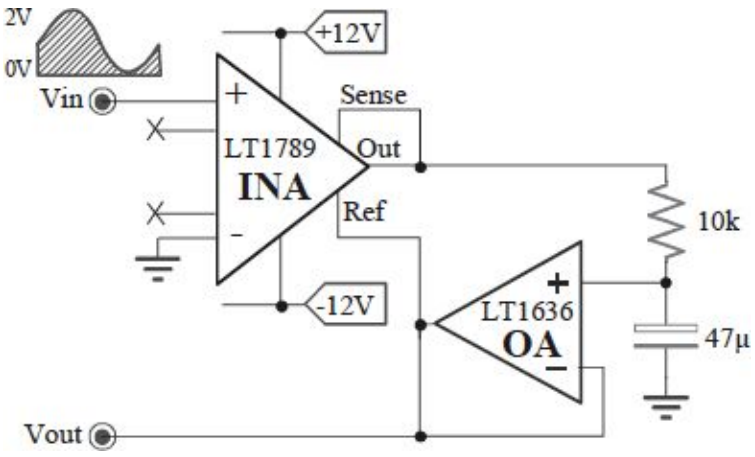


5.51.

An INA is used with a normal OpAmp with slow changing input (*<1kHz*).

a) Determine the performed task.

b) Modify the circuit to assure the complete discharge of the capacitance to *–12V* every *100ms* or when its voltage overcomes *+10V*.

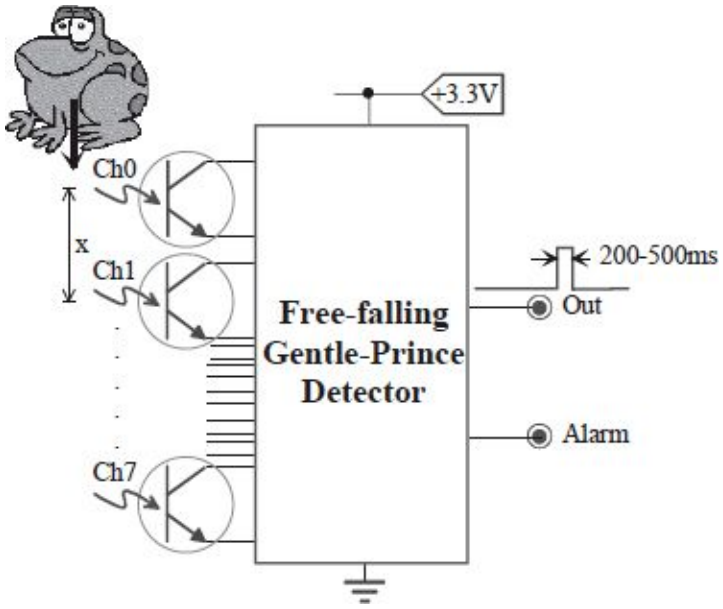


5.52.

Develop the free-fall “Prince Charming” detector using *8* phototransistors placed at distance *x*. The network provides a pulse every *200*, *300*, *400*, *500*, *600ms* depending on the toad dimension (*2x*, *3x*, *4x*, *5x*, *6x* rispectively). Contemporary falls for more toads are not possible; one frog (dimensions less then *2x*) must trigger an alarm.

a) Design the overall digital network using standard MSI components.

b) Propose an analog implementation.

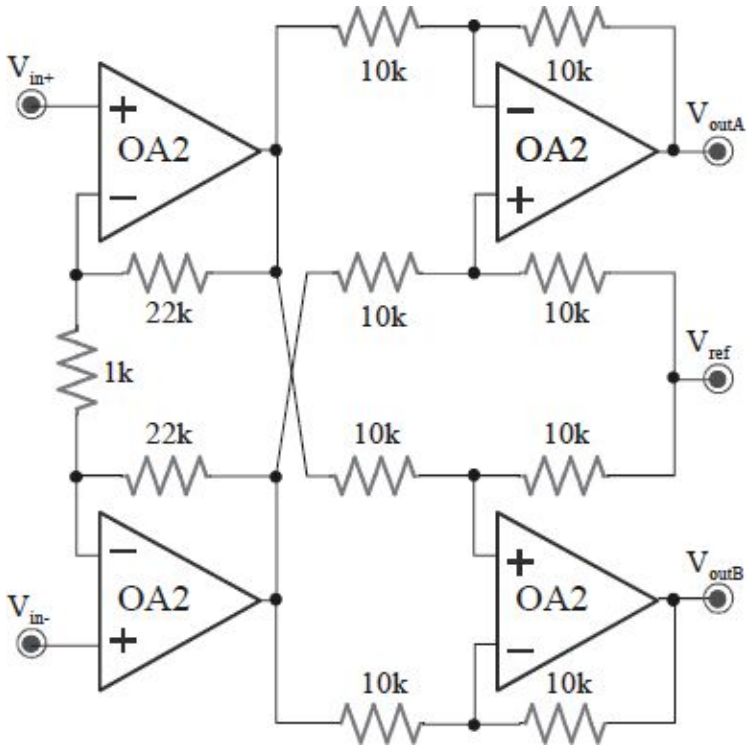


5.53.

A quad OpAmp is $\pm 12V$ supplied with the pin $V_{ref}=+5V$.

a) Extract the relationship of V_{outA} and V_{outB} as function of the input differential signal.

b) Assuming an offset current of $I_{OS}=10nA$, determine the worst case for the output offset over-imposed to V_{outA} and V_{outB}.



6

Sampling circuits

“Everywhere the sun is shining.
All around the world it’s shining.
But cold winds blow across your mind.
Confusion--it’s such a terrible shame.
Confusion--you don’t know what you’re sayin’.
You’ve lost your love and you just can’t carry on.
You feel there’s no-one there for you to lean on.
Every night you’re out there darlin’.
You’re always out there runnin’.
And I see that lost look in your eyes.
Confusion, I don’t know what I should do.
Confusion, I leave it all up to you.
You’ve lost your love and you just can’t carry on.
You feel there’s no-one there for you to lean on.

“Confusion”, Electric Light Orchestra

6.1. SAMPLING THEOREM

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. Previously, H. Nyquist and D. Gabor had published results related to the Sampling Theorem.

The original statement of the Sampling Theorem, as appears in the 1949 Shannon's paper, is the following: "*If a function x(t) contains no frequencies higher than BW hertz, it is completely determined by giving its ordinates at a series of points spaced 1/2BW seconds apart*". Practically, it states that a continuous-time signal can be reconstructed starting from a series of discrete-time samples, clocked at a distance T_S=1/2BW, if every sample is replaced with a cardinal sine sinc(.)=sin(.)/(.) with the same width as the distance between two samples.

Thus, we use the following reconstruction formula:

$$s(t) = \sum_{n=-\infty}^{+\infty} s(n \cdot T_S) \cdot \frac{\sin\left(\pi \cdot \left(t/T_S - n\right)\right)}{\left(\pi \cdot \left(t/T_S - n\right)\right)}$$

where s(t) is the analog time-continuous function and $s(n \cdot T_S)$ is the samples sequence.

Fig. 6.1 shows an example of the original signal, the sampled signal, the action of the reconstruction formula (i.e. the sum of the *sinc* functions weighted and translated as the sampled signal), and finally the reconstructed signal. The reconstructed signal is significantly distorted at the initial and final moments because the interpolating functions are truncated to their extreme while the theory uses the infinite *sinc*.

An interesting interpretation of the reconstruction formula is derived from the signal theory; in fact, the interpolating *sinc* function is the pulse response of an ideal low-pass filter with bandwidth BW while the reconstruction formula is the convolution between BW and the sample sequence. For this reason, practically, the reconstructed signal is obtained from the ideal low-pass filtering of the samples, which can be thought of as a comb of pulses modulated by the original signal. The concepts just described about sampling and about reconstruction are summarized in Fig. 6.2.

The properties of duality and symmetry of the Fourier Transform validate the Sampling Theorem in the opposite direction, i.e. from the frequency domain to time domain: "*If F (ω) is the spectrum of a time-limited function f*

(t), which is zero for /t/ >T, then it is completely determined by the values taken at frequencies spaced 1/2T Hz from one another". Notice that T is T_S, the sampling period, but represents the total duration of the signal in question.

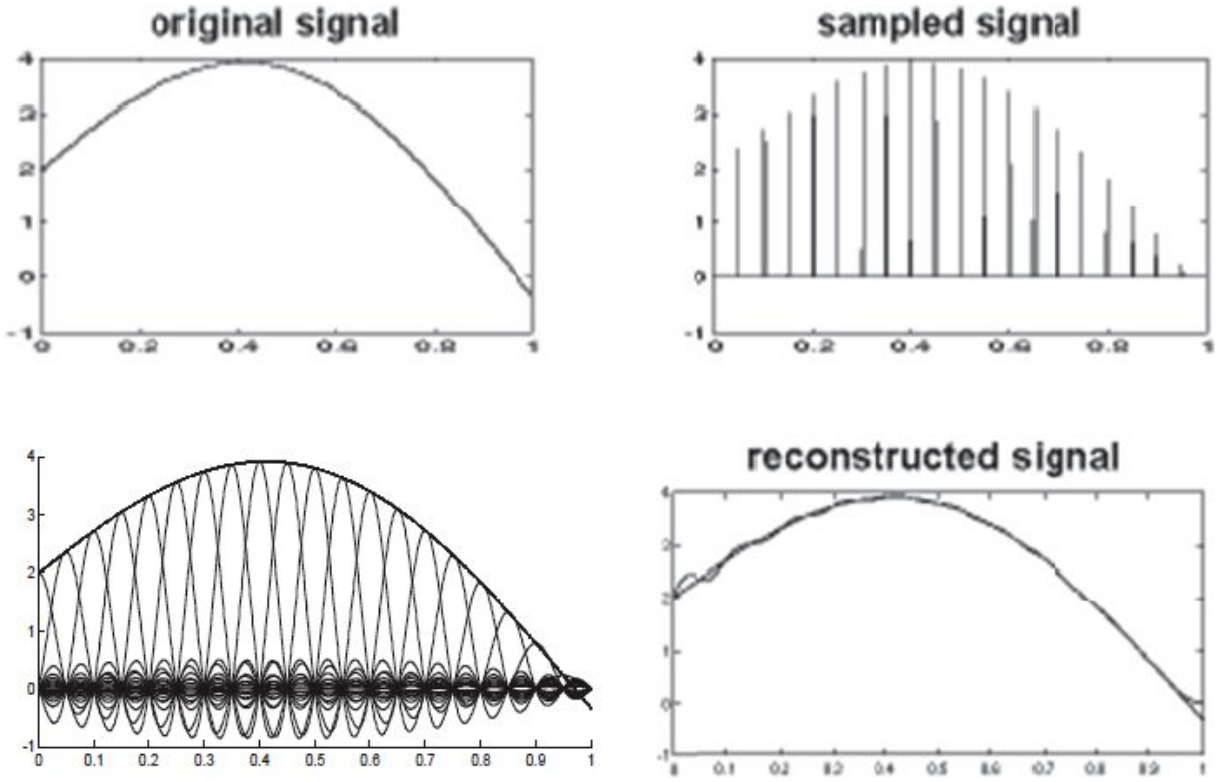


Fig. 6.1: Example of signal reconstruction. The signal has inaccuracies at the edges of the time range due to the truncation of the cardinal sine.

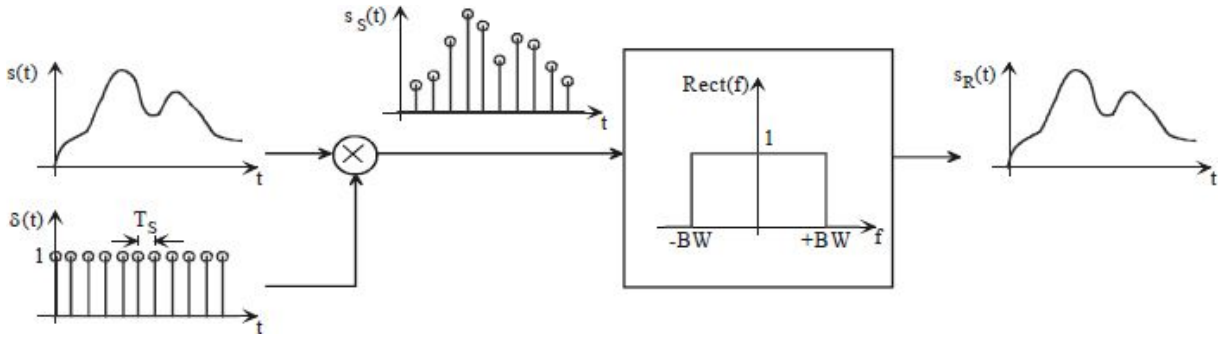


Fig. 6.2: Signal sampling and reconstruction.

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.

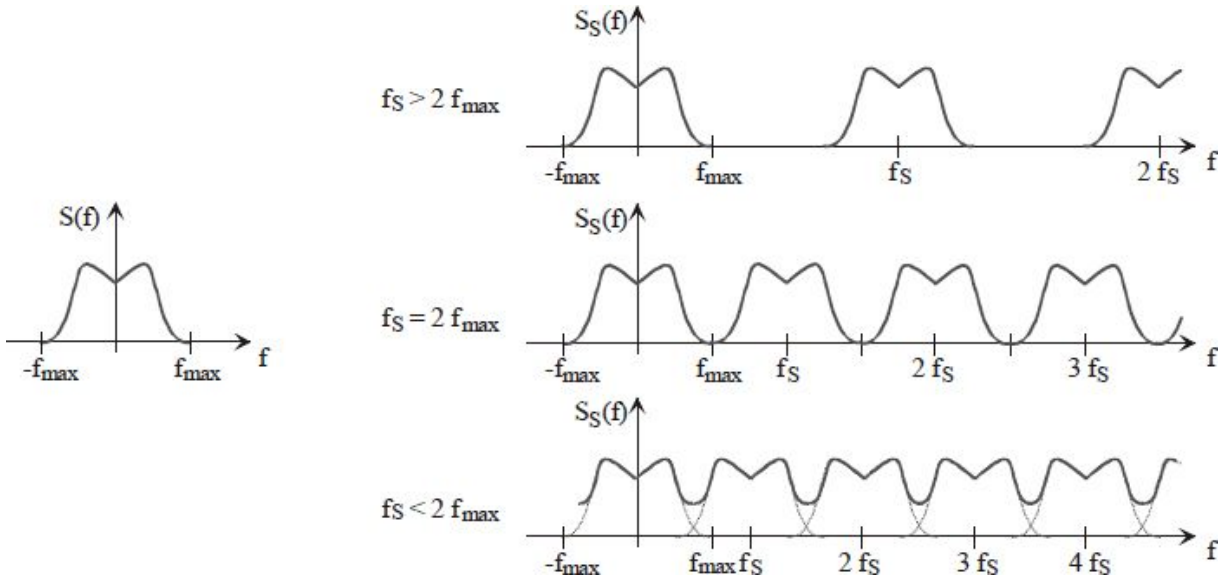


Fig. 6.3: Original signal (on the left) and sampled signal (on the right) spectra for various f_S.

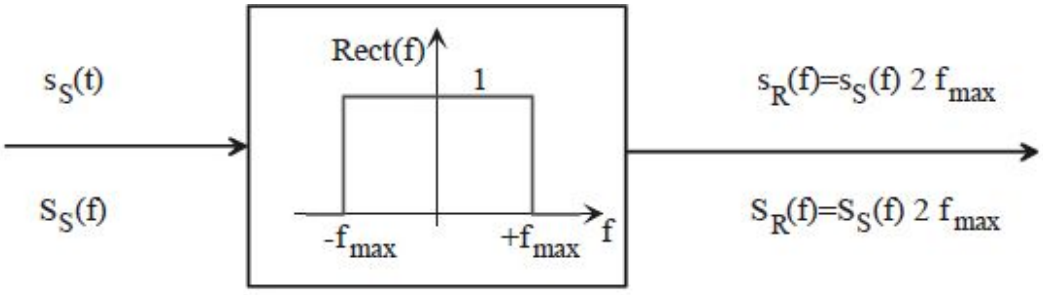


Fig. 6.4: Ideal reconstruction filtering.

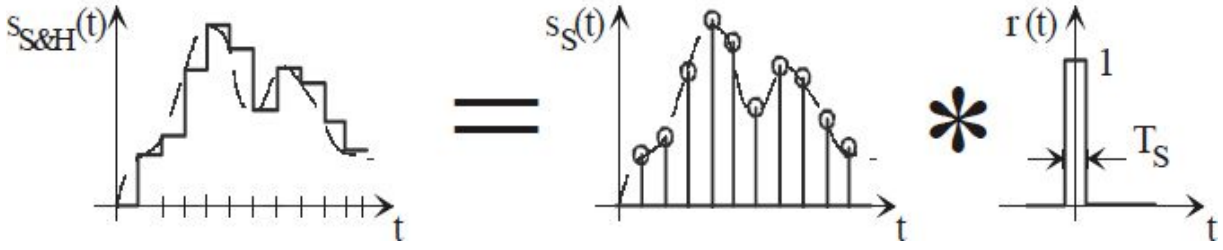


Fig. 6.5: Sampler S&H output signal seen as the convolution between pulse samples and ideal rectangle.

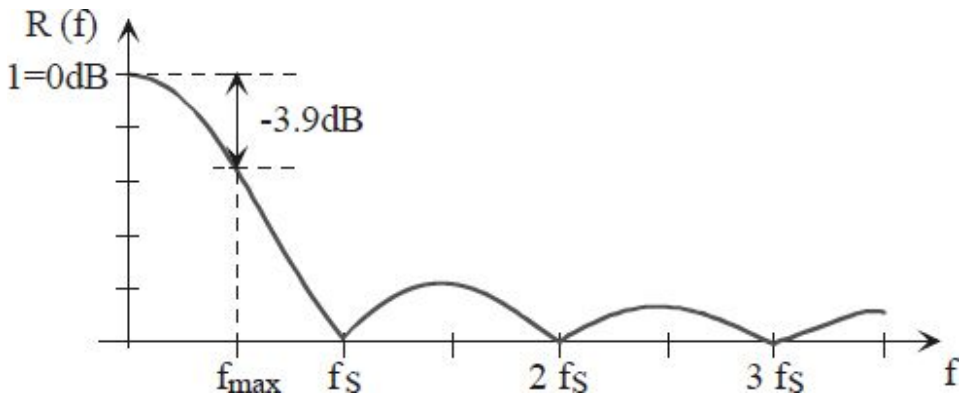


Fig. 6.6: Amplitude distortion on the sampled signal.

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.

Before concluding, it is better to hint to a convention used in digital processing, which is applied to the sampled signal converted into digital sequences. Because the sampling frequency is a fundamental parameter in the subsequent steps (processing and reconstruction), we normally use to normalize frequency axis matching f_s with the single frequency of the sampled

sequence, as shown in Fig. 6.9. To obtain the spectrum of the digital sequence starting from the sampled signal, you can just rescale the axis with a factor equal to f_S; thus, the range between $\pm f_S/2$ is brought between ±½.

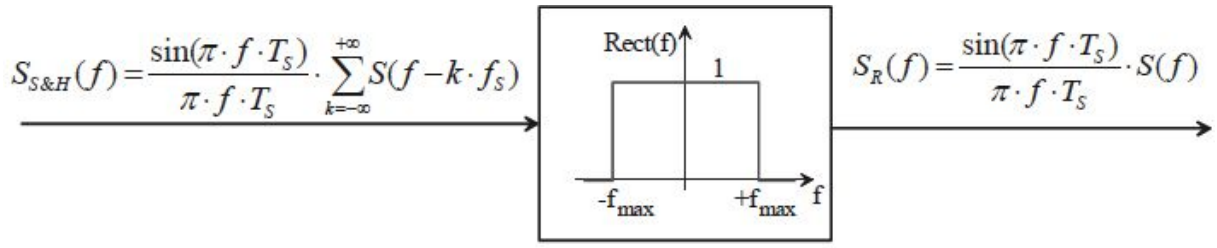


Fig. 6.7: Reconstruction filter.

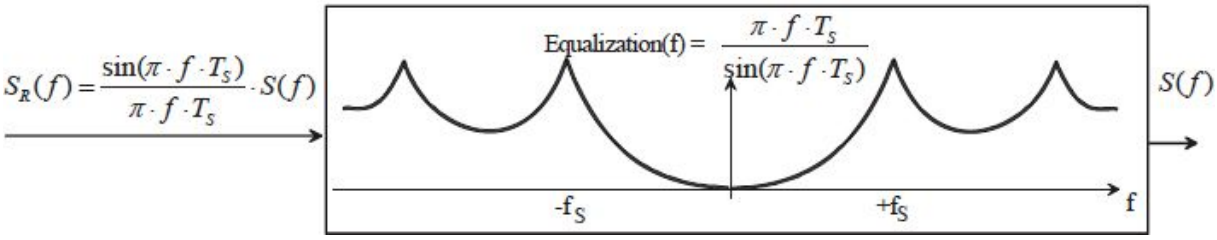


Fig. 6.8: Equalization filter.

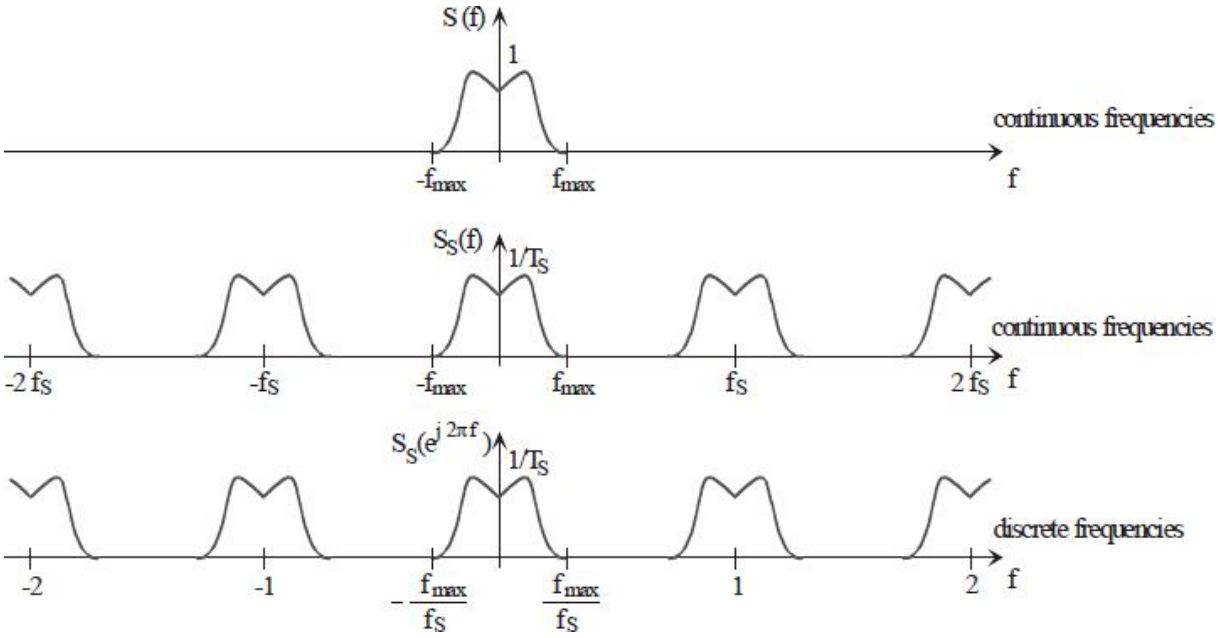


Fig. 6.9: From signal to sequence in the frequency domain.

It should be noted that the transformation process from a sampled signal to a digital sequence involves an amplitude discretization (*quantization*) associated with an error. This problem will be described in detail in the chapter about the analog-digital conversion.

6.2. Hints on Sampling

We will now consider some practical problems that may arise as the result of a sampling process. Some of these problems have already been reported in the previous section whereas the others are new and often underestimated in the designing step or the analysis of circuits.

A family of signals useful for the study of the behavior of systems is the complex exponential family, described as follows:

$$x(n) = e^{s \cdot n} \qquad s = \sigma + j\omega \in C \qquad \rightarrow \qquad x(n) = e^{\sigma \cdot n} \cdot e^{j\omega \cdot n} \qquad \sigma, \omega \in R$$

From the Euler's Formula, imposing σ=0, we obtain:

$$\mathrm{Re}\left\{e^{s \cdot n}\right\} = \mathrm{Re}\left\{e^{j\omega \cdot n}\right\} = \cos(\omega \cdot n) \qquad \mathrm{Im}\left\{e^{s \cdot n}\right\} = \mathrm{Im}\left\{e^{j\omega \cdot n}\right\} = \sin(\omega \cdot n)$$

The frequency f=2π/ω is named the *sequence frequency*. Because of this periodicity, a remarkable characteristic of the sequences is that they cannot represent frequencies less than –½ and greater than +½ because the complex exponential $x(n) = e^{j\omega \cdot n}$ has unity module and phase ω·n (with n=0, ±1, ±2…), can assume values between –π and +π, and is periodic.

To clarify the above-written point, in Fig. 6.10, some complex exponentials are shown with σ=0, i.e. sinusoidal oscillations. The frequencies depicted vary between 0 and 1, and the observation allows verifying that the maximum representable frequency is equal to ½. In fact, in this case, the sinusoid is the same every two samples, i.e. alternatively assumes the values +1 and -1. We can observe the symmetrical behavior around the maximum frequency (½) and the degeneration of a constant sequence for the sinusoids with frequencies equal to 0 and 1. This effect of the sampling is named *aliasing* and will be explained in the following.

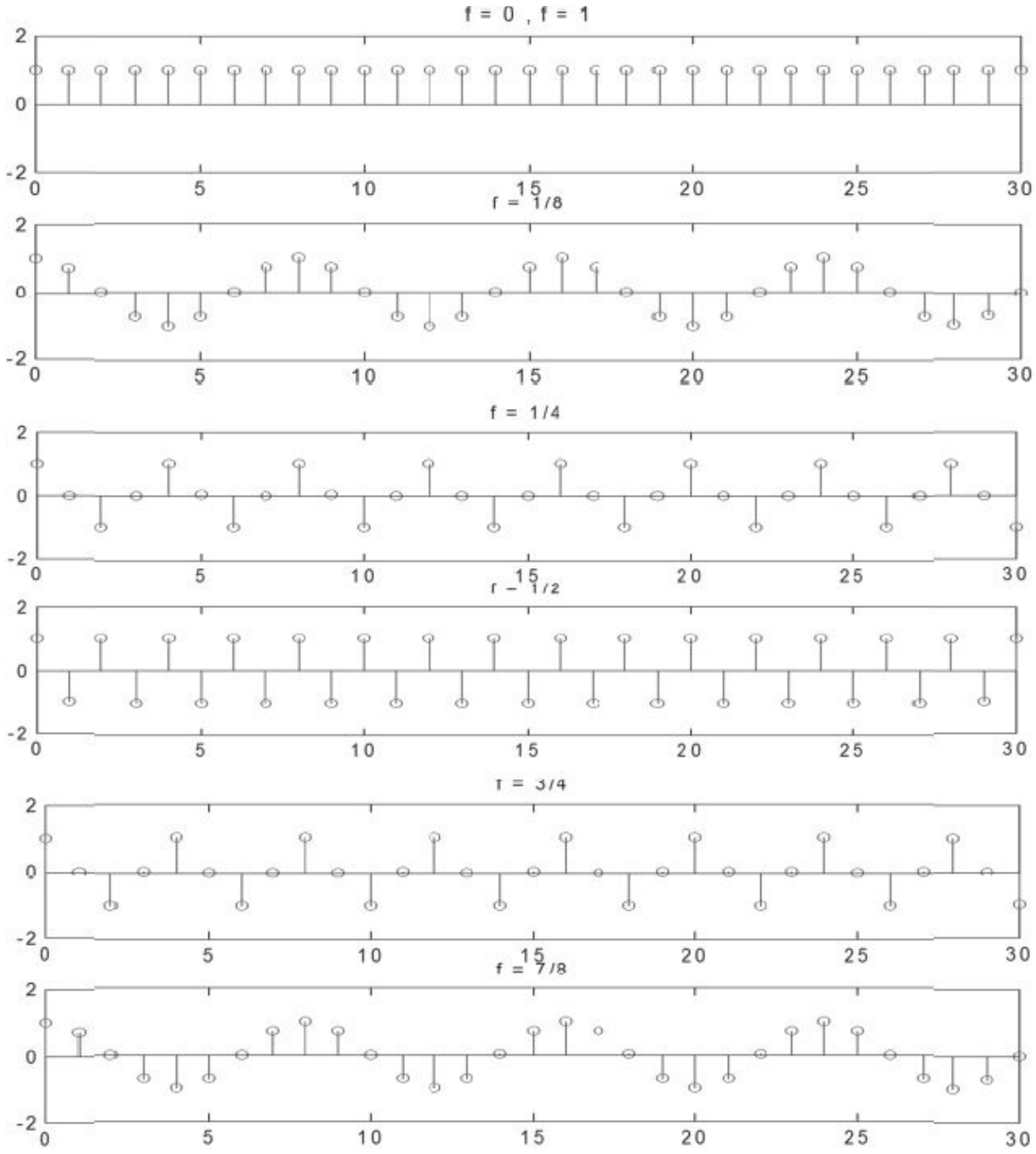


Fig. 6.10: Sampled sinusoids: the frequency is normalized to 1, corresponding to f_S.

Another difference between the time-continuous and discrete-time complex exponentials is the periodicity. In the continuous case, if there is periodicity, the period is equal to the inverse of the frequency; on the contrary, in the discrete case, if there is periodicity, the period is not necessarily equal to the inverse of the frequency.

In fact, in order to have a sequence which is repeated every N samples, the equation $e^{j\omega\cdot(n+N)}=e^{j\omega\cdot n}$ must be satisfied; which implies that the frequency must be a rational number:

$$e^{j\omega(n+N)} = e^{j\omega n} \rightarrow e^{j\omega N} = 1 \rightarrow \omega N = 2\pi k \rightarrow f = k/N.$$

Ultimately, a discrete complex exponential is periodic when the frequency is a rational number. The period is not necessarily the inverse of the frequency; that is, since it must be an integer, it is the smallest multiple of the inverse of the frequency. Fig. 6.11 is intended to illustrate that a complex exponential is periodic only when its frequency is a rational number. In the first case, there is periodicity after 29 samples whereas in the latter case there is no periodicity.

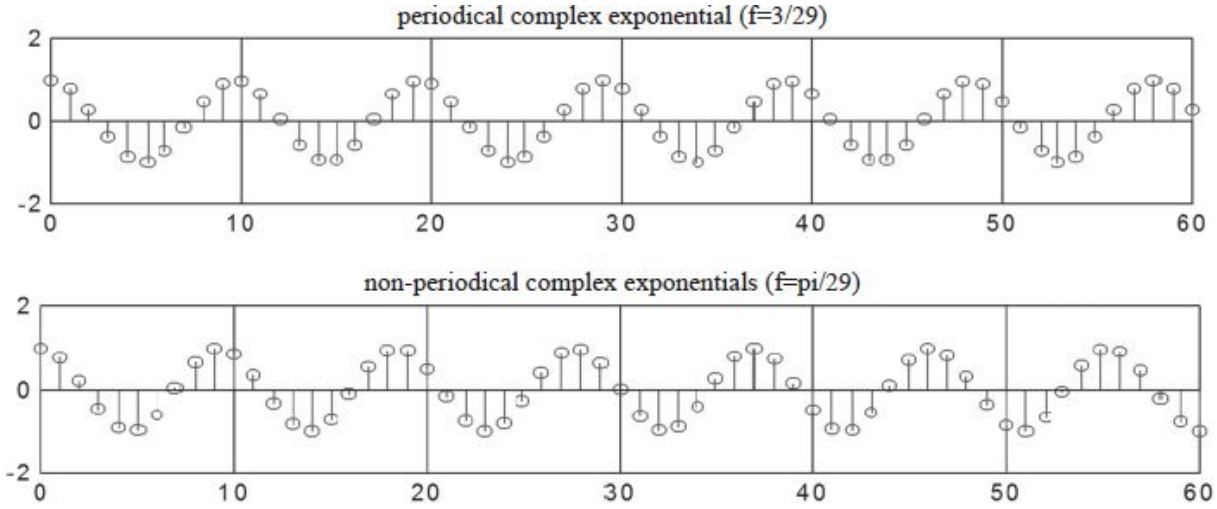


Fig. 6.11: Complex exponentials (σ=0): periodical (at the top), non-periodical (at the bottom).

The parameter σ, as for the continuous-time, acts on the sequence amplitude, amplifying it if positive or attenuating it if negative, as shown in Fig. 6.12.

We can now analyze some effects due to the sampling.

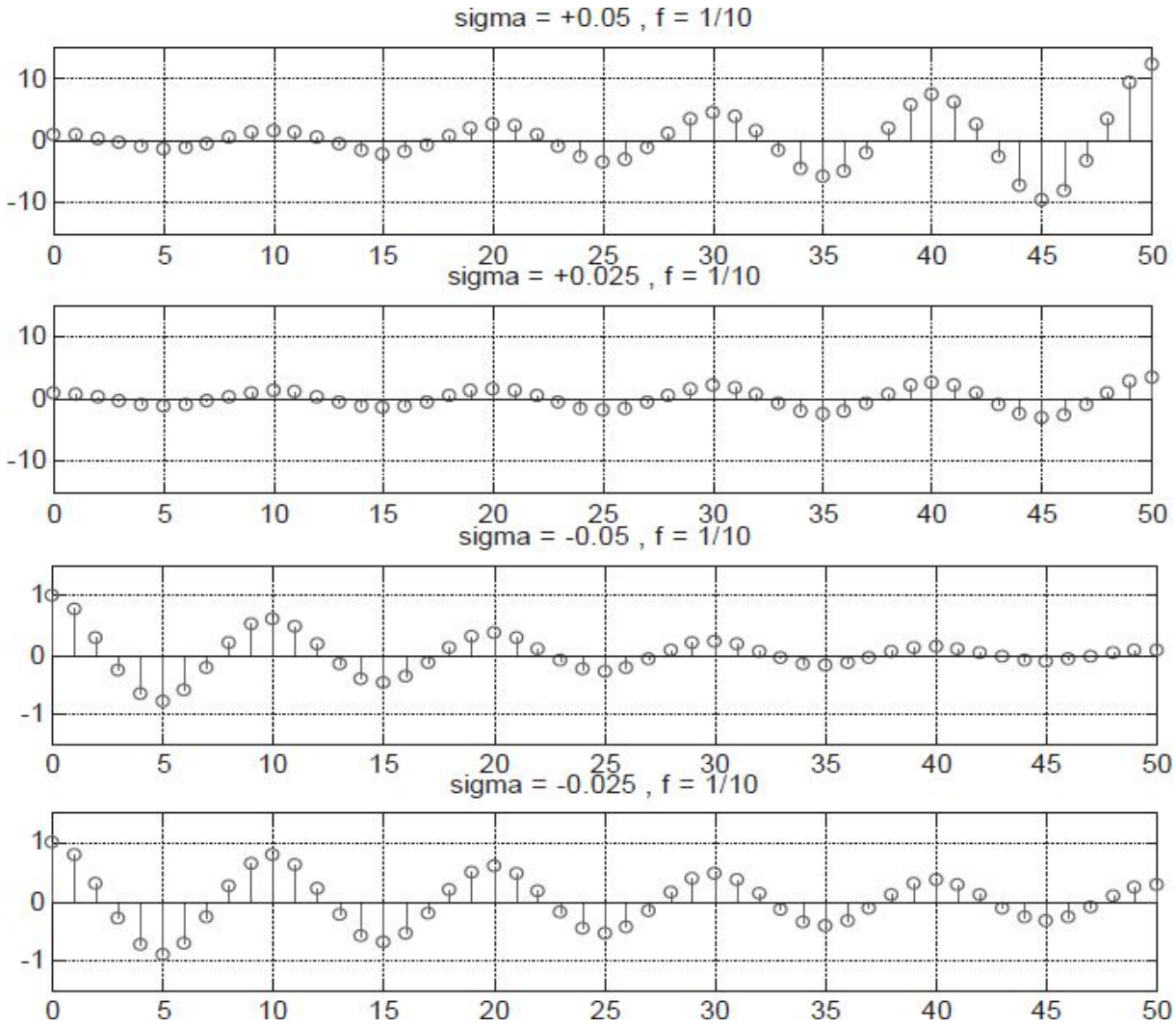


Fig. 6.12: Parameter σ effect.

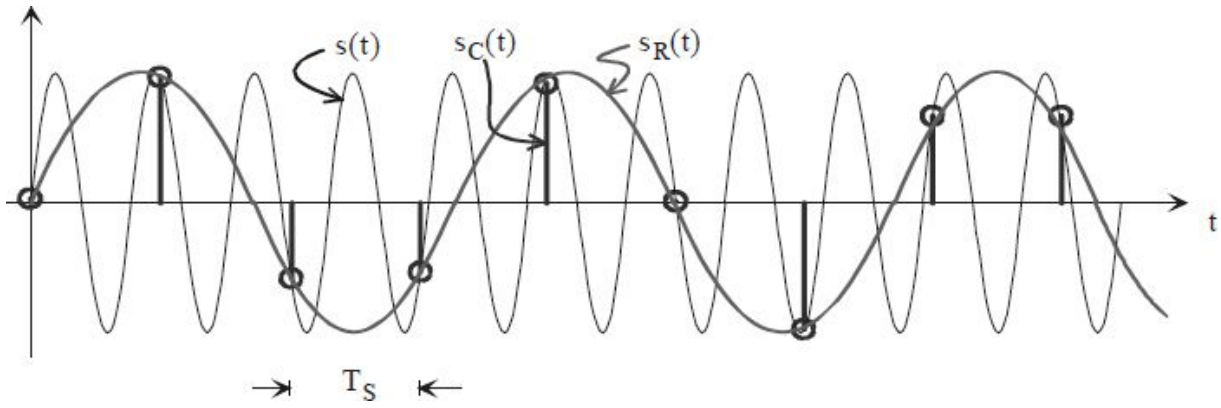


Fig. 6.13: Sampling with a low frequency f_S, and hence the aliasing problem rises.

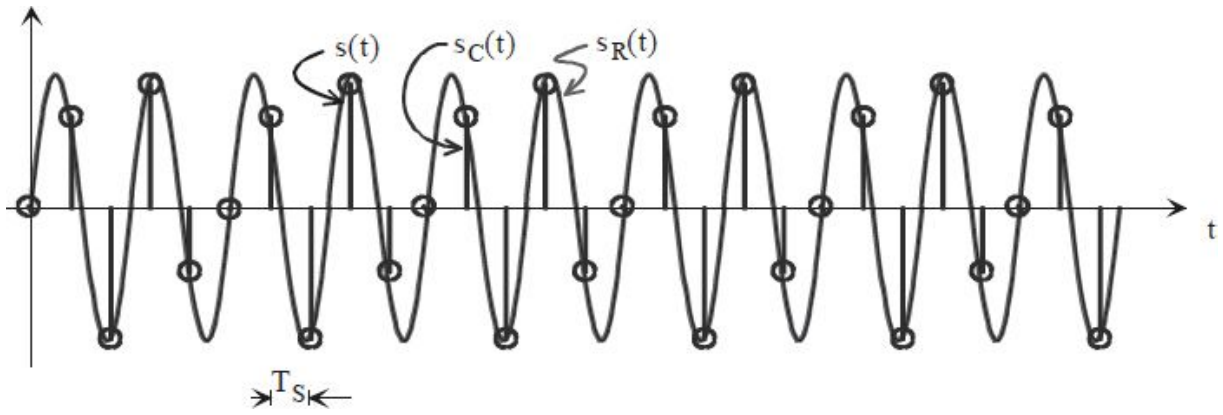


Fig. 6.14: Sampling with a frequency f_S high enough, we can avoid the aliasing depicted in Fig. 6.13.

6.2.1 Aliasing

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.

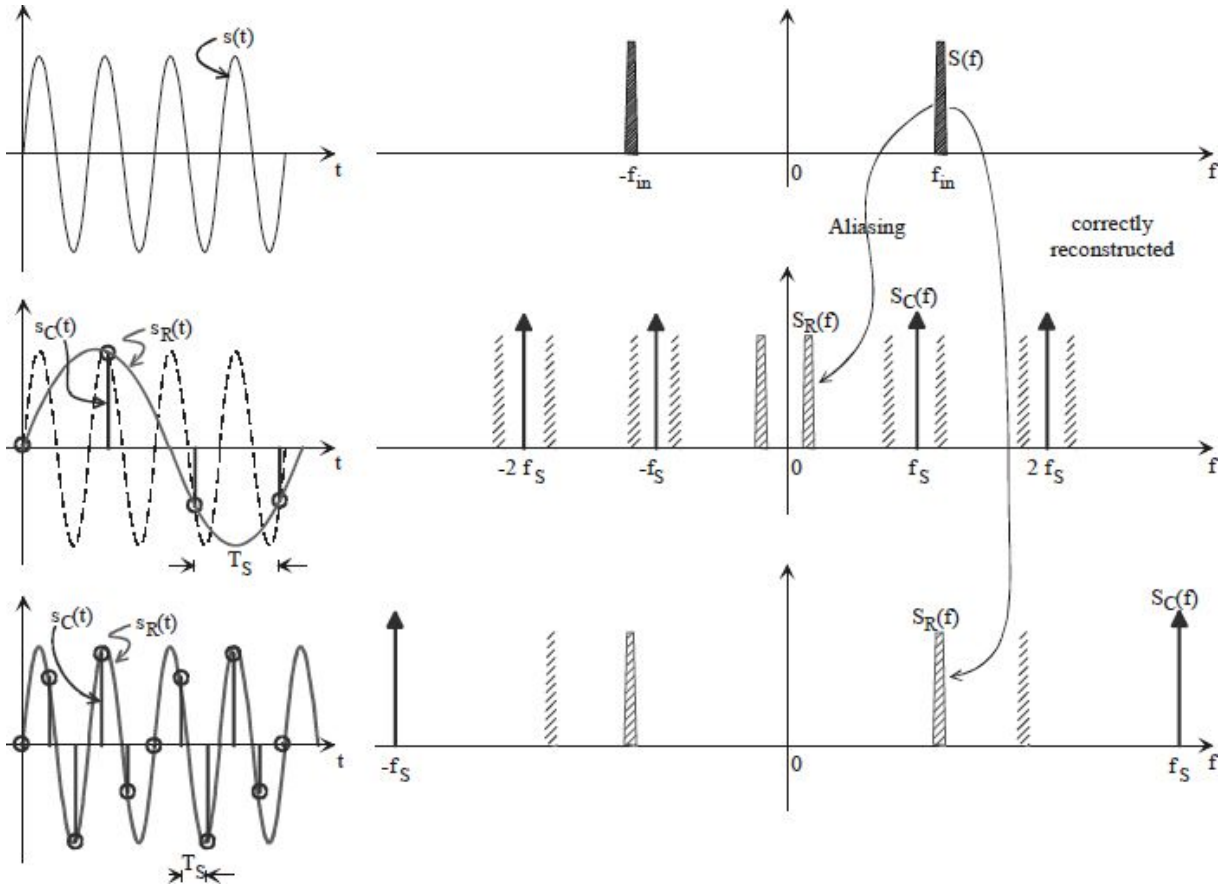


Fig. 6.15: Analog time-continuous signal (at the top), sampled with $f_S<2 \cdot f_{max}$ and, thus, with aliasing (in the center), and correctly sampled with $f_S>2 \cdot f_{max}$ (at the bottom).

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.

6.2.2 Anti-aliasing filtering

As seen previously, given the input signal bandwidth f_{max}, you must choose a sampling frequency f_s at least twice the former. Unfortunately, in addition to the input signal, there are other components, such as noise and unwanted harmonics. Even though some of them have components at frequencies higher than f_s / 2, they could be included in the base band due to the sampling operation, causing aliasing and distortion. These spurious components are misunderstood, could have a frequency in the signal bandwidth, and could be definitively interpreted as the original signal. The example in Fig. 6.18 is explanatory: a disturbance at 130kHz, non-audible, is superimposed on the input sinusoid at 10kHz.

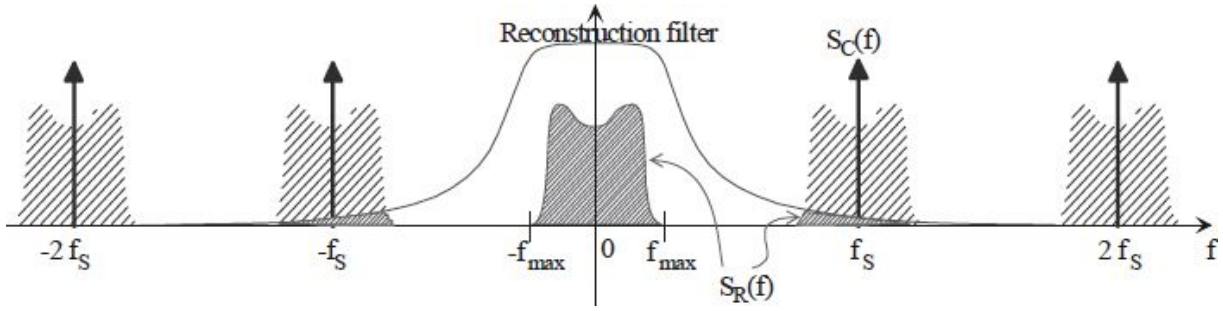


Fig. 6.16: Reconstruction filter selectivity for a sampled signal with $f_S>>2\cdot f_{max}$.

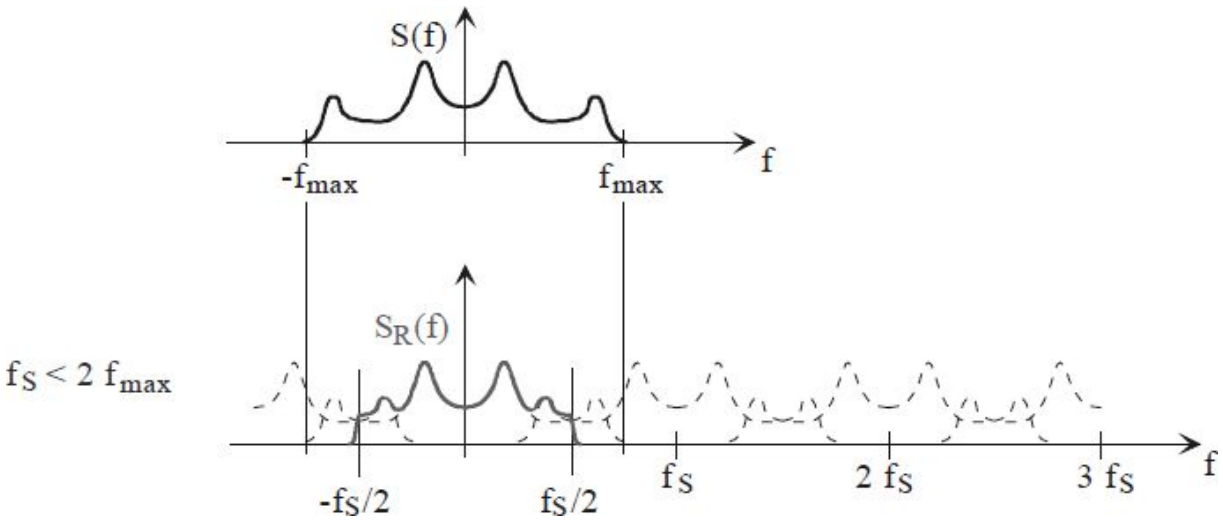


Fig. 6.17: Superimposition of the sampled spectra with a consequent spectrum deformation of the reconstructed signal, also in the case of ideal filtering with a cut-off frequency $f_S/2$.

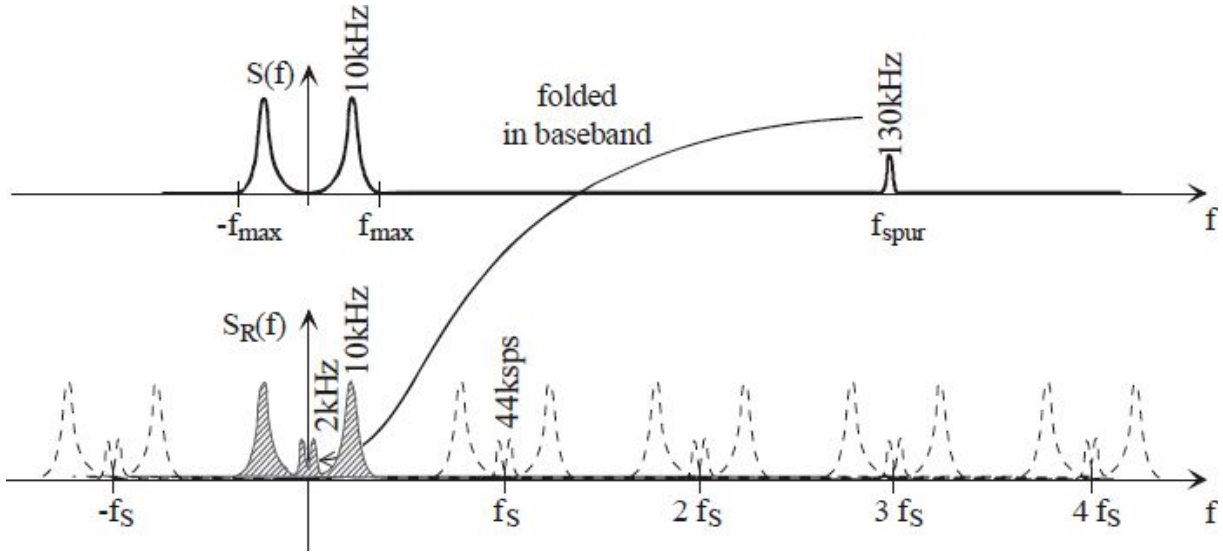


Fig. 6.18: Effect due to the input sampler disturbances with a frequency higher than $f_S/2$.

However, sampling correctly at 44ksps, we will obtain the starting sinusoid with an unattended oscillation at 2kHz, which is indistinguishable from the useful signal. To this end, we need to insure that only the desired signals are at the sampler input with limited bandwidth lower than $f_S/2$. We must use a selective filter indicated as the *anti-aliasing* filter. Hence, the Fig. 6.19 shows the whole processing system.

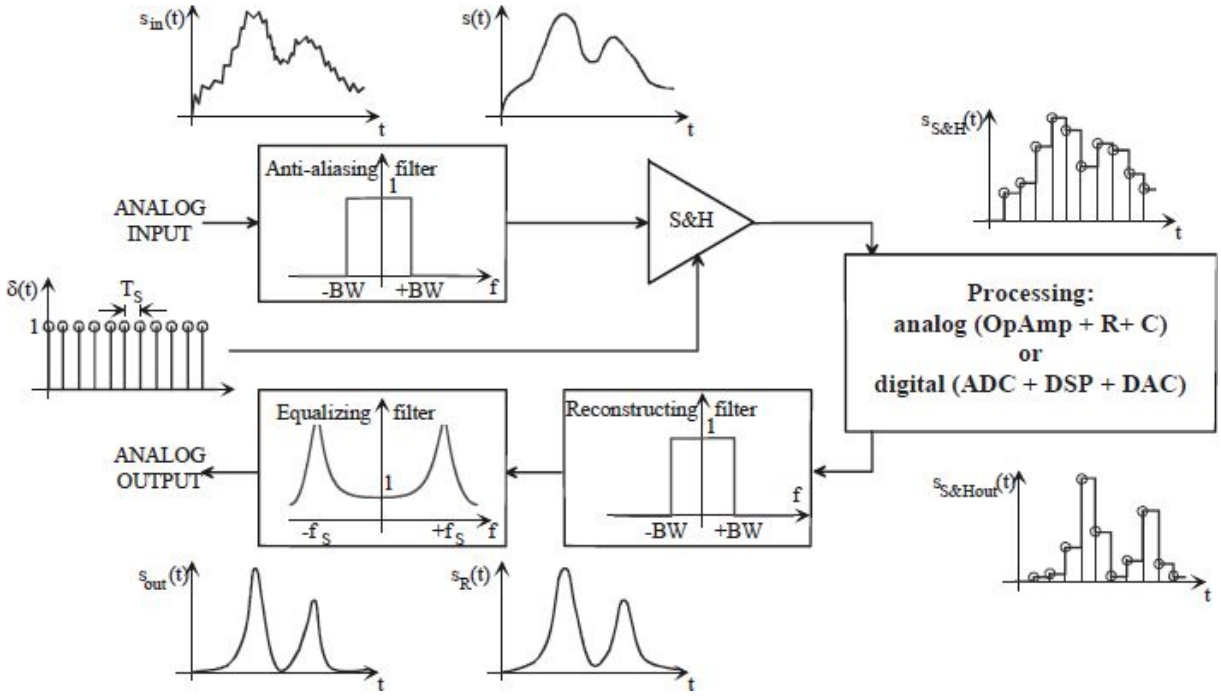


Fig. 6.19: Complete processing system with sampled data, including an output equalizer filter.

6.2.3 Windowing for the spectrum computation

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.

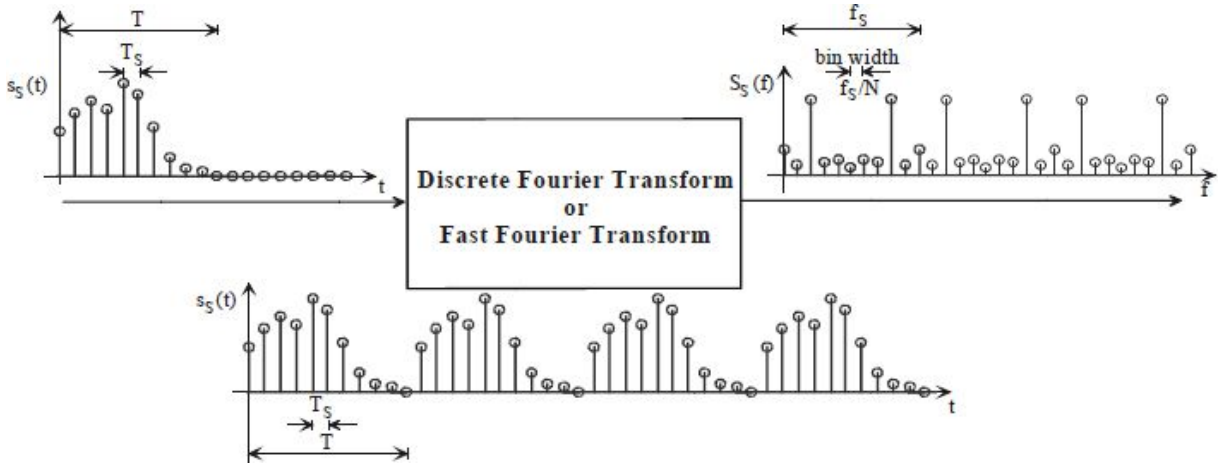


Fig. 6.20: FFT algorithm for the computation of the spectrum on an input sequence with N samples repeated periodically, as for the N values provided by the FFT, every f_S.

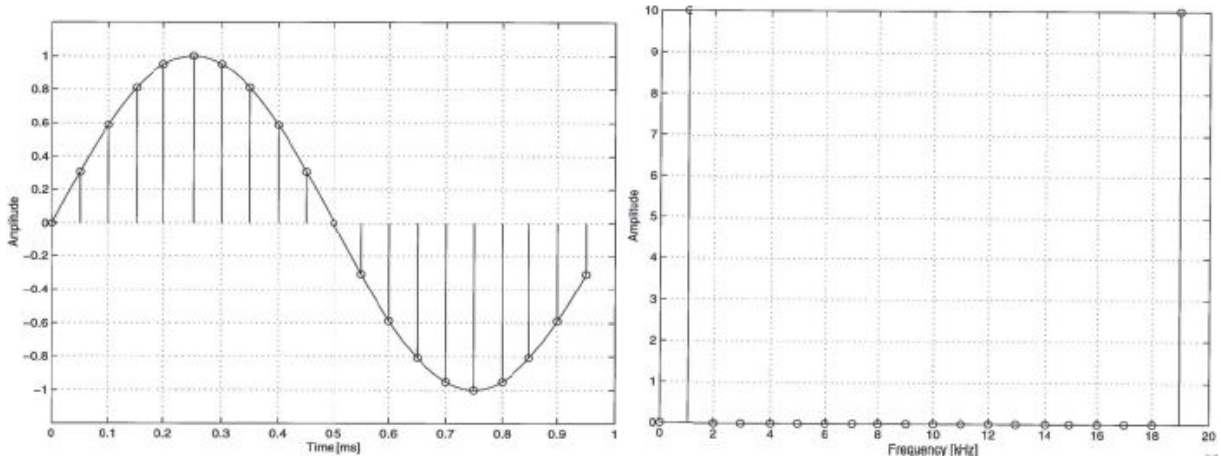


Fig. 6.21: Sinusoid and its FFT computed with N=20 samples.

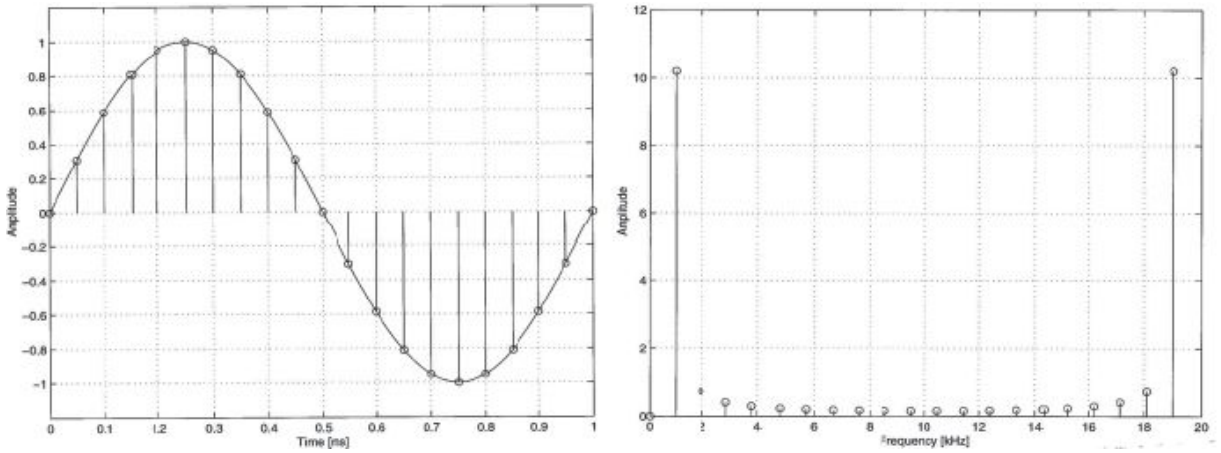


Fig. 6.22: Deformation of the FFT for a sinusoid.

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.

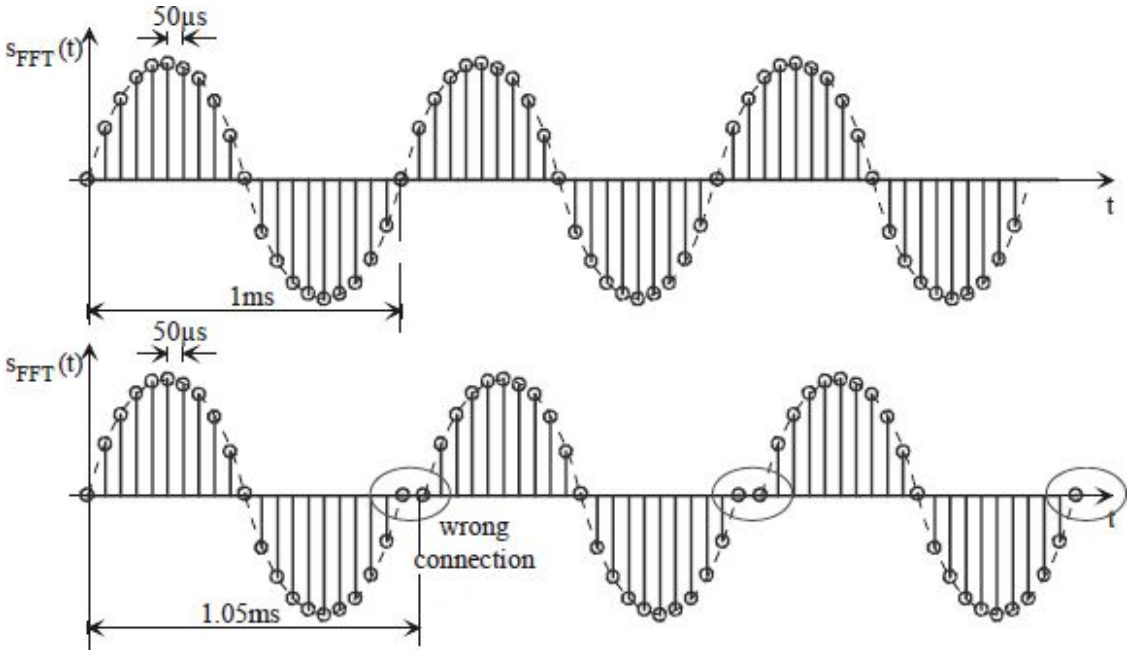


Fig. 6.23: Sequences of Fig. 6.21(a) and Fig. 6.22(b), as effectively considered by the FFT algorithm.

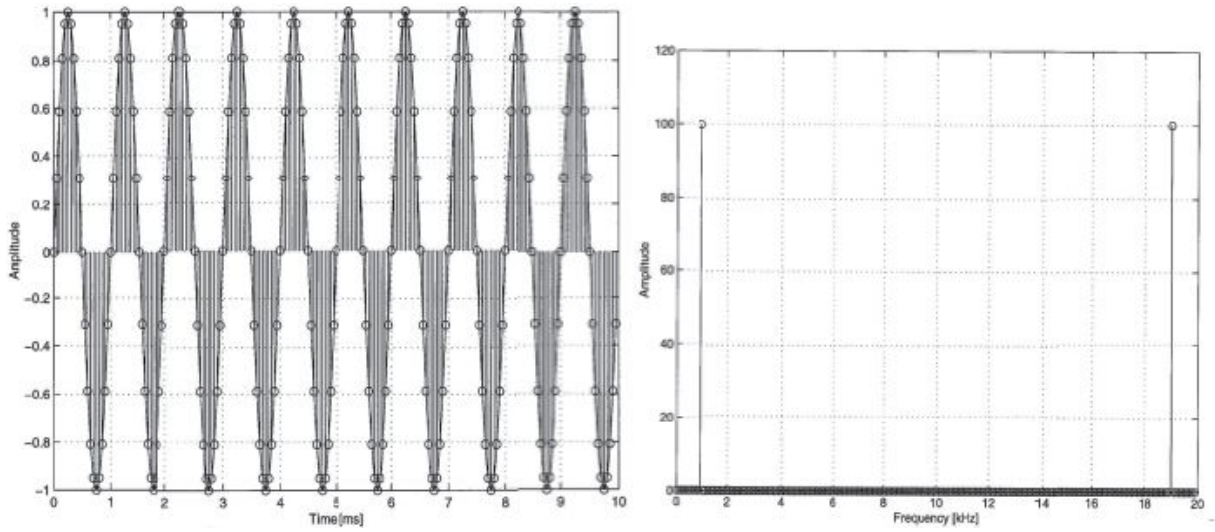


Fig. 6.24: Original and sampled signal, but with N=200 samples.

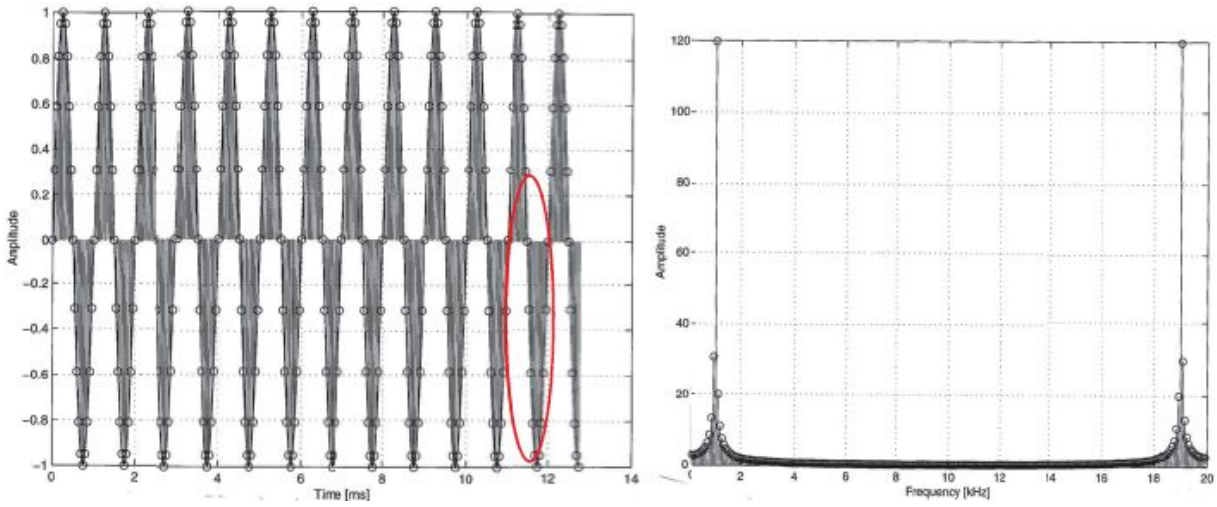


Fig. 6.25: Original and sampled signal, but with N=256 samples. Notice the abrupt ending of the sequence, which prevents a correct connection.

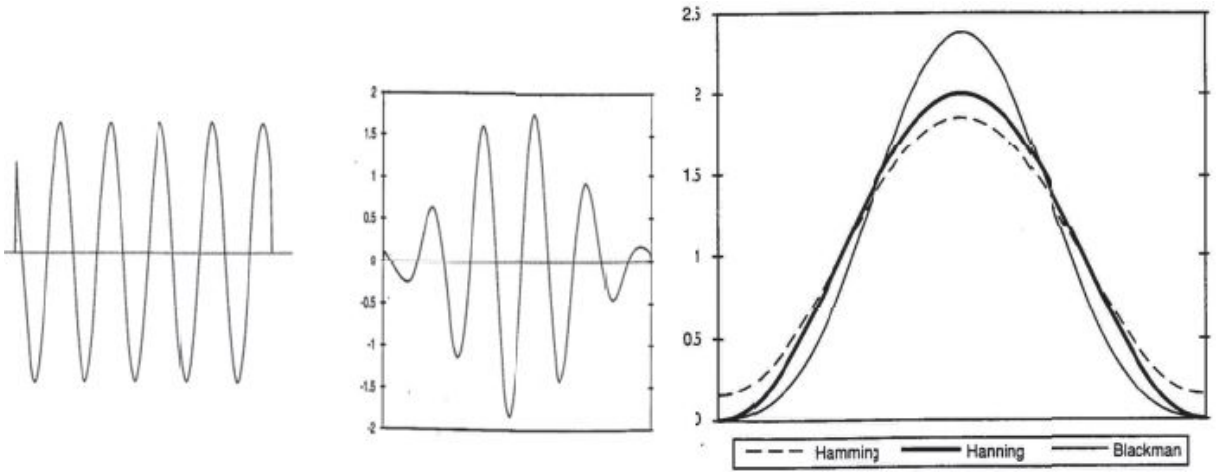


Fig. 6.26: In the center, an example of windowing with values smoothing the original sequence (on the left). On the right, some of the most common windows.

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:

$$\text{window}(n) = 1 - \frac{|n|}{M} \qquad \text{if} \quad -M \le n \le M \qquad \text{otherwise} \quad 0$$

In the case of N samples, other windows (Fig. 6.27) are the Hanning (raised cosine):

$$\text{window}(n) = \frac{1}{2} \cdot \left(1 + \cos\frac{2\pi \cdot n}{N}\right) \qquad \text{if} \quad -\frac{N-1}{2} \le n \le \frac{N-1}{2} \qquad \text{otherwise} \quad 0$$

the Hamming:

$$\text{window}(n) = 0.54 + 0.46 \cdot \cos\frac{2\pi \cdot n}{N} \qquad \text{if} \quad -\frac{N-1}{2} \le n \le \frac{N-1}{2} \qquad \text{otherwise} \quad 0$$

and the Blackman:

$$\text{window}(n) = 0.42 + 0.5 \cdot \cos\frac{2\pi \cdot n}{N} + 0.08 \cdot \cos\frac{4\pi \cdot n}{N} \qquad \text{if} \quad -\frac{N-1}{2} \le n \le \frac{N-1}{2} \qquad \text{otherwise} \quad 0$$

The spectral trends can be found in textbooks.

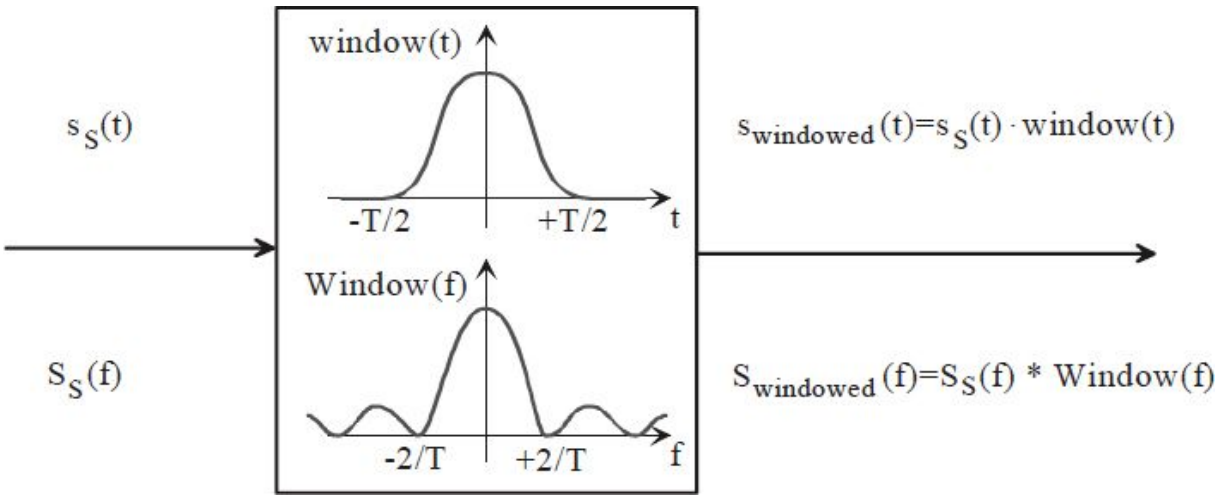


Fig. 6.27: Windowing effect on original sequence samples.

6.3. Sample and Hold

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 $\Delta V < ½LSB = 610\mu V$. Assuming to apply as an input a sinusoid with a peak-to-peak amplitude equal to the whole range of the ADC (V_{pmax}=½FSR) and to use a conversion time of T_{conv}=10μs, the maximum frequency allowed for the input signal is:

$$f_{max} = \frac{\Delta V}{T_{conv}} \cdot \frac{1}{2\pi \cdot V_{p\,max}} = 3.9Hz$$

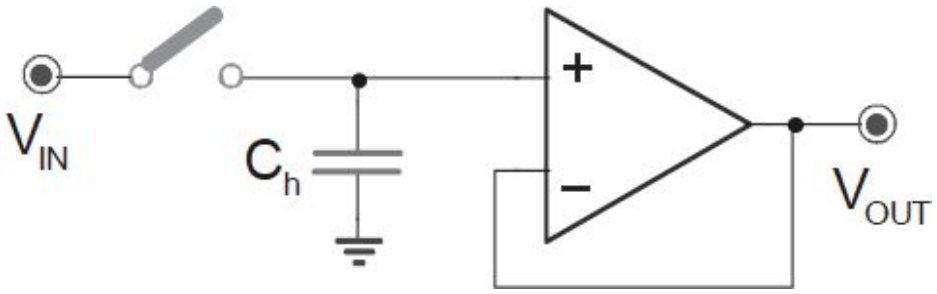


Fig. 6.28: Principle scheme of a Sample&Hold.

It is obvious that this is a very restrictive condition, much more restrictive than the theoretical limit imposed by the Sampling Theorem of $f_{max} < ½f_S = (2 \cdot T_{conv})^{-1} = 50kHz$.

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 $f_{max} = ½\ f_S$.

6.4. S&H Accuracy and Speed

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.

6.4.1 Charge-injection induced offset or pedestal error

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.

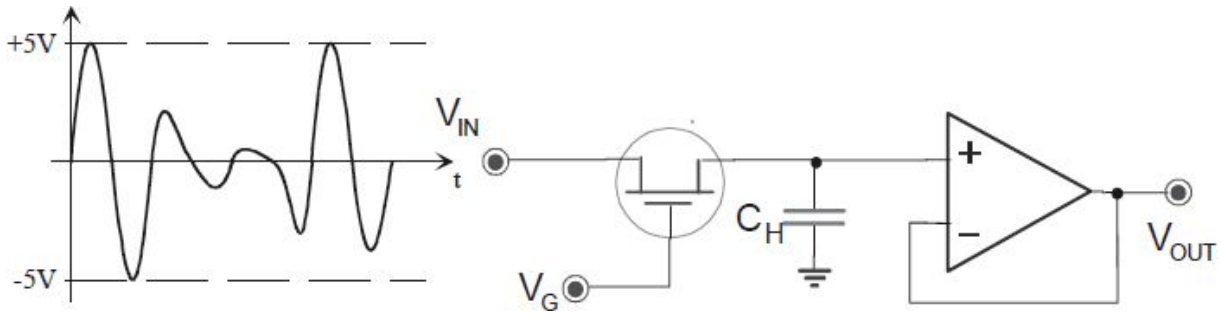


Fig. 6.29: S&H scheme.

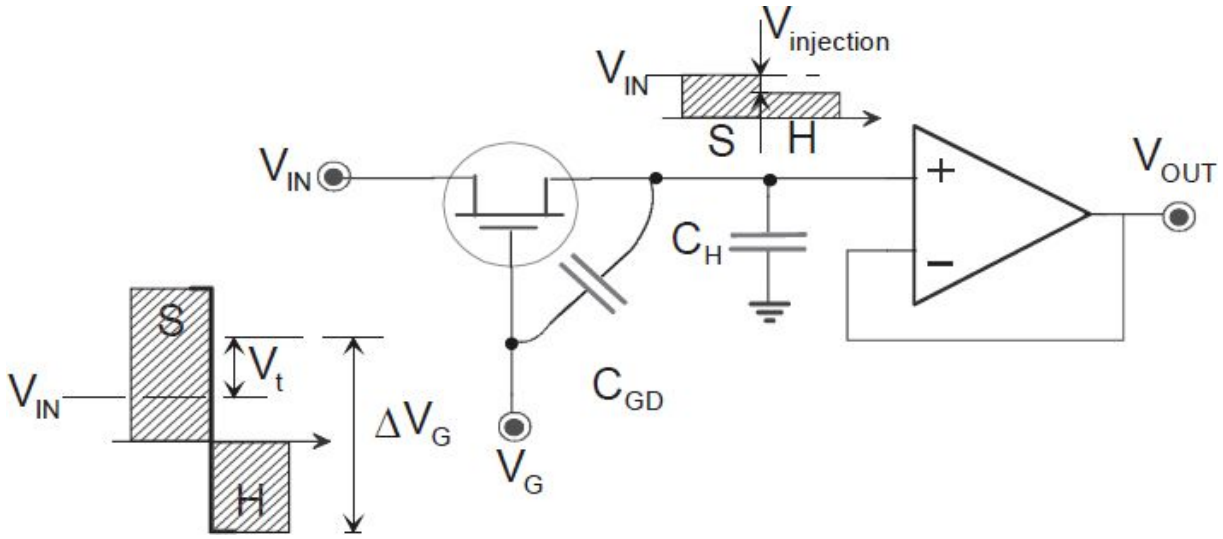


Fig. 6.30: Capacitive coupling causing a voltage induced error on the capacitance and the subsequent non-linearity of the transfer function of the S&H.

6.4.2 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 $V_{injection}=0.28mV$).

6.4.3 Signal feed-through

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:

$$V_{feedthrough} = \Delta V_{IN} \cdot \frac{C_{ds}}{C_{ds} + C_H}$$

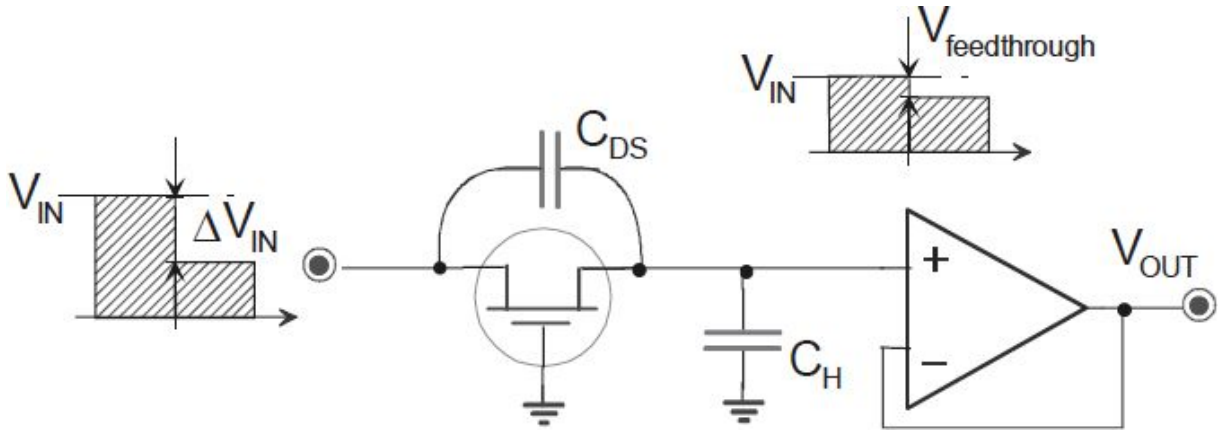


Fig. 6.31: Capacitive coupling between the input and Hold capacitance during the Hold phase.

With a capacitor of 9nF and in the worst condition, i.e. the input fluctuation equal to the whole FSR ($\Delta V_{in}=10V$), 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.

6.4.4 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.

6.4.5 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.

6.4.6 Aperture delay time

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 $V_G\cong0.2-0.7-5.7=-6.2V$ (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 $V_G \cong 15-0.7-5.7=+8.6V$ (sampling phase). Both levels are suitable to drive the S&H.

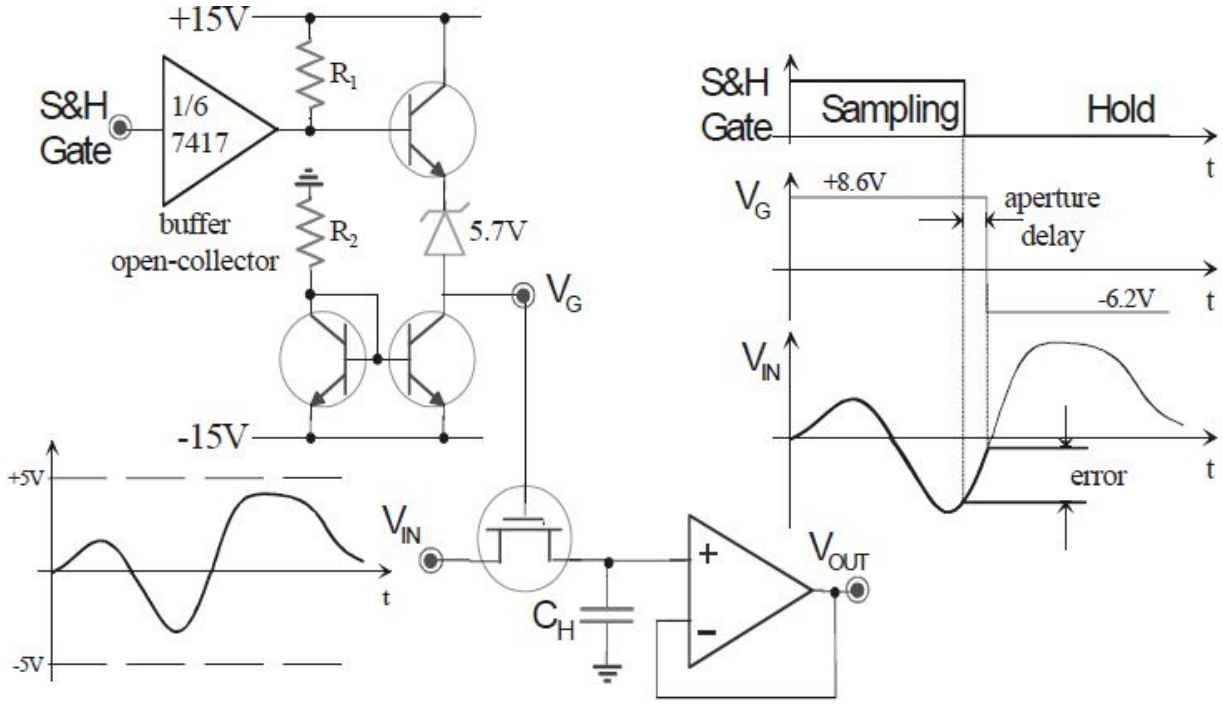


Fig. 6.32: Driving stage for the sampling command of S&H.

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).

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:

$$\Delta V_{aperture} = \left.\frac{dV_{in}}{dt}\right|_{max} \cdot T_{aperture} = 2\pi \cdot f_{max} \cdot V_{in,max} \cdot T_{aperture}$$

For example, with an average delay $T_{aperture}$=1ns, we have $\Delta V_{aperture}$=2π·20kHz·5V·1ns=628μV, close to the limit of the accuracy imposed by the specifications.

6.4.7 Aperture time jitter

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*.

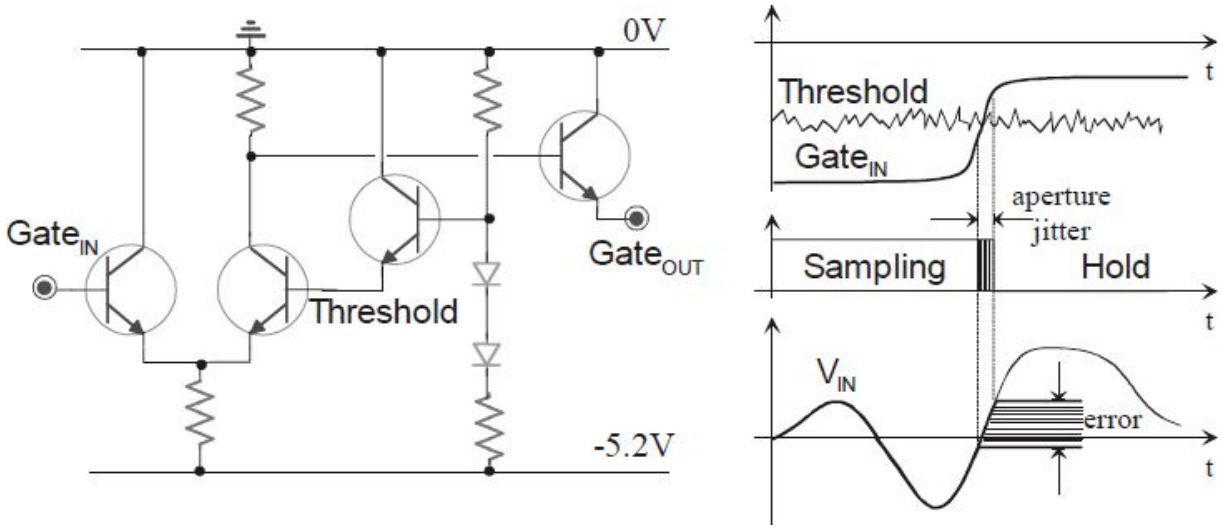


Fig. 6.33: Example of a differential stage for the sampling command.

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:

$$\sigma_{aperture} = \left.\frac{dV_{in}}{dt}\right|_{max} \cdot t_{aperture} = 2\pi \cdot f_{max} \cdot V_{in,max} \cdot t_{aperture}$$

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.

6.4.8 Acquisition time

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:

$$\left.\frac{dV_{C_H}}{dt}\right|_{max} = \left.\frac{d\left[\Delta \cdot \left(1 - e^{-t/\tau}\right)\right]}{dt}\right|_{max} = \frac{\Delta}{\tau}$$

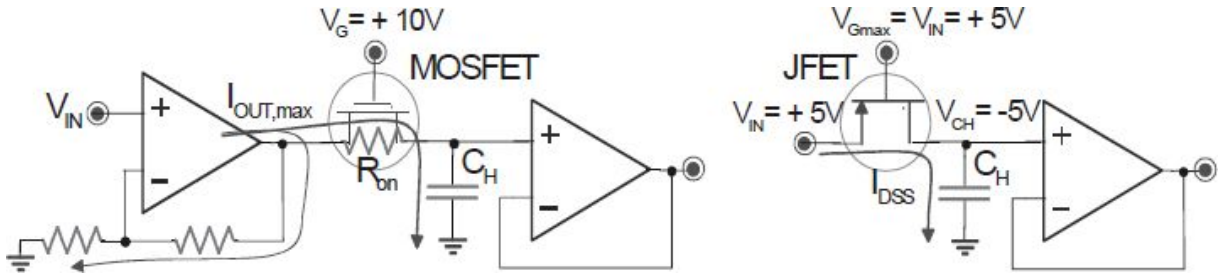


Fig. 6.34: Two cases in which the current available for the C_H is limited.

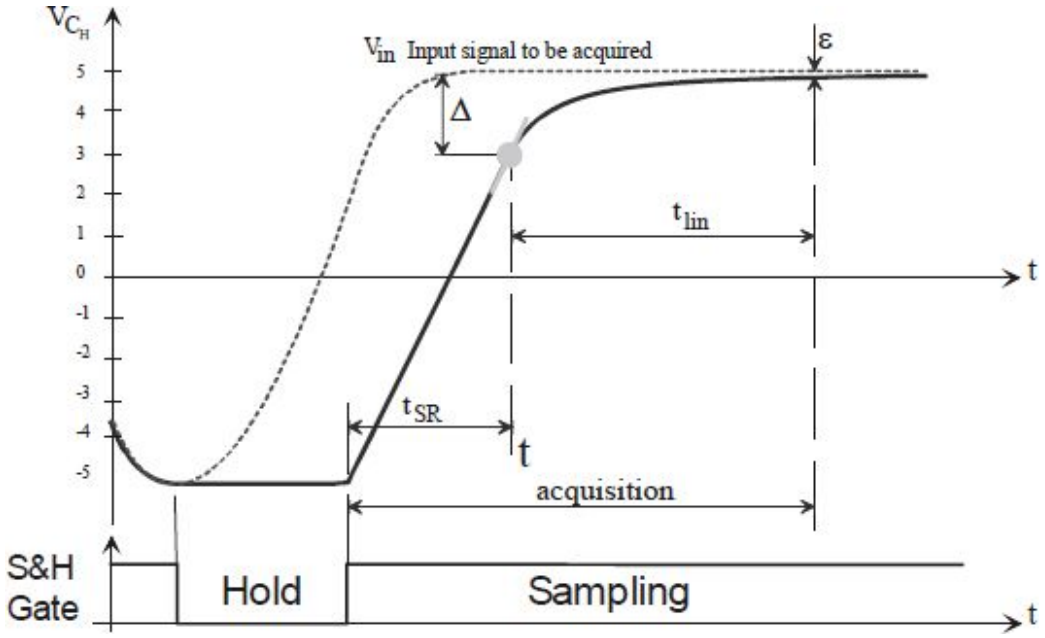


Fig. 6.35: Acquisition phase for an S&H, with the Slew-Rate limit highlighted.

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 $\Delta=50\Omega\cdot 25mA=1.25V$. 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.

6.5. Simulation Examples

To better understand the limits introduced in the preceding section, consider the circuit in Fig. 6.36 and simulate it with SPICE. The input is connected to a sinusoid with a frequency of 15kHz and peak amplitude of 5V (i.e. $10V_{pp}$). The n-MOSFET has the following characteristics: threshold voltage V_t=2.5V, parasitic capacitance Cgd=Cgs=10pF, and Sample-phase conduction resistance R_{on}=100Ω. The bulk is connected to -5V to avoid the direct conduction of the transistor parasitic diode. The gate command voltage (V_G) varies from -7.5V to +10V while the sampling frequency is f_S=65kHz, corresponding to $T_S=1/f_S$=15μs. As a buffer, we use the OpAmp LM324. In Fig. 6.36 are shown the input and command V_G voltages.

To evaluate the pedestal error, we can connect the input to the ground and overlook the voltage across the capacitance. Ideally, it should always be zero; actually, the simulation shows the trend in Fig. 6.37. When the gate is HI, the voltage is effectively 0V, but when the switch is open (V_G goes from +10V to -7.5V), the voltage decreases by -1.6mV. Such an error is caused by the charge injection on the gate command and can be estimated as:

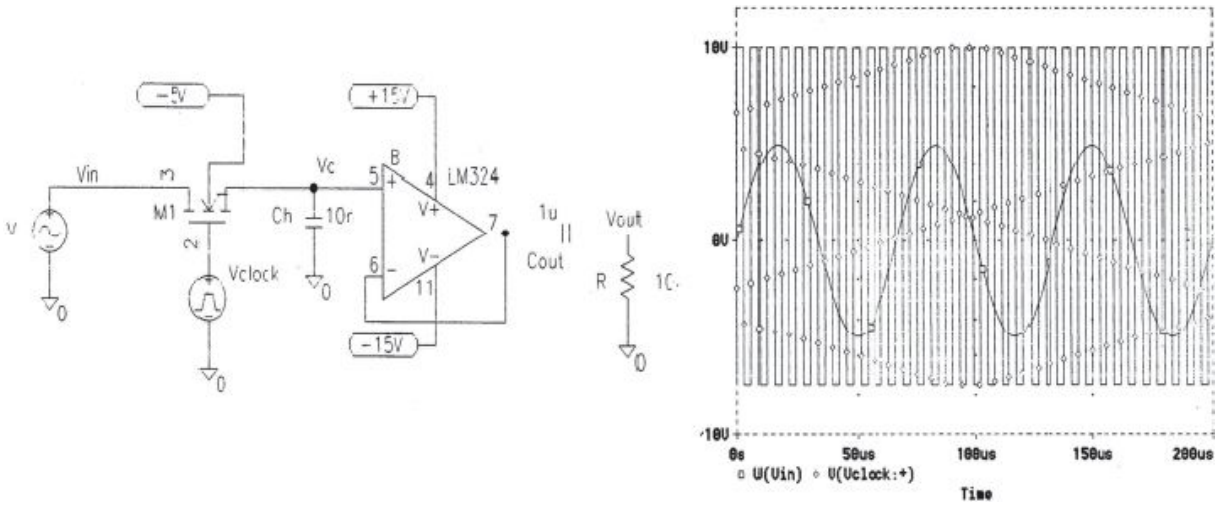


Fig. 6.36: Input voltage and command voltage for the MOSFET gate.

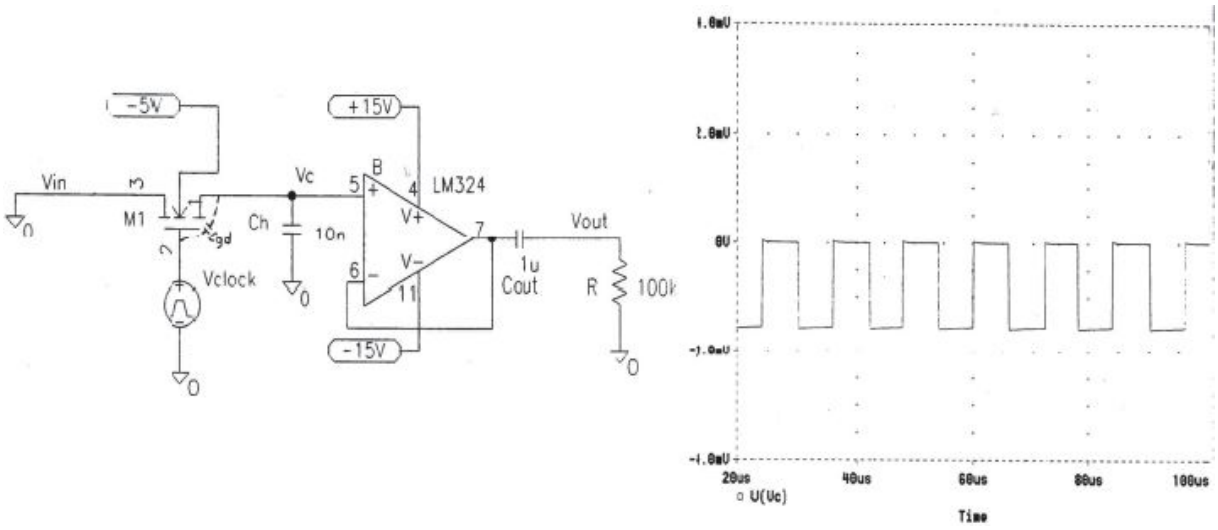


Fig. 6.37: Pedestal error induced on the capacitance on the command gate transitions, although the input is steadily connected to the ground.

$$\left|V_{injection}\right| = \Delta V_G \cdot \left(\frac{C_{gd}}{C_H + C_{gd}}\right) = 17.5V \cdot \left(\frac{10pF}{10nF + 10pF}\right) = 17.5mV$$

If one admits an error less than 0.01% of the maximum peak-to-peak input voltage (i.e. 1mV), such an error is non-negligible. To reduce this error, it should be better to use a MOSFET with very low parasitic capacitance values or use a larger C_H (at the expense of the acquisition speed, as we will see in the following).

Evaluate now the error due to the droop. In Fig. 6.38 (on the right), one can note two effects: the deep charge injection of about 1mV due to the MOSFET switching off and the droop of 67µV in 15µs, which is verified to correspond to rate of 4.5V/s.
Such an error is due to the OpAmp bias current equal to I_{bias}=45nA and to the MOSFET leakage current of 50pA; thus, the discharging rate is:

$$\left.\frac{dV_{C_H}}{dt}\right|_{droop} = \frac{I_{bias}+I_{leakages}}{C_H} = \frac{45.05nA}{10nF} = 4.5V/s$$

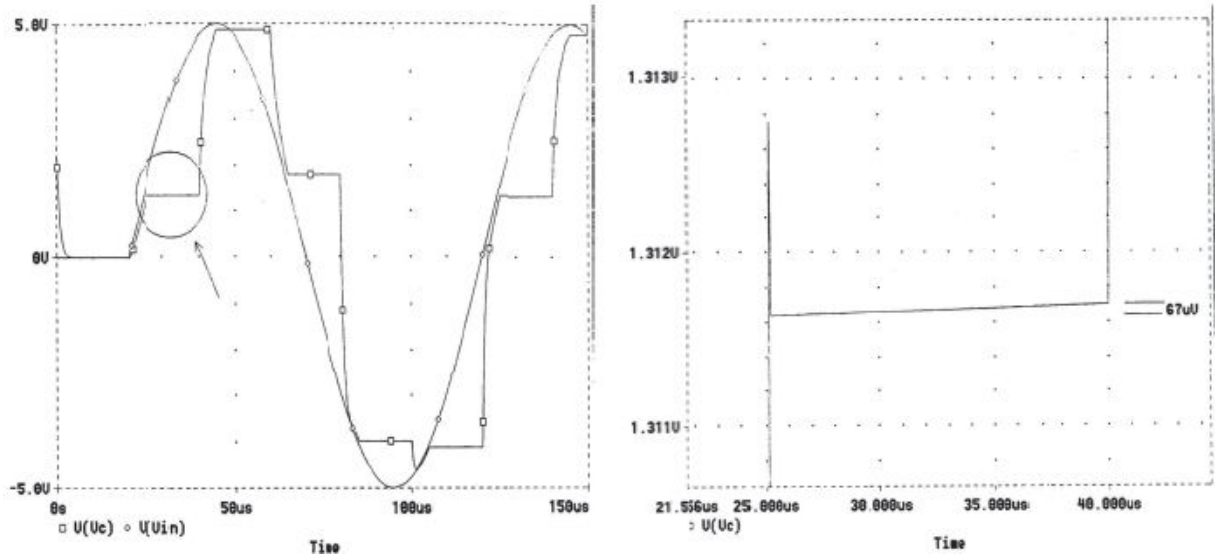


Fig. 6.38: Hold phase (on the left) and the droop (on the right). Note the spurious charge injection.

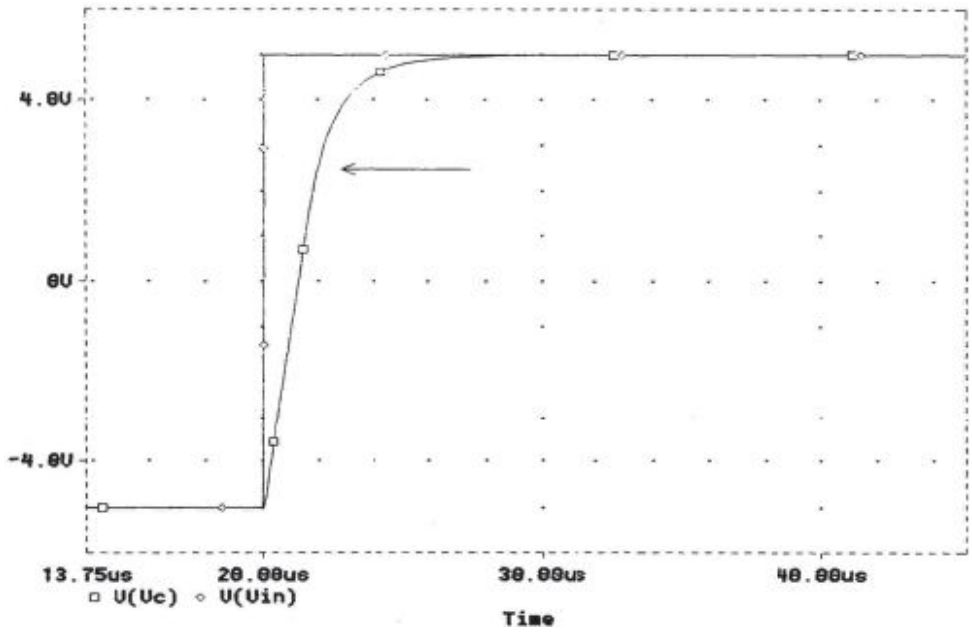


Fig. 6.39: S&H acquisition time: the first linear part of the ramp is limited by the OpAmp SR and by the OpAmp maximum output current, the second phase is exponential.

Another aspect which must be considered on an S&H is the acquisition phase duration, shown in Fig. 6.39. To reach the asymptotic value within an error range less than 0.01% of the FSR, the required time is $T_{acquisition}$=11µs. To reduce the offset shown in Fig. 6.37 by a factor of 10, we should increase C_H by the same factor (i.e. up to 100nF), but consequently the acquisition time should be 10 times greater (see the Fig. 6.35).

6.6. Improved S&H

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: $t_{acq,TOT} = \sqrt{t_{acquisition}^2 + t_{settling}^2} = \sqrt{(300ns)^2 + (300ns)^2} = 424ns$

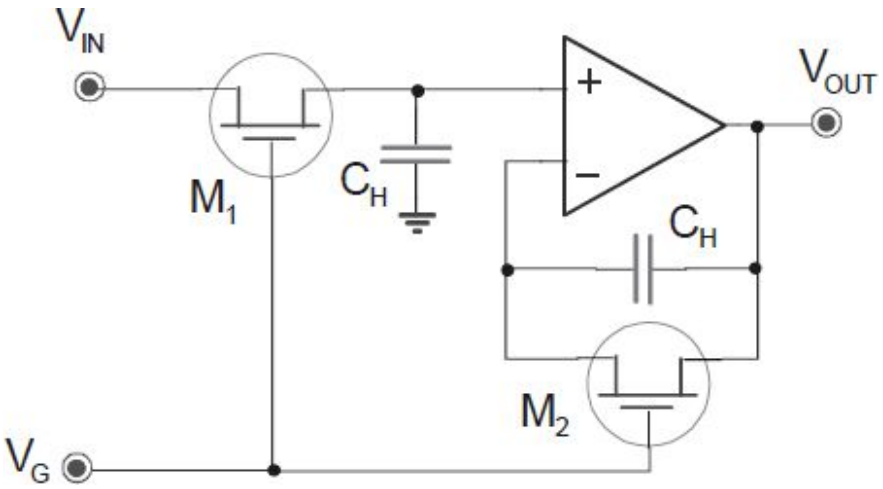


Fig. 6.40: Capacitor and transistor couples which reduce charge injection.

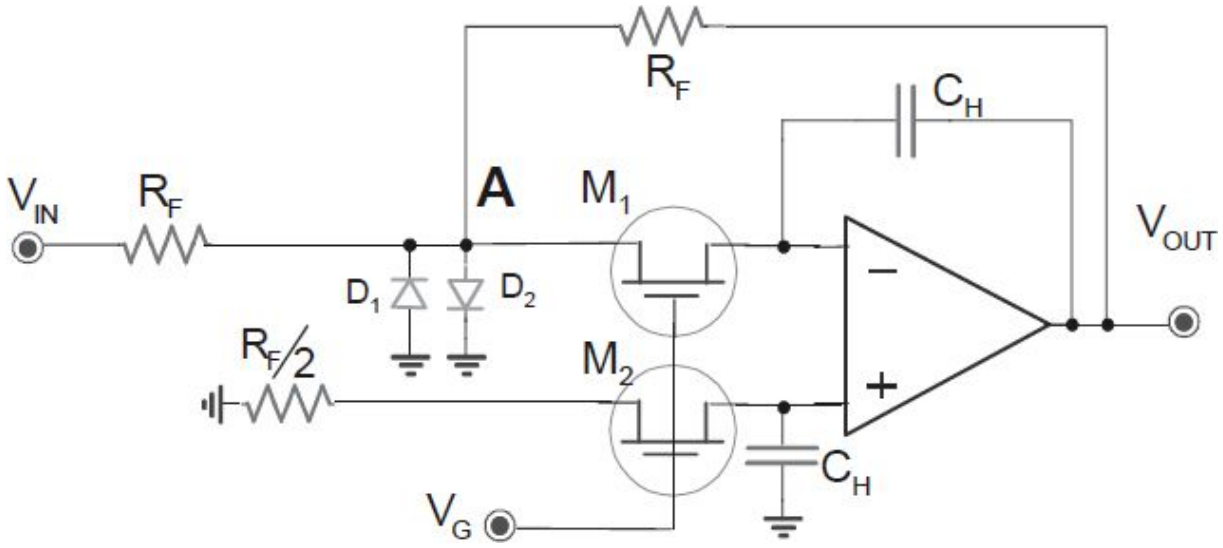


Fig. 6.41: Thanks to the virtual ground, to drive this S&H, standard logic levels can be employed whatever the input signal amplitude is.

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.

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. 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.

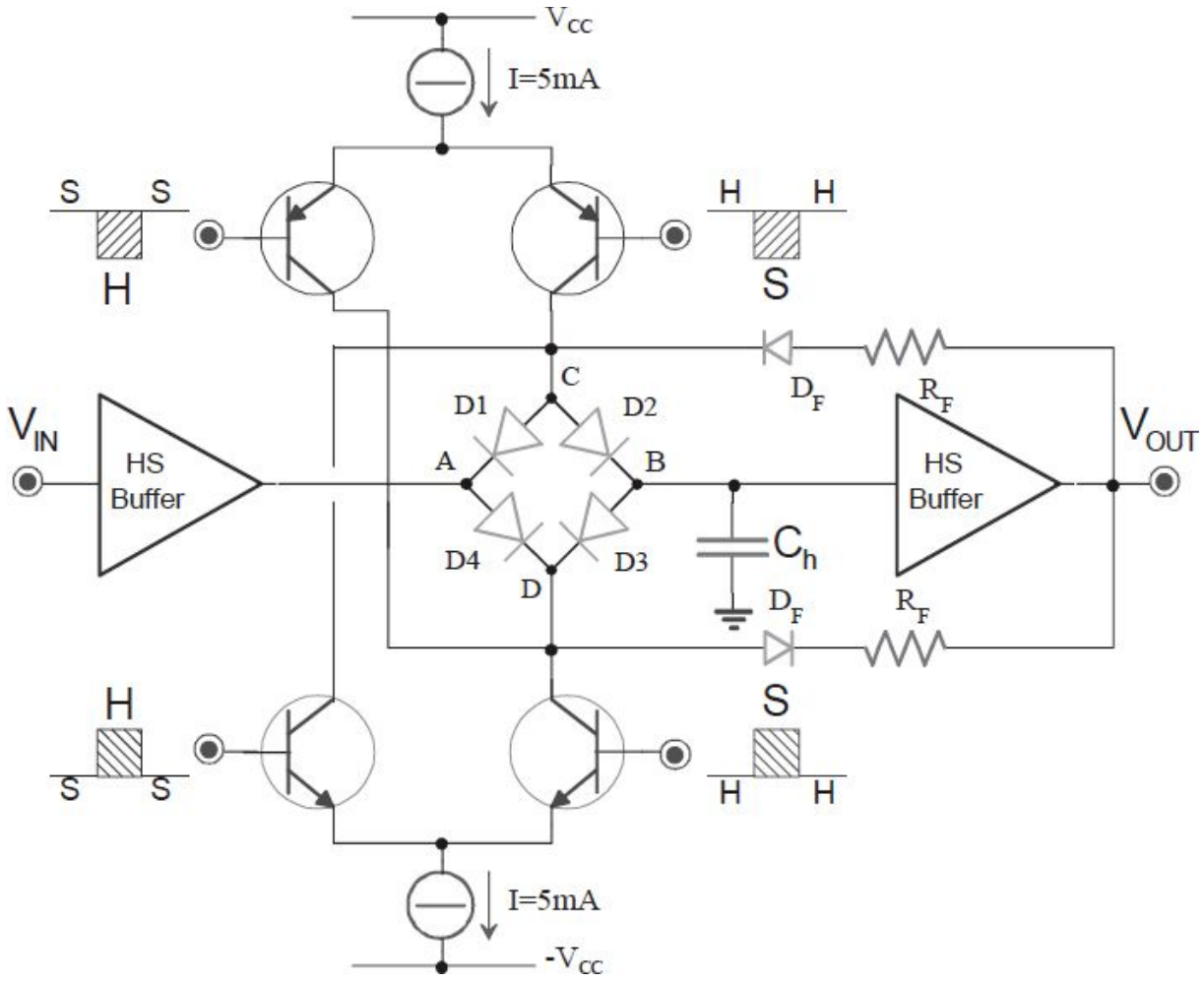


Fig. 6.42: Extremely fast S&H structure where the switch is made by a diode bridge.

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.

6.7. FEEDB 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.

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 $\beta=1$ 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 $1/\beta=1+R_2/R_1$.

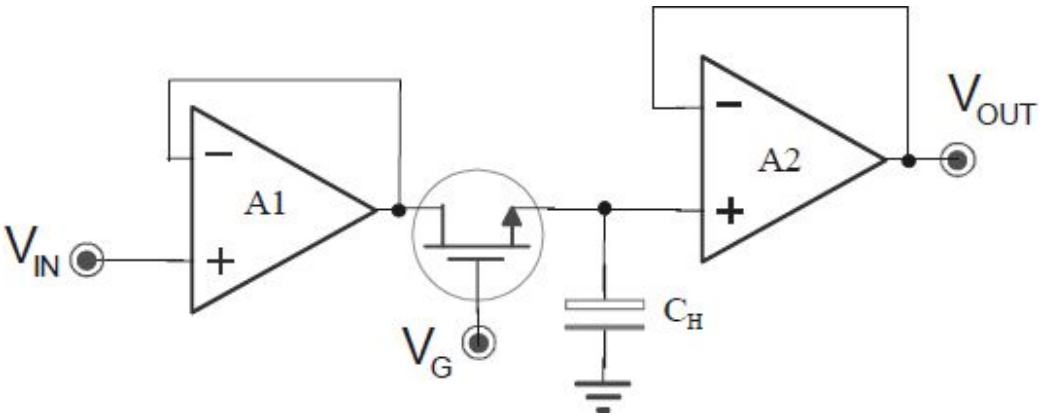


Fig. 6.43: S&H with an input buffer to improve the acquisition speed.

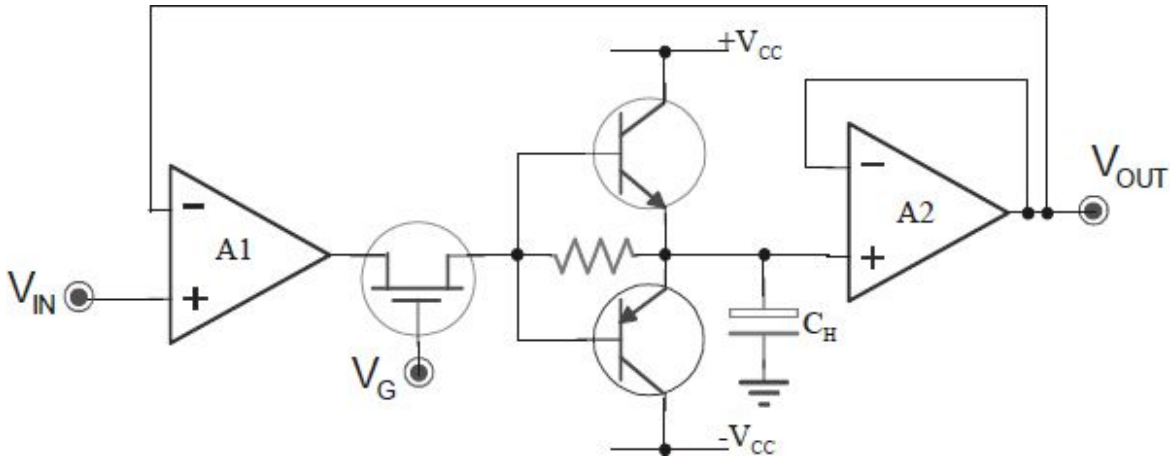


Fig. 6.44: Feedback configuration S&H of Fig. 6.43.

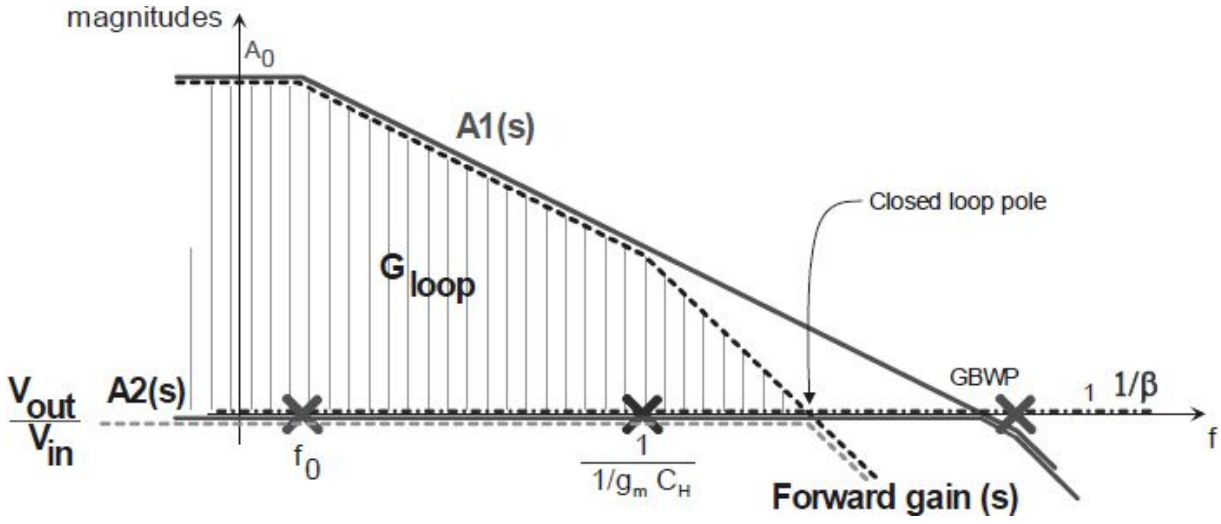


Fig. 6.45: Bode diagram for the stability analysis of S&H shown in Fig. 6.44.

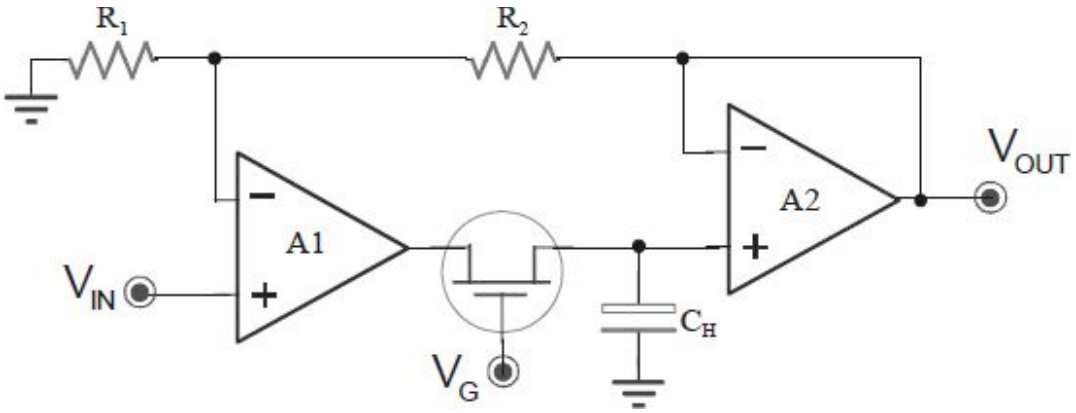


Fig. 6.46: Feedback S&H with gain.

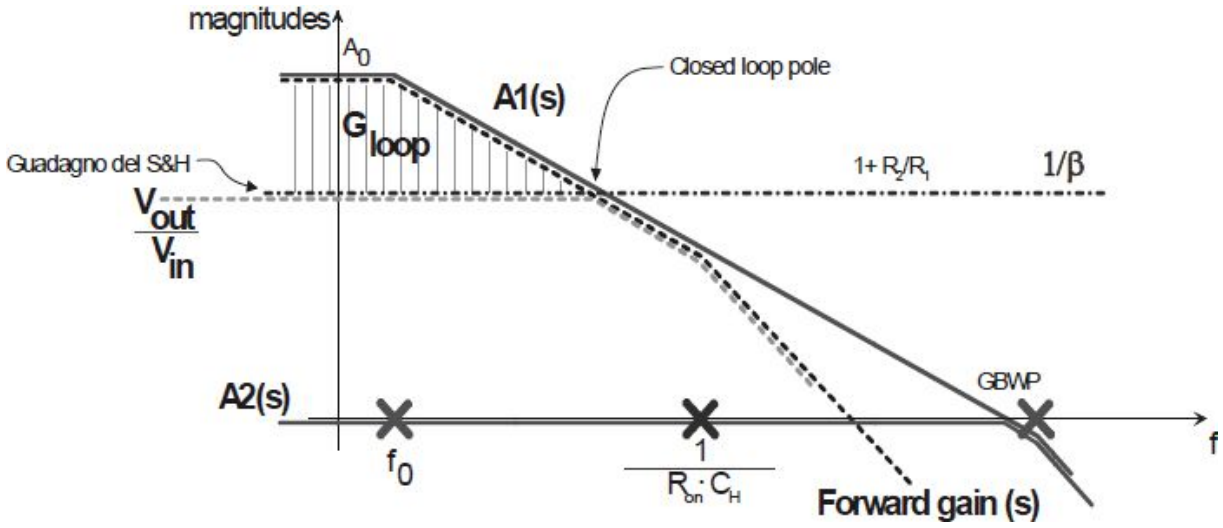


Fig. 6.47: Bode diagram and the stability analysis of the S&H shown in Fig. 6.46.

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}.

For example, in Fig. 6.49, the first page of the data-sheet of a commercial S&H, the SHC5320 Burr-Brown, is shown. We can see that this S&H allows

implementing various circuital configurations from the internal structure, some of which are shown in Fig. 6.48.

The data-sheet frontispiece of another commercial Track&Hold (T&H - SHC605, Burr-Brown) is shown in Fig. 6.50. Fig. 6.51 shows the trend of the most important characteristics. The internal simplified scheme for a T&H is shown in Fig. 6.52: it is possible to identify various blocks and the Hold capacitance values, doubled to compensate the charge injection error and the droop and to improve the T&H stability.

In addition to the common inverting and non-inverting connections, shown on the left and in the center of Fig. 6.53, this T&H can be connected in a differential configuration to sample a double-ended input and to translate it into a single-ended output, as shown in Fig. 6.53, on the right.

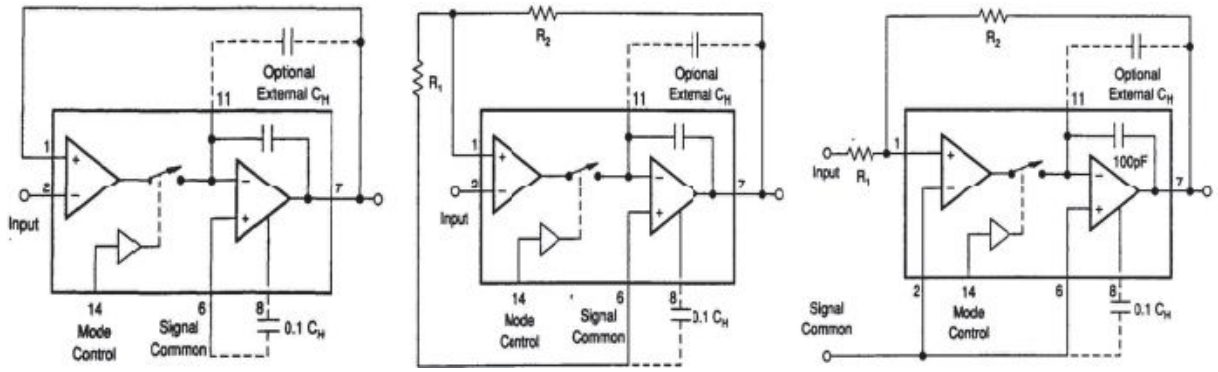


Fig. 6.48: Some configurations for a commercial S&H (SHC5320, Burr-Brown): non-inverting with a unity gain (on the left); non-inverting with a gain $1+R_2/R_1$ (in the center); inverting with a gain $-R_2/R_1$ (on the right).

BURR - BROWN®

SHC5320

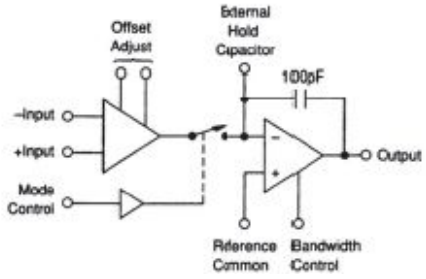
High-Speed Bipolar Monolithic SAMPLE/HOLD AMPLIFIER

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

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.

Fig. 6.49: Data-sheet frontispiece of a commercial S&H (SHC5320, Burr-Brown), with the most important characteristics and the internal structure.

BURR - BROWN®

SHC605

DEMO BOARD AVAILABLE
See Appendix A

High-Speed Operational TRACK-AND-HOLD AMPLIFIER

FEATURES

- VERY GOOD SPURIOUS FREE DYNAMIC RANGE:
  90dB at 1MHz F_{IN} and 20MSPS
  86dB at 2MHz F_{IN} and 20MSPS
  77dB at 5MHz F_{IN} and 20MSPS
- LOW ACQUISITION TIME: 30ns to 0.01%
- LOW DROOP RATE: 8mV/µs max T_{MIN} to T_{MAX}
- LOW POWER CONSUMPTION: 335mW
- EXTREMELY VERSATILE ARCHITECTURE: Noninverting, Inverting, and Differential Gains
- LOGIC FLEXIBILITY: TTL and ECL Compatible
- SMALL PACKAGE: 16-Lead SOIC
- EXTENDED TEMPERATURE SPECS: –40°C to +85°C

APPLICATIONS

- A/D CONVERTER FRONT ENDS
- MULTIPLE CHANNEL SIMULTANEOUS SAMPLING
- IMPROVING FLASH ADC PERFORMANCE
- PEAK DETECTORS
- DAC DEGLITCHING

DESCRIPTION

The SHC605 is a monolithic hig… racy track-and-hold amplifier. It co… sition and low distortion to provid… tion for a wide range of sampling applications. Its proprietary closed-loop architecture provides a single-chip solution to many data acquisition problems formerly requiring more than one device. Noninverting, inverting, and differential gain configurations are easy to apply with the SHC605. An on-board logic reference circuit makes the SHC505 compatible with both single-ended and differential ECL or TTL clock inputs. An internal track-mode lockout circuit allows edge-triggered operation in data acquisition systems. The SHC605 is available in a 16-lead SOIC package specified for the –40C to +85C industrial temperature range.

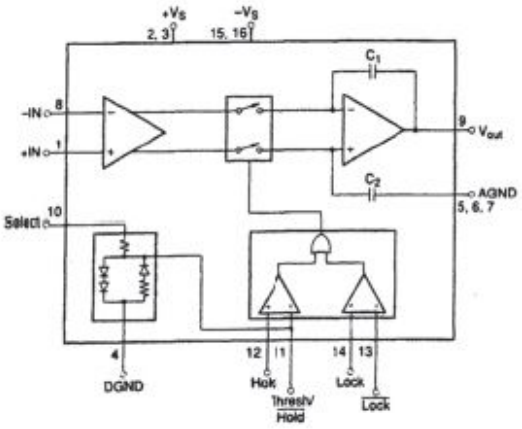
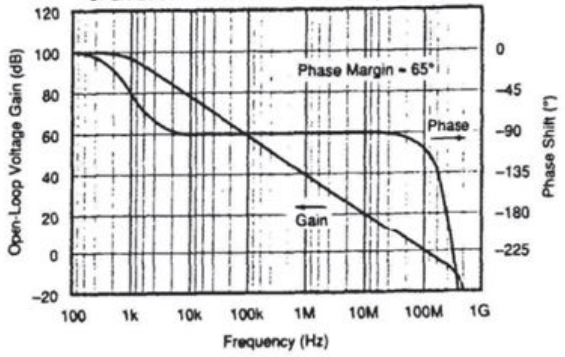


Fig. 6.50: Data-sheet frontispiece of the T&H SHC605, Burr-Brown.

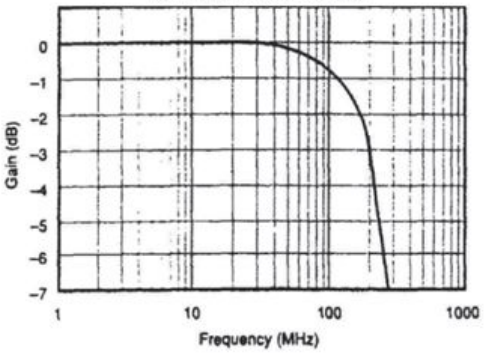
TYPICAL PERFORMANCE CURVES

At T_A = +25°C, $\pm V_S$ = ±5V, G = +1V/V, R_L = 100Ω, C_L = 5pF, and ECL Hold/$\overline{\text{Hold}}$ inputs, unless otherwise noted.

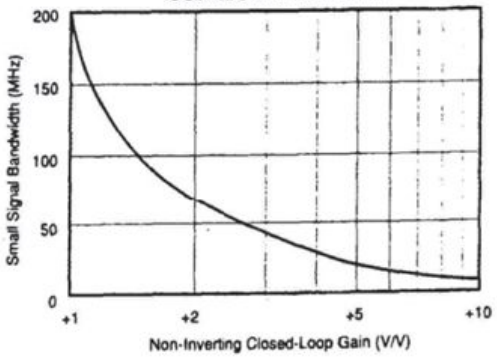
OPEN-LOOP FREQUENCY RESPONSE (Track Mode)



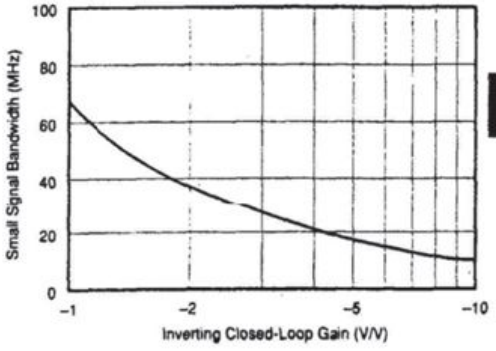
CLOSED-LOOP GAIN = +1V/V
vs FREQUENCY (Track Mode)



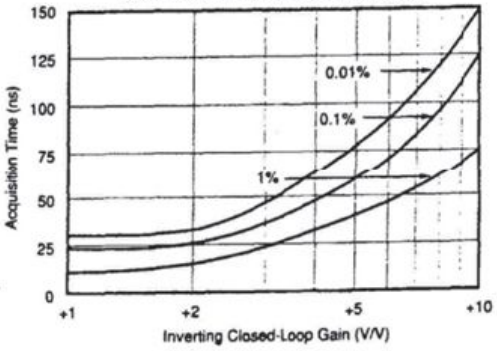
NON-INVERTING CLOSED-LOOP GAIN
vs SMALL SIGNAL BANDWIDTH



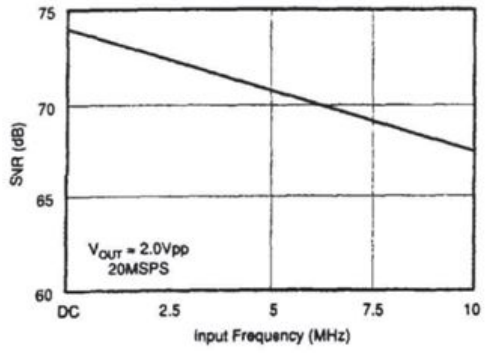
INVERTING CLOSED-LOOP GAIN
vs SMALL SIGNAL BANDWIDTH



ACQUISITION TIME
vs NON-INVERTING CLOSED-LOOP GAIN



SNR vs INPUT FREQUENCY



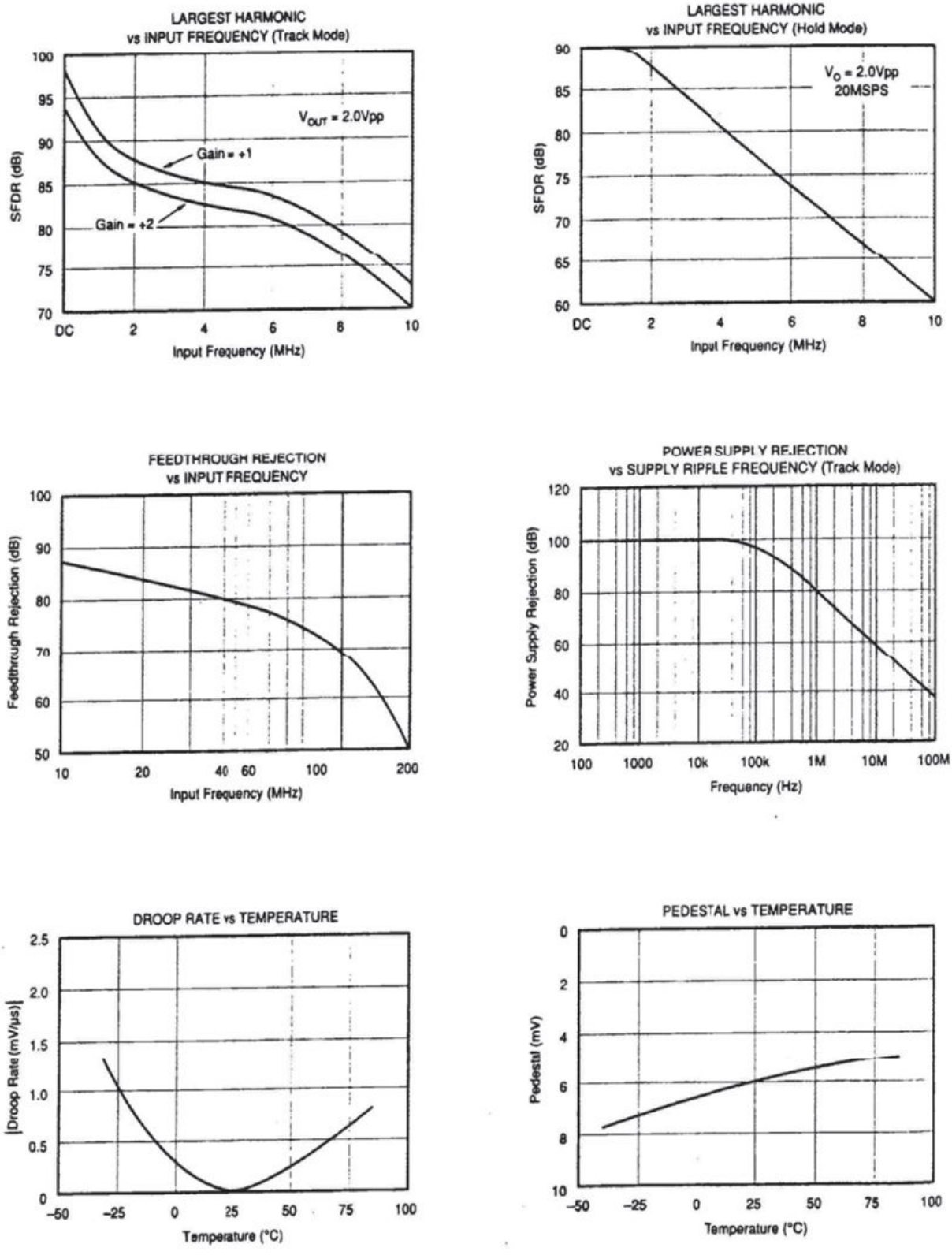


Fig. 6.51: T&H Burr-Brown SHC605 characteristics.

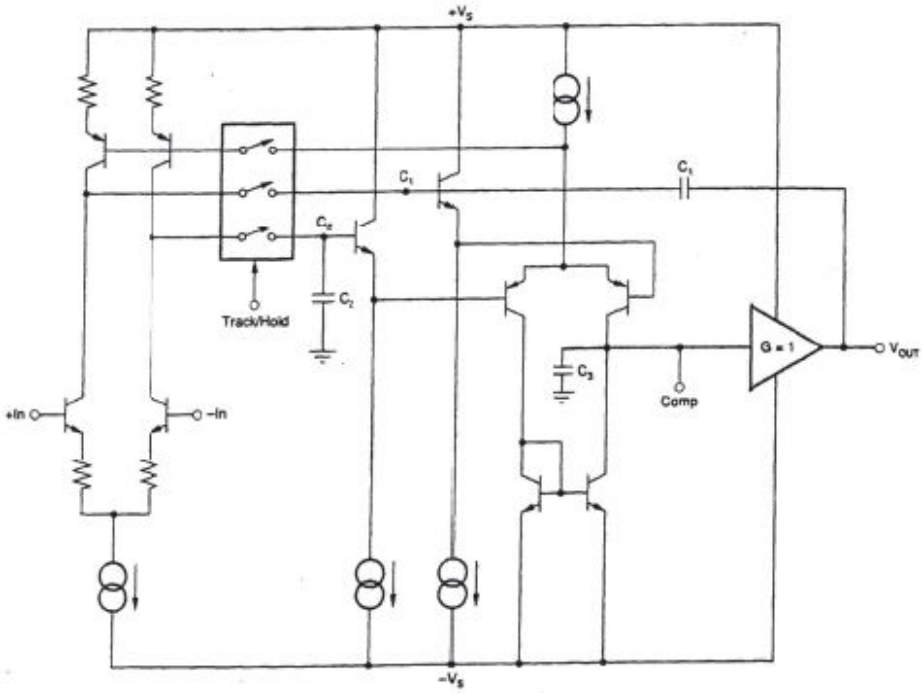


Fig. 6.52: Internal structure of the T&H SHC605, Burr-Brown.

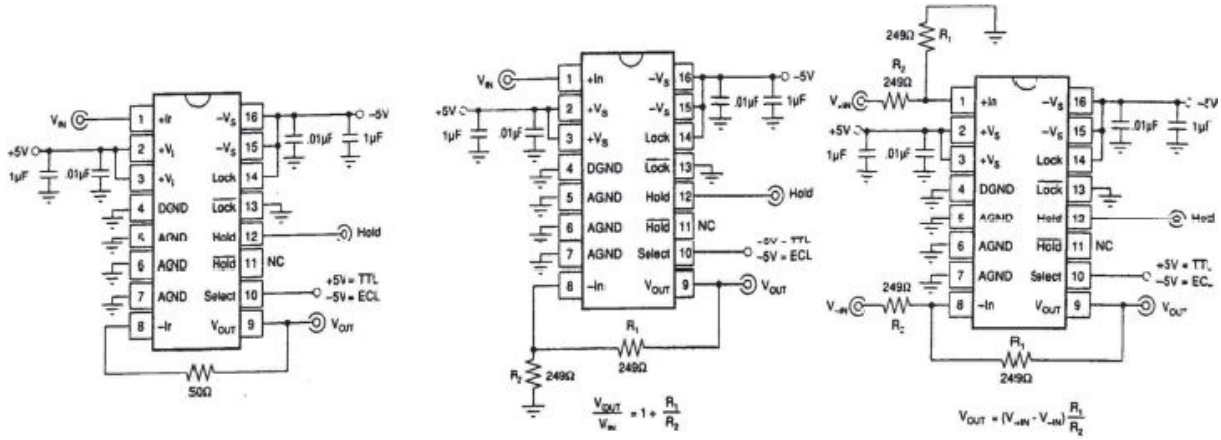


Fig. 6.53: T&H configurations with a gain +1 (on the left) and +2 (in the center), and differential input conversion to a single ended output (on the right), with the SHC605, Burr-Brown.

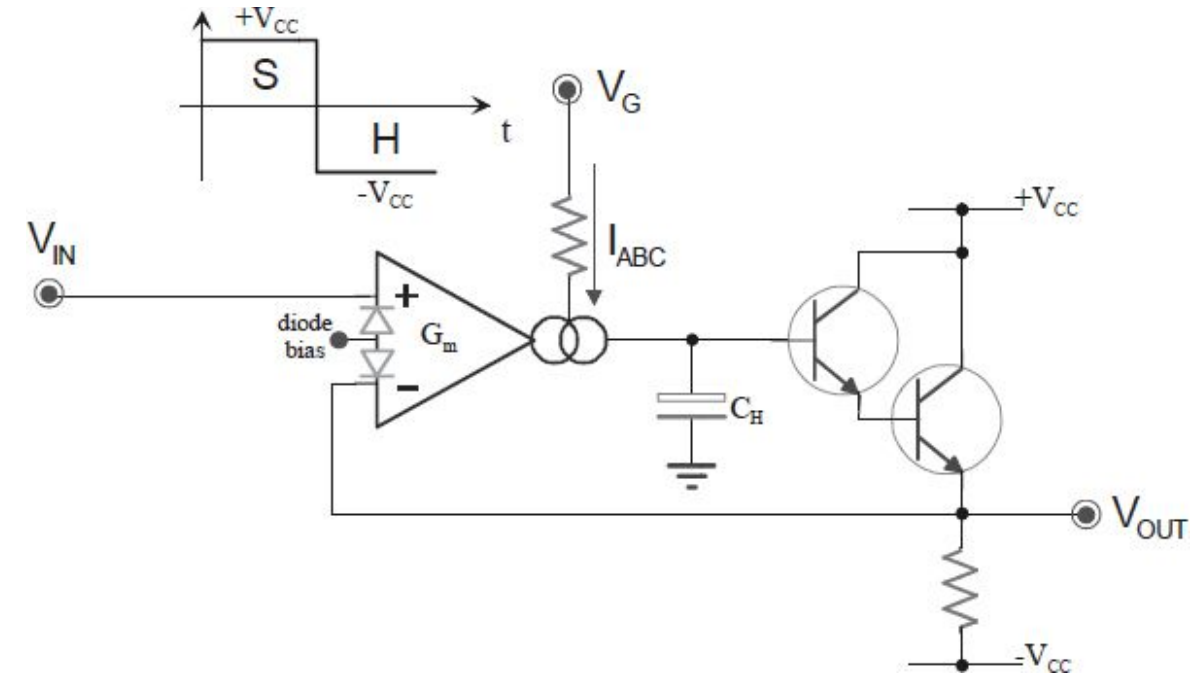


Fig. 6.54: S&H realized with an OTA and an output buffer.

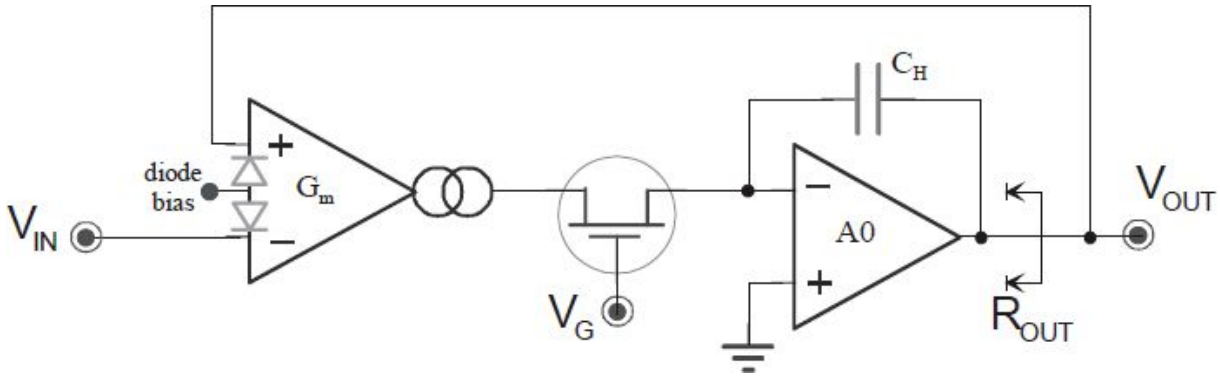


Fig. 6.55: Another S&H with an OTA, but with C_H on feedback on the second OpAmp.

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.

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 $C_{Hold}=50{,}000pF=50nF$.

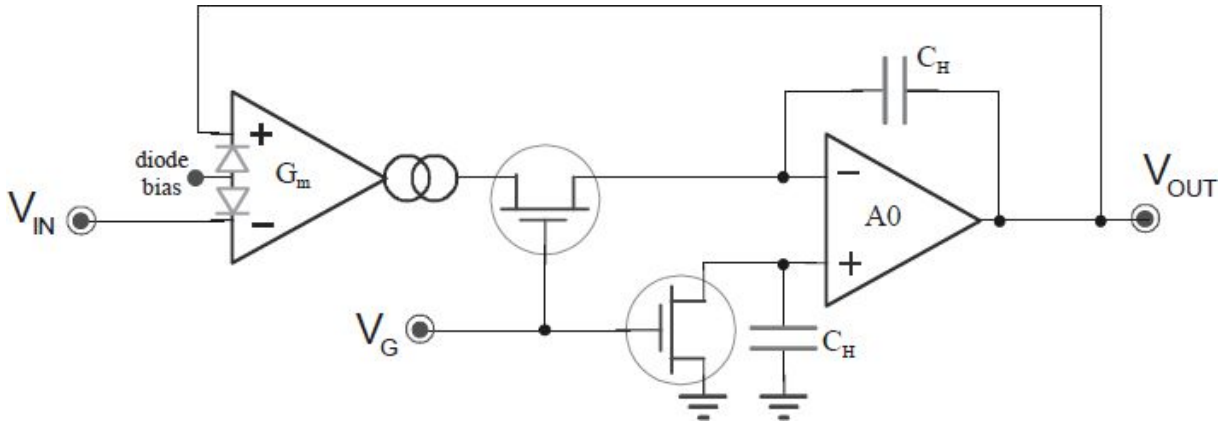


Fig. 6.56: Charge-injection effect reduction by means of a differential compensation.

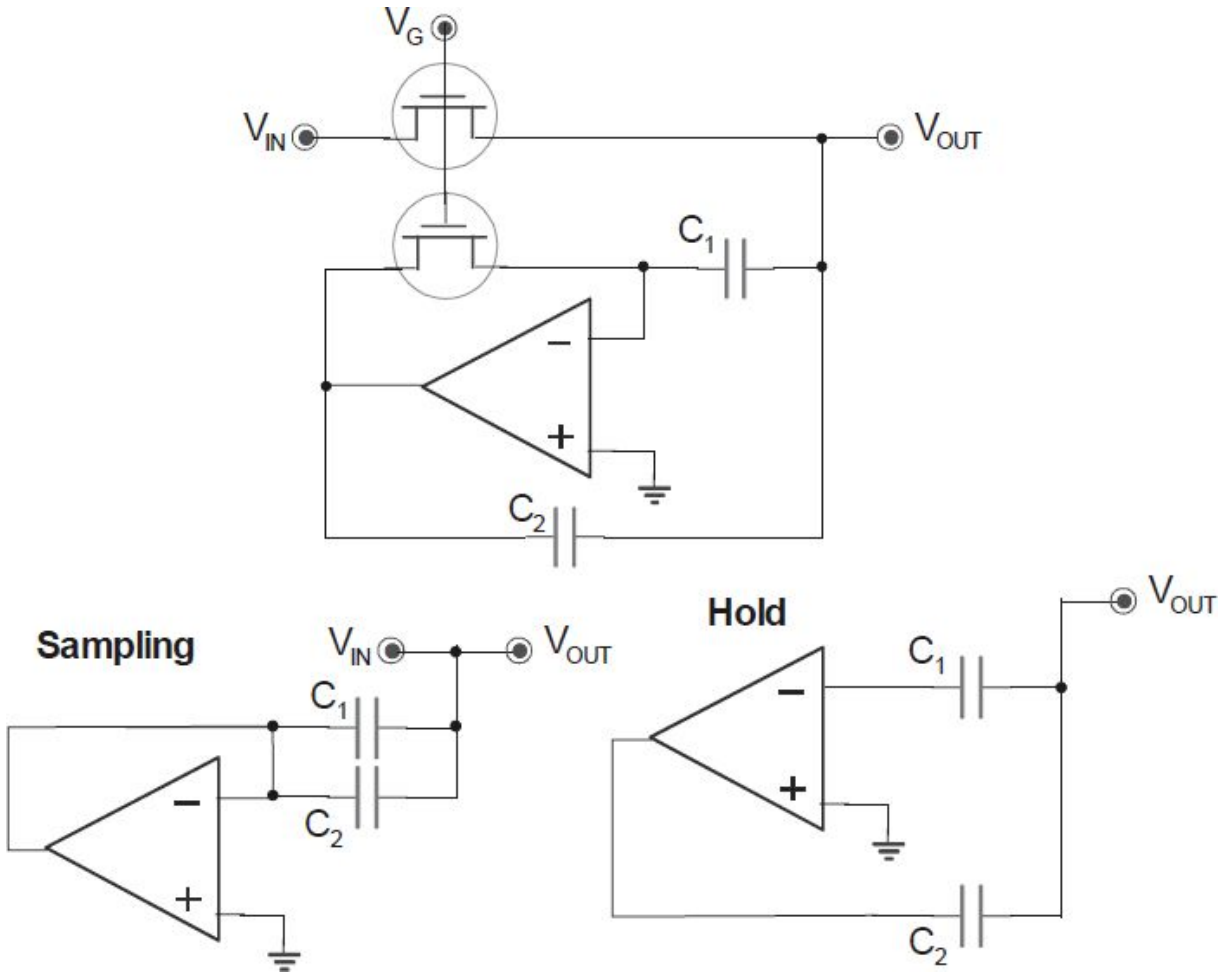


Fig. 6.57: Thanks to the feedback, this circuit features a hold capacitance much larger than the one experienced during the sampling phase.

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}=1µA determines a V_{out} discharge with a speed $dV_{out}/dt=I/C_{Hold}$=20V/s=20µV/µs. 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 $\Delta t = \Delta V \cdot C_2 / I = 25\mu s$.

6.8. Digital Potentiometer

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.

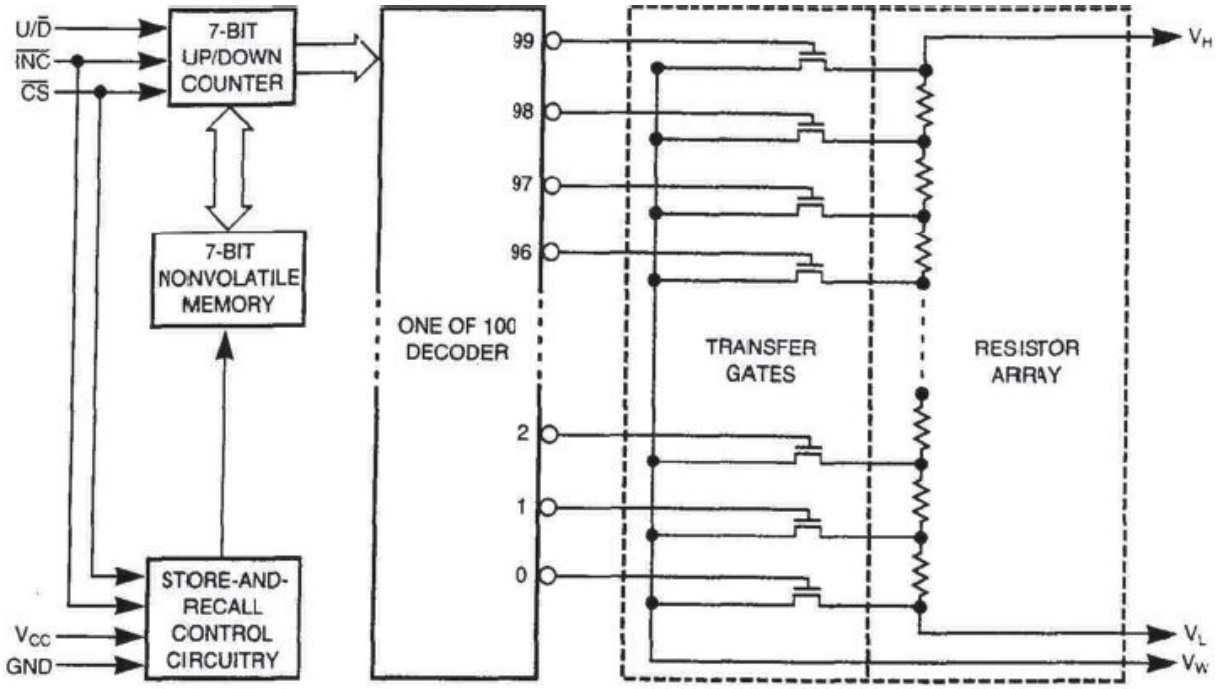


Fig. 6.58: DigPot internal structure (Xicor Inc. X9312).

Company	Model	Steps	Resistance (kΩ)	Configur-ation	Non-volatile	Packages	Interface	Operating power	Features	Price (1000)
Analog Devices Inc	AD8402	256	10, 50, 100	Dual	No	DIP-14, SO-14	Three-wire	2.7 to 5.5V, 5 µA	Full ac specs, nA shutdowncurrent	$1.66
	AD8403	256	10, 50, 100	Quad	No	SOL-24	Three-wire	2.7 to 5.5V, 5 µA	Full ac specs, nA shutdowncurrent	$2.51
Dallas Semi-conductor Corp	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
Xicor Inc	X9511	32	1, 10	Single	Yes	DIP-8, SO-8	Pushbutton	5V, 3 mA 200 µA standby	Auto-storeor user-store wipe; X9514 is a log-taper version	$0.85
	X9221	64	2, 10, 50	Dual	Yes	DIP-20, SO-20	Two-wire (I^2C-compatible)	5V, 2 mA 500 µA standby	Multiple wper settings, can be cascaded	$2.80
	X9312	100	1, 2, 10, 50	Single	Yes	DIP-8, SO-8	Three-wire	5V, 3 mA	Up to 0 to15V across terminals, up/down, clock selec 1-mA shut-down	$1.45

Fig. 6.59: Comparison between some commercial digital potentiometers.

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.

6.9. Analog Multiplexers

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.

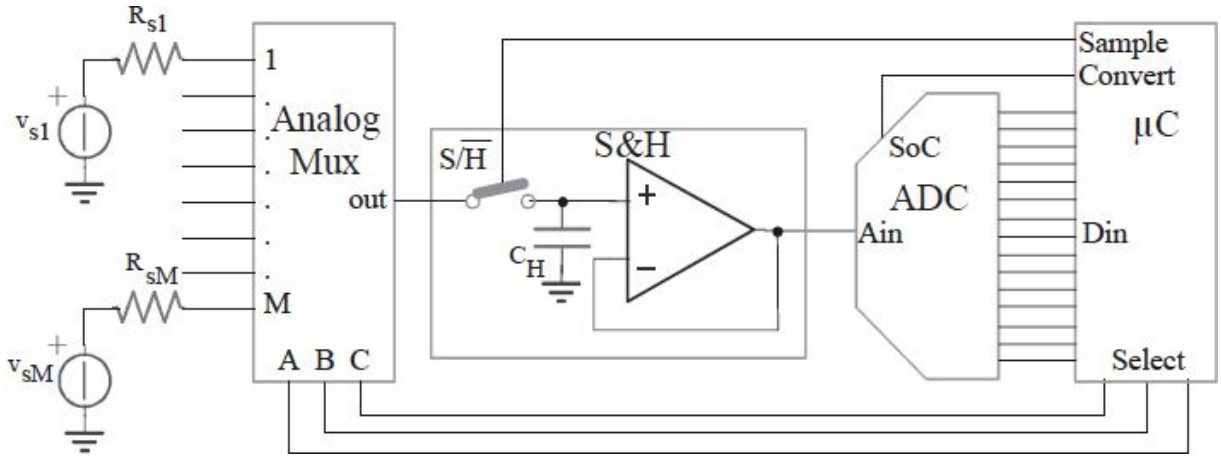


Fig. 6.60: Typical acquisition circuit with many inputs and separate analog channels.

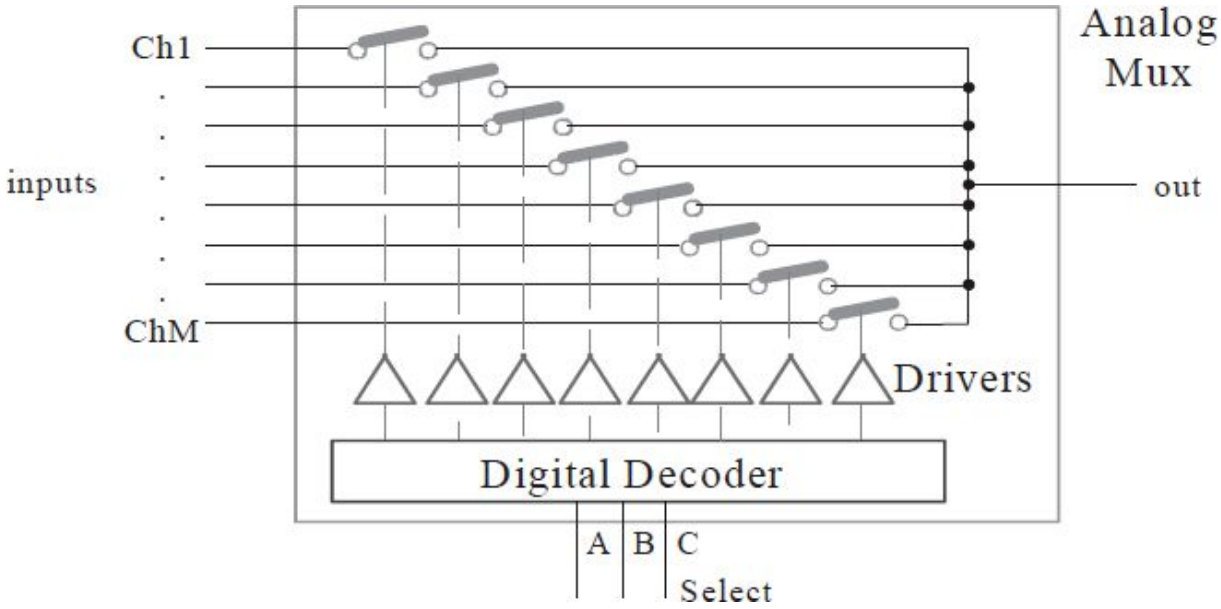


Fig. 6.61: Analog multiplexer internal structure.

19-0596; Rev 0; 5/95

MAXIM

Low-Voltage, 8-Channel/Dual 4-Channel Multiplexers with Latchable Inputs

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

TOP VIEW

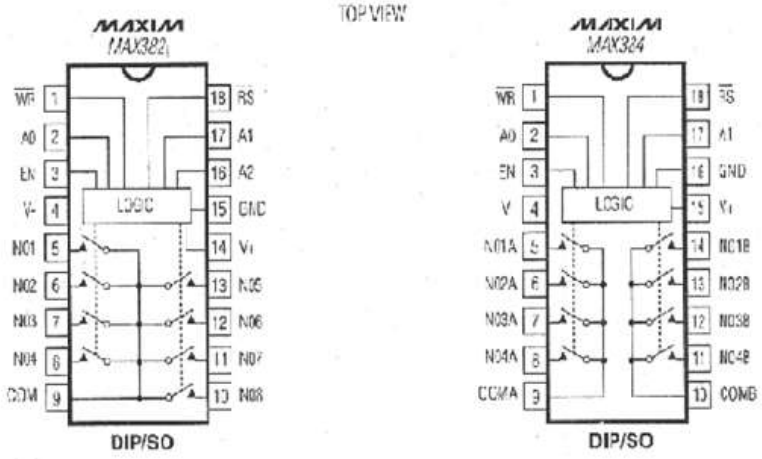


Fig. 6.62: Commercial analog mux characteristics.

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, 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:

$$\Delta V = \pm v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{ R_{in} \left\| \left[R_{off} / (M-1) \right] \right. \right\}} \pm (I_{off} \cdot M \pm I_b) \cdot \left\{ (R_s + R_{on}) \left\| R_{in} \right\| \left[R_{off} / (M-1) \right] \right\}$$

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:

$$v_s \cdot \frac{R_s + R_{on}}{R_s + R_{on} + \left\{ R_{in} \parallel \left[R_{off} / (M-1) \right] \right\}} = \left[I_{off} \cdot M + I_b \right] \cdot \frac{(R_s + R_{on}) \cdot \left\{ R_{in} \parallel \left[R_{off} / (M-1) \right] \right\}}{R_s + R_{on} + \left\{ R_{in} \parallel \left[R_{off} / (M-1) \right] \right\}}$$

i.e. when:

$$R_{in} \parallel \left[R_{off} / (M-1) \right] = \frac{v_s}{I_{off} \cdot M + I_b}$$

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:

$$\Delta V_{max} \approx \left| v_{s,max} \right| \cdot \frac{R_s + R_{on}}{R_s + R_{on} + R_{off}/M} + \left| I_{off} \right| \cdot M \cdot (R_s + R_{on})$$

With the preceding values, we obtain $\Delta V_{max} \approx 17mV+1.1mV$, in accordance with the results acquired by the most precise computations.

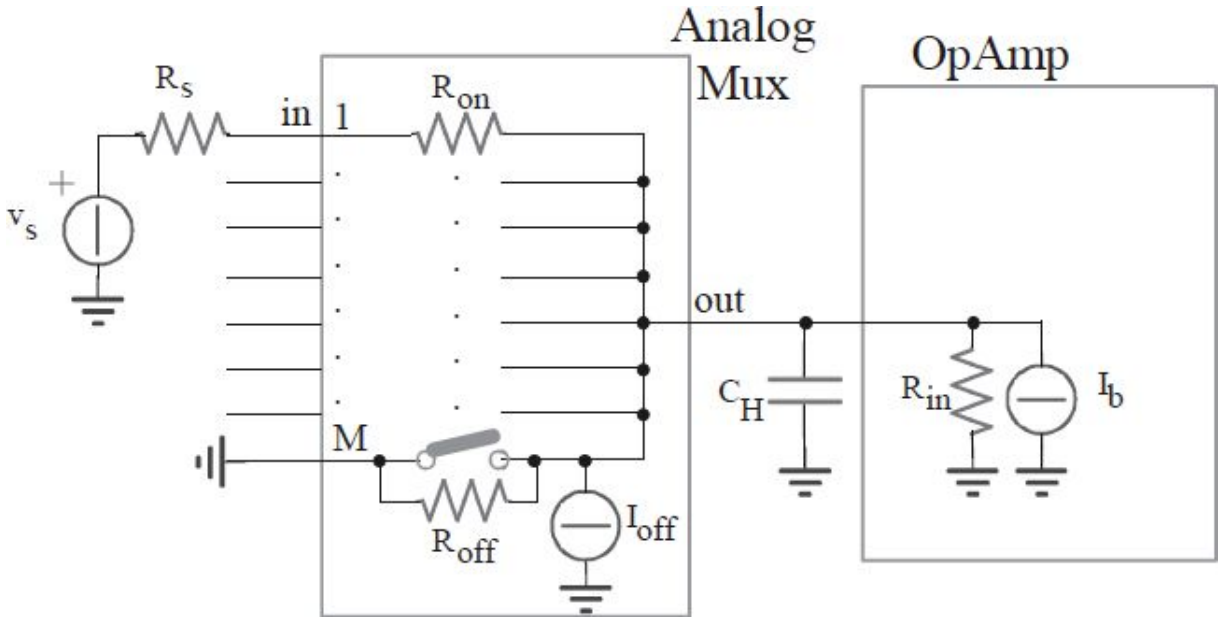


Fig. 6.63: Analog mux equivalent circuit for the computation of the output error.

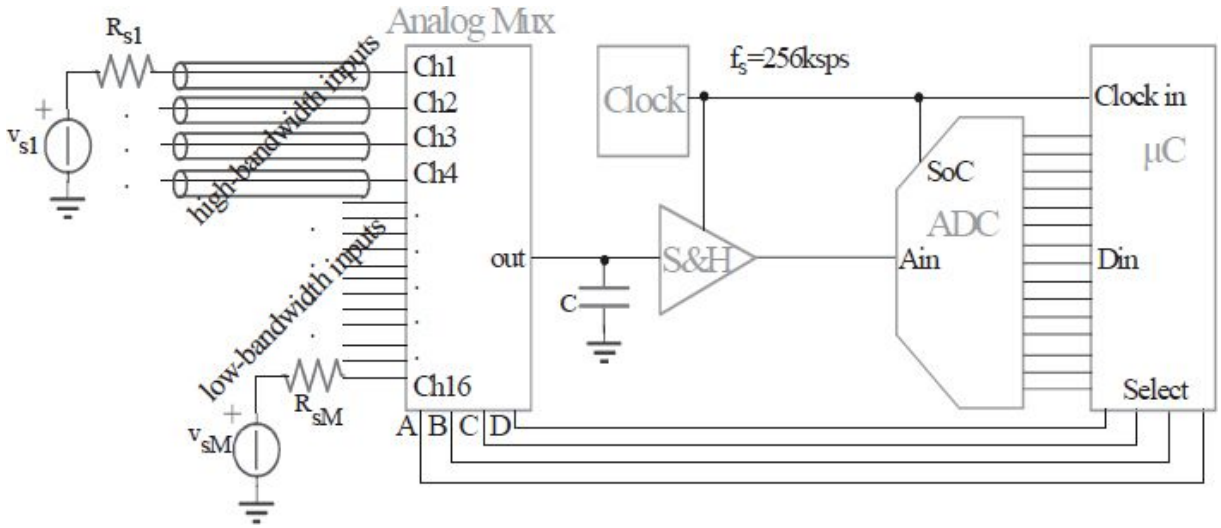


Fig. 6.64: Evaluation of the time required to acquire many analog multiplexed inputs.

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.

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:

$$\tau \le \frac{T}{\ln\frac{FSR}{0.5\cdot LSB}} = \frac{T}{\ln\left(2\cdot 2^n\right)}$$

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.

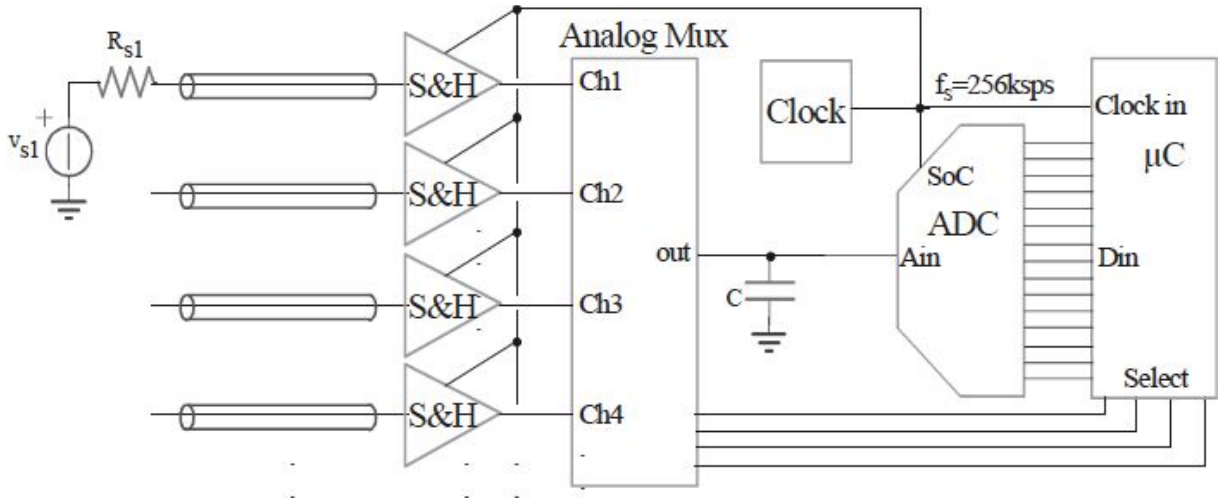


Fig. 6.65: Many S&Hs used ensures the simultaneous sampling of multiple inputs.

MOTOROLA
SEMICONDUCTOR
TECHNICAL DATA

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

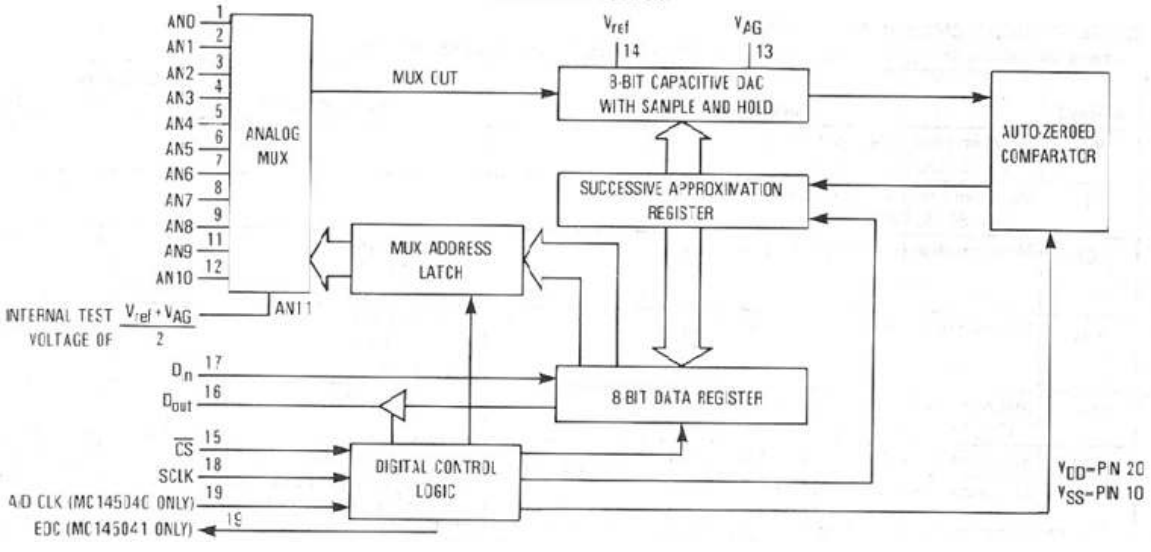


Fig. 6.66: Commercial ADC example with an 8 inputs mux and an S&H.

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.

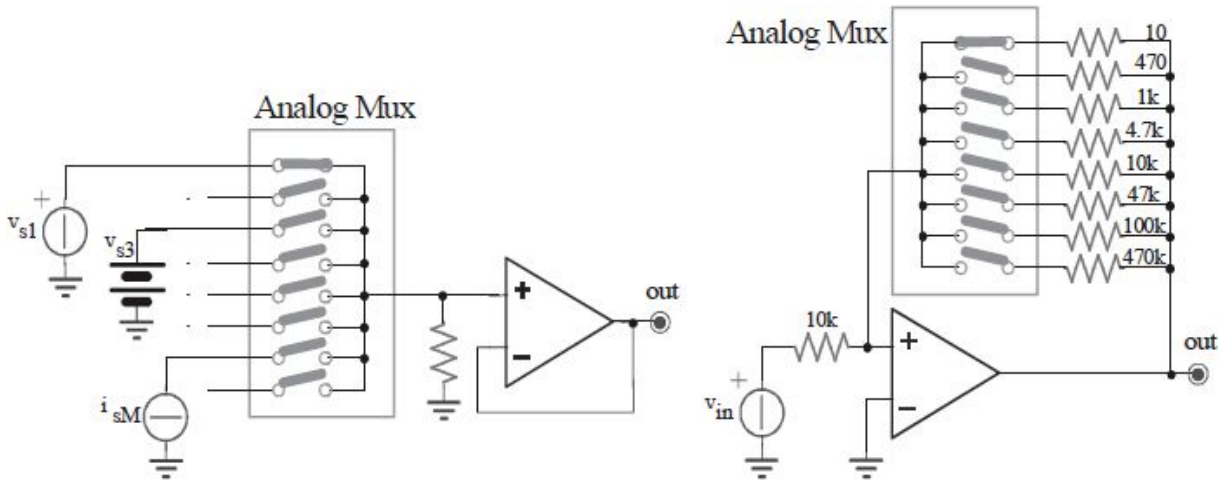


Fig. 6.67: Various uses of an analog mux: on the left, the mux Break-before-Make is recommended while on the right, the Make-before-Break typology is more suitable.

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.

6.10. Universal Active Filters

Once the signal is acquired (in hold phase), it brings the “good” information and all the other undesired things (errors, noises, tolerances, capacitive injections…). The following processing or, more likely, the digital processing downstream the ADC should improve the Signal-to-Noise Ratio (SNR). However, it is important to remove every possible error cause to avoid that undesired frequencies can be confused with the useful signal and be unavoidable also in the following digital processing.

For these reasons, it is important to make a good filtering upstream the S&H, both for the anti-aliasing purpose and for the signal conditioning (spectral cleanness, noise reduction, harmonics, equalization…). For example, Fig. 6.68 (on the left) shows the spectrum for an input sinusoid at 10kHz with a high spectral distortion due to the saturation phenomena. Practically, it is the same as a “squared” sinusoid applied to the input. Sampling this signal

without particular precautions, the spurious harmonics should be folded in the baseband, causing undesired tones, as shown in Fig. 6.69 (on the left).

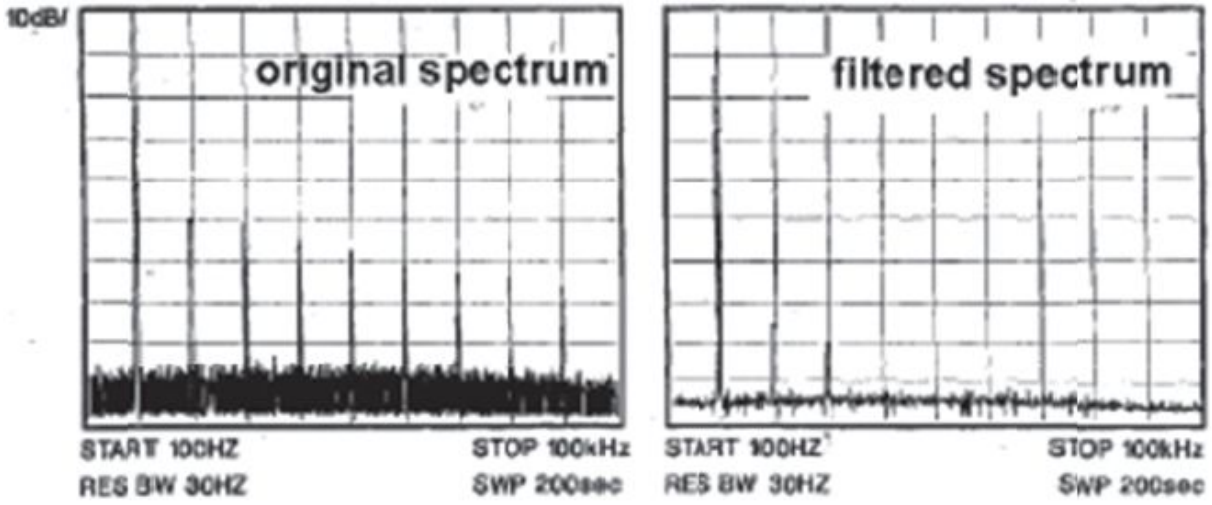


Fig. 6.68: Original input spectrum for a distorted sinusoid (on the left) and its filtered version (on the right).

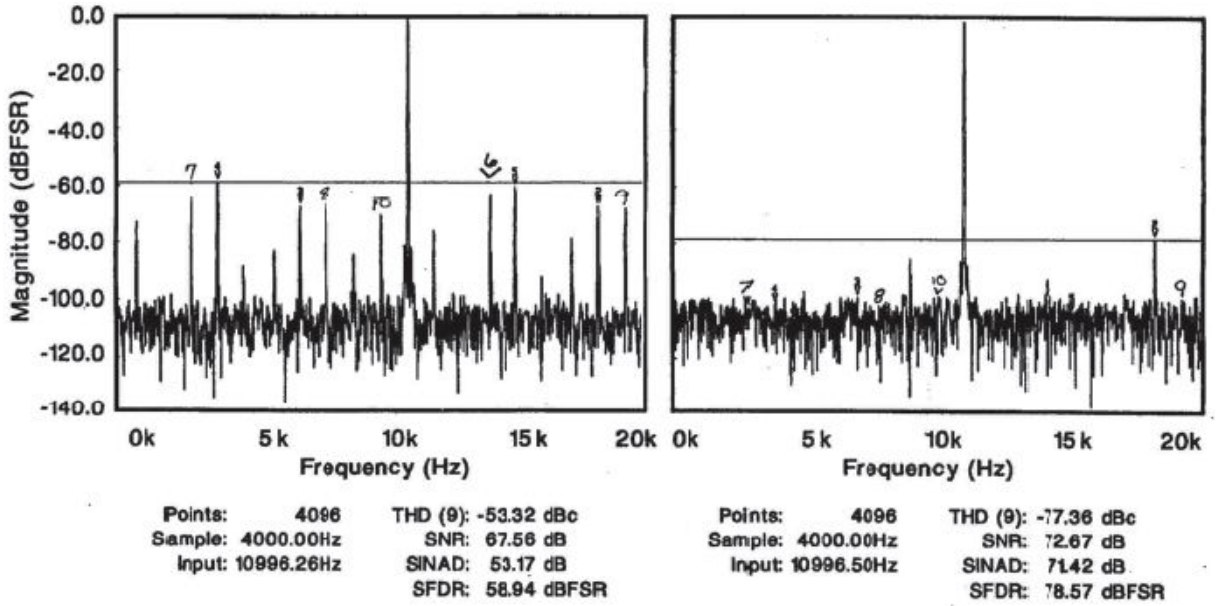


Fig. 6.69: Spectrum of the sampled distorted sinusoid (on the left) and filtered version (on the right).

To avoid this problem, you must pre-filter the input signal, for example with a Butterworth low-pass filter with a pole at 10kHz, which provides the spectrum shown in Fig. 6.68 (on the right). Now the harmonics have been reduced, and, probably, they do not pose problems downstream the sampling.

For example, if the signal is sent by a 16bit ADC with a sampling frequency of 40ksps (over the Nyquist limit due to f_{in}=10kHz), there will be the spectra shown in Fig. 6.69 at the output. All the 10kHz sinusoid harmonics should be folded in the baseband and are undistinguishable from the useful tones. In the case in the figure, thanks to the pre-filtering, one can improve the THD of 24dB, the SiNAD of 18dB, and the SFDR of 20dB.

The filtering should be low-pass or band-pass (never high-pass because an anti-aliasing filter will be necessary) and will be realized with different frequency masks (filter order, shaping, poles, and zeroes position), various typologies (passive, active, and mixed), and different devices (OpAmp general-purpose filters or specific integrates).

Just the necessity to have filters adaptable to different needs pushed manufacturers to develop integrated circuits dedicated to the realization of analog filters, known as Universal Active Filter (UAF). Many of these commercial devices are internally made of three OpAmps (or eventually another additional OpAmp to realize the output buffer), various resistors, and capacitors to make every type of filter (low-pass, high-pass, band-pass, and stop-band) with two or three poles, as in the classical feedback configurations. Fig. 6.70 shows the Burr Brown UAF42.

Opportunely changing the three outputs by means of external resistors, it is possible to obtain any two poles' frequency response. Depending on the poles' and zeros' positions, obtainable simply from the table provided by the UAF Producer, it is possible to make the typical frequency masks shown in Fig. 6.71. This choice is determined by needs of out-band attenuation or by the presence of undesired ripple within the bandwidth. The time domain response must be analyzed to choose the mask.

In this way it is possible to appreciate the speed of the filter, the progression towards the steady-state condition, and the presence of eventual ringing. This is showed in Fig. 6.72.

Depending on the application, you can choose, for example, a Bessel filter, or the Chebyshev configuration, with an abrupt frequency transition.

Modifying the external resistance values, one can realize an extremely adaptable filter. For example, employing programmable resistors (obtained with a DAC and an OpAmp), as shown in Fig. 6.73, it is possible to make a completely controllable digital filter. All the obtainable responses (Fig. 6.73) are simultaneously available at the different IC outputs.

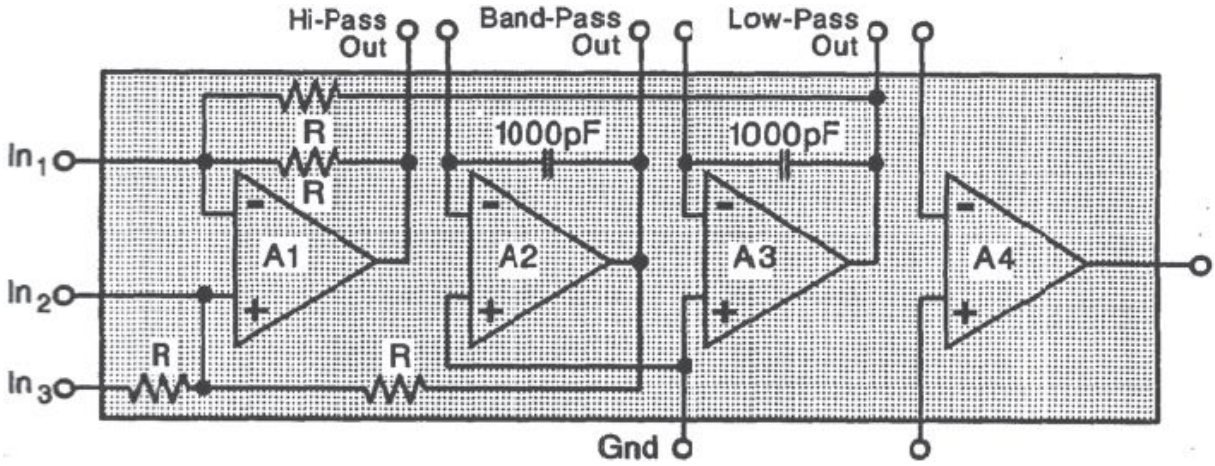


Fig. 6.70: UAF42 internal structure.

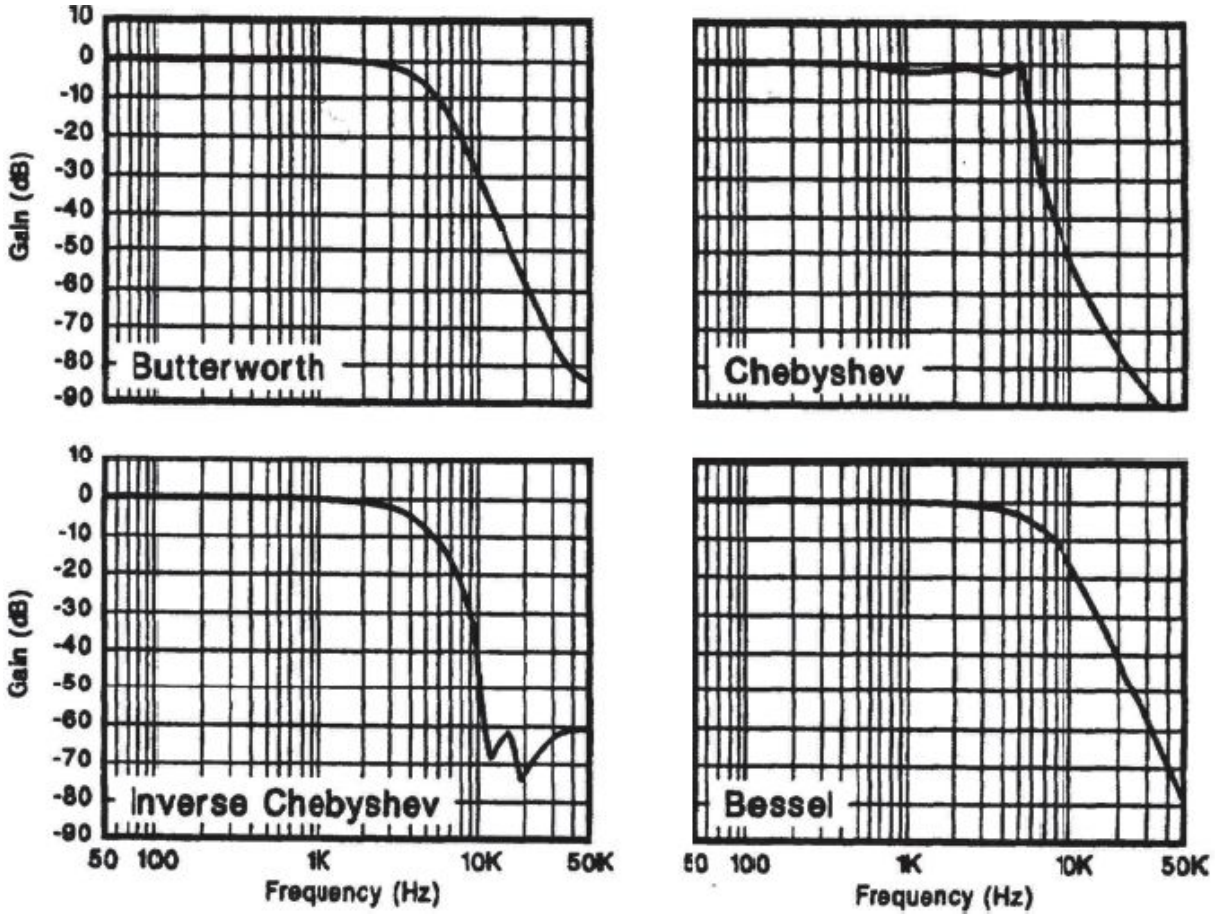


Fig. 6.71: Frequency standard masks for second order filters.

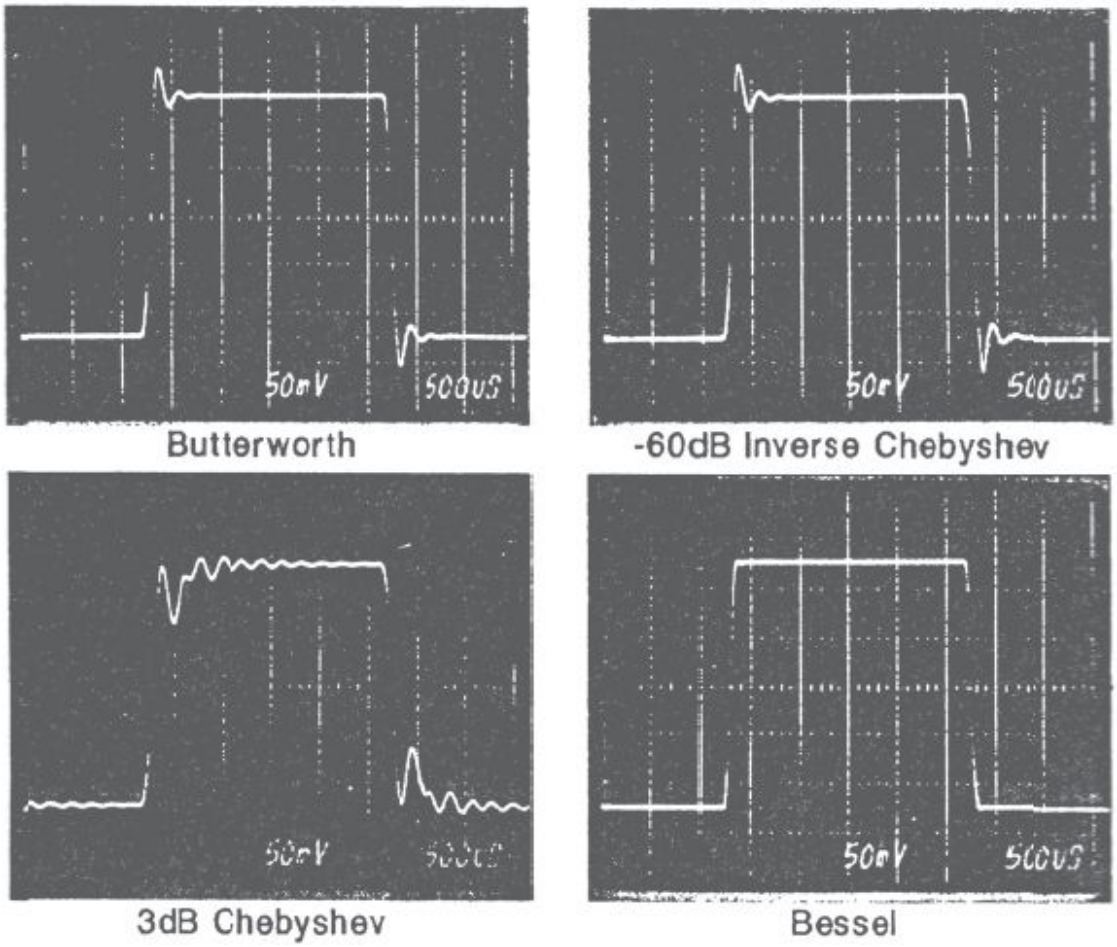


Fig. 6.72: Temporal responses for second order standard filters.

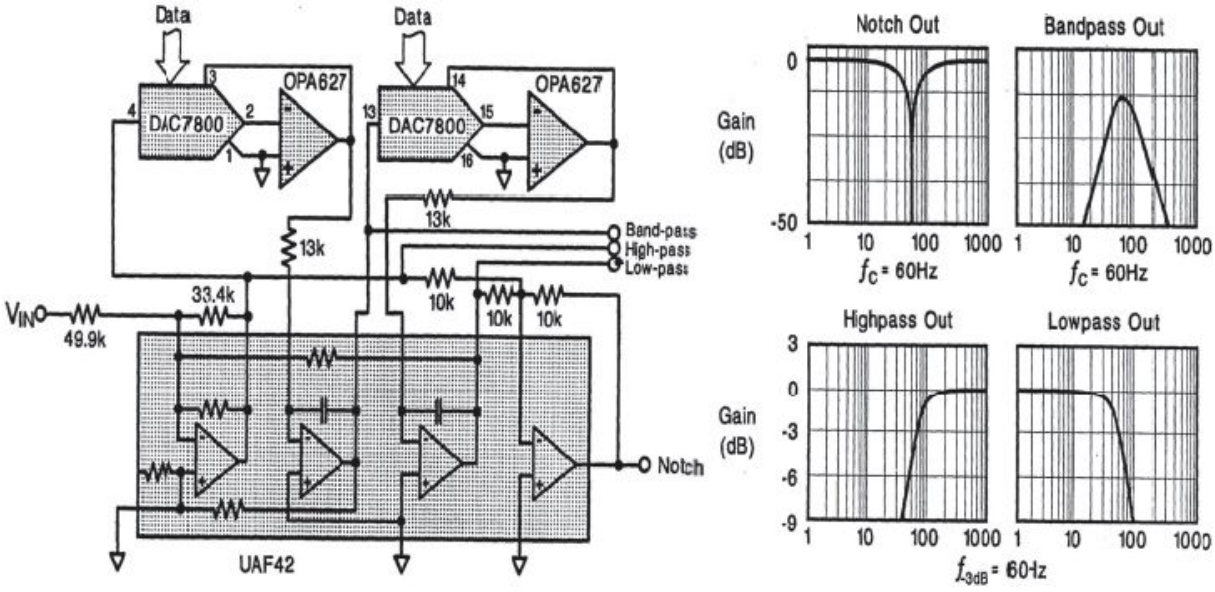


Fig. 6.73: Usage example for a Universal Active Filter (on the left) and the available output for the UAF (on the right).

Also UAFs containing a second order double filter, such as the Maxim MAX270, exist, as shown in Fig. 6.74. The two filters can be used separately (for example, one as an anti-alias and the other as a reconstruction filter in a conversion chain ADC+processing+DAC, as depicted in the same figure) or in cascades (to realize a fourth order filter).

19-3701; Rev 1; 8/91

MAXIM

Digitally-Programmed, Dual 2nd-Order Continuous Lowpass Filters

General Description

The MAX270/MAX271 are digitally-programmed, dual second-order continuous-time lowpass filters. Their typical dynamic range of 96dB surpasses most switched capacitor filters which require additional filtering to remove clock noise. The MAX270/MAX271 are ideal for anti-aliasing and DAC smoothing applications and can be cascaded for higher-order responses.

The two filter sections are independently programmable by either microprocessor (µP) control or pin strapping. Cutoff frequencies in the 1kHz to 25kHz range can be selected.

The MAX270 has an on-board, uncommitted op amp, while the MAX271 has an internal track-and-hold (T/H).

Applications

Lowpass Filtering

Anti-Aliasing Filter

Output Smoothing

Low-Noise Applications

Anti-Aliasing and Track-and-Hold (MAX271)

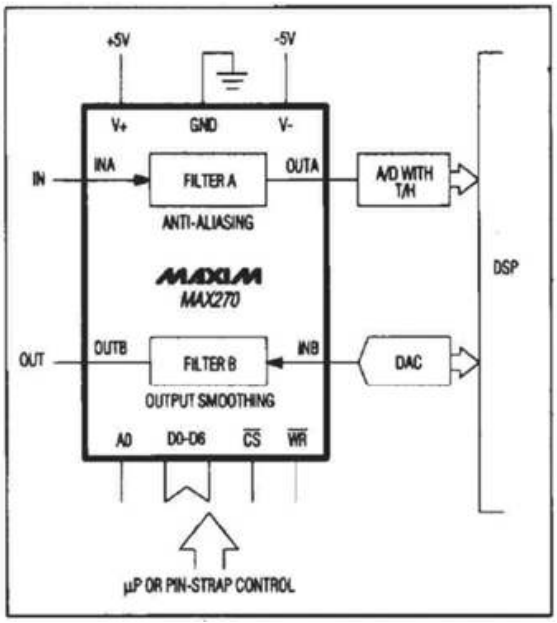
Features

- ♦ **Continuous-Time Filtering - No Clock Required**
- ♦ **Dual 2nd-Order Lowpass Filters**
- ♦ **Sections Independently Programmable: 1kHz to 25kHz**
- ♦ **96dB Dynamic Range**
- ♦ **No External Components**
- ♦ **Cascadable for Higher Order**
- ♦ **Low-Power Shutdown Mode**
- ♦ **Track-and-Hold (MAX271)**

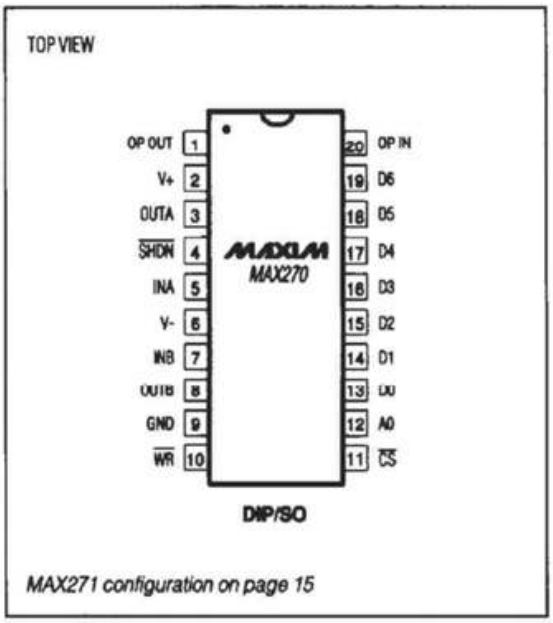
Ordering Information

PART	TEMP. RANGE	PIN-PACKAGE
MAX270CPP	0°C to +70°C	20 Plastic DIP
MAX270CWP	0°C to +70°C	20 Wide SO
MAX270EPP	-40°C to +85°C	20 Plastic DIP
MAX270EWP	-40°C to +85°C	20 Wide SO
MAX270MJP	-55°C to +125°C	20 CERDIP
MAX271CNG	0°C to +70°C	24 Plastic DIP
MAX271CWG	0°C to +70°C	24 Wide SO
MAX271ENG	-40°C to +85°C	24 Plastic DIP
MAX271EWG	-40°C to +85°C	24 Wide SO
MAX271MRG	-55°C to +125°C	24 CERDIP

Typical Operating Circuit



Pin Configurations



MAX271 configuration on page 15

Fig. 6.74: Frontispiece of a data-sheet of the double time-continuous UAF MAX270 (Maxim).

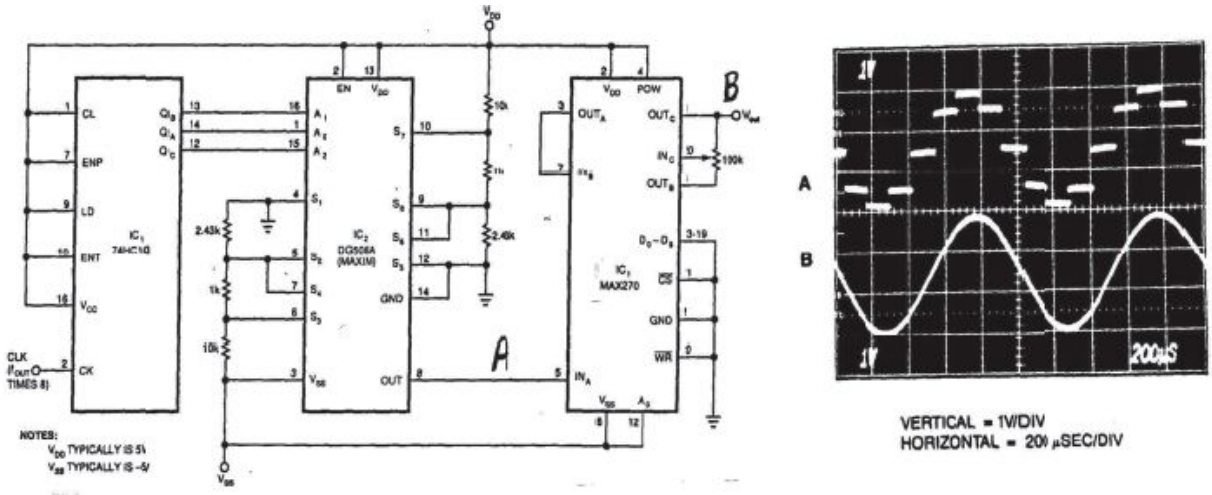


Fig. 6.75: Usage example for a universal filter to create a sinusoid starting from sinusoidal steps.

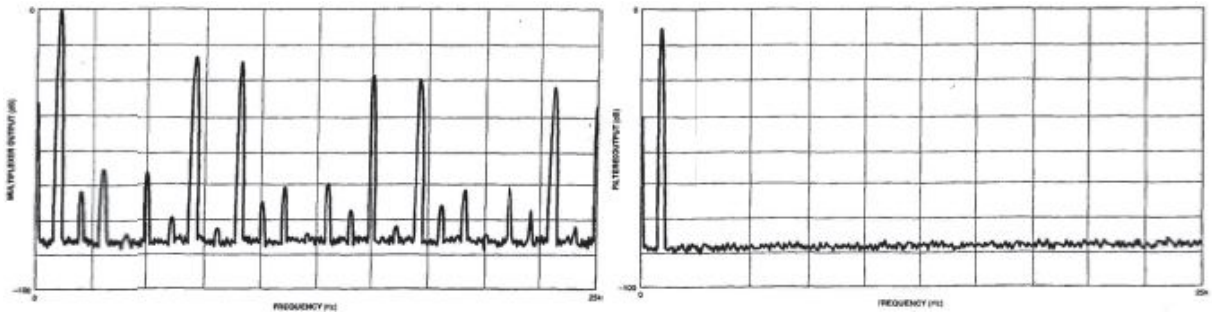


Fig. 6.76: Sinusoidal steps spectrum (on the left) of Fig. 6.75 and the sinusoid sine (on the right) obtained after the UAF.

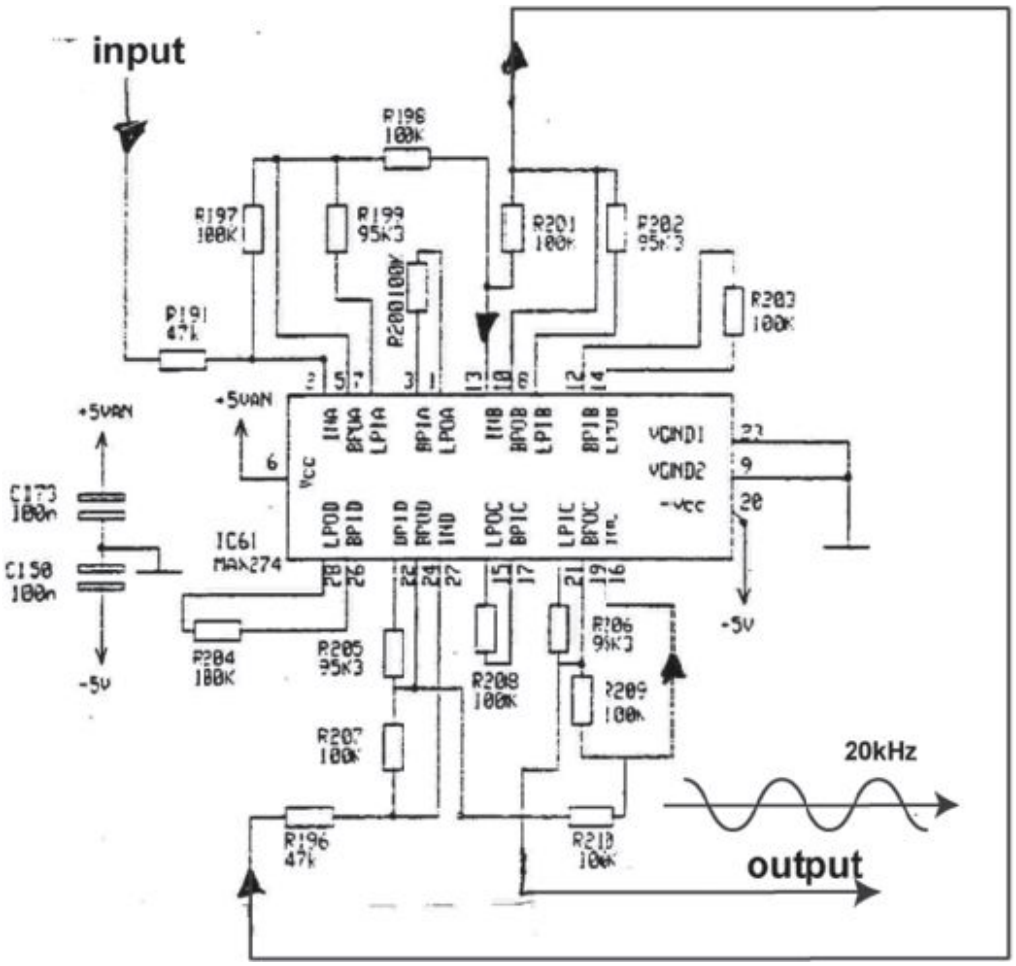


Fig. 6.77: 20kHz signal recovery with noise, using an IC with 4 second order-continuous-time filters. The circuit realizes an eight-order band-pass filter.

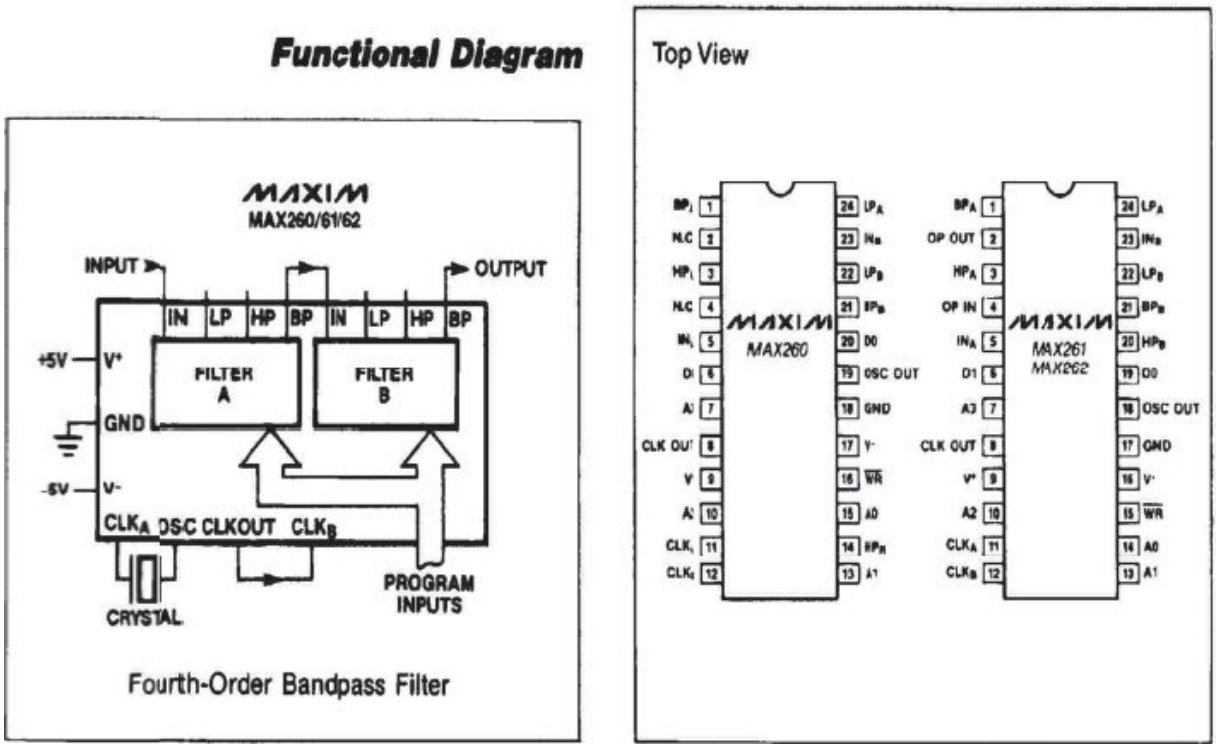


Fig. 6.78: Example of an IC containing a couple of switching capacitor filters.

A simple example of the use of these ICs is reported in the scheme in Fig. 6.75: a digital counter (74HC163) drives an eight-channel analog multiplier (DG508A) to obtain calibrated steps with a sinusoidal trend (see the steps shown in the upper trace of the oscilloscope shown in Fig. 6.75).

To obtain the continuous sine (and without distortion), the steps go towards the universal filter MAX270, connecting in cascade two second-order internal filters, obtaining a low-pass fourth-order Sallen-Key filtering. Spectral purity is more satisfactory (observe the lower trace in Fig. 6.75 and the output spectrum in Fig. 6.76). The THD is better than –80dB. The frequency is adjustable between 1kHz and 25kHz, simply modifying the counter command frequency (equal to eight times that required at the output).

The circuit in Fig. 6.77 shows another example of a circuit that employs, this time, a quadruple UAF (Maxim MAX274) to realize an eight-order (!) band-pass centered around 20kHz, simply with an IC and some external resistors. In this way, it is possible to recover the useful input signal at 20kHz, which is completely corrupted by white noise.

To obtain more selective filters, i.e. with a higher order, it should be possible to cascade many second order circuits, as done in the preceding

examples, with a consequent higher dimension of the overall circuit. Producers offer ICs with a new type of high order filters that are not based on OpAmps, time-continuous capacitance values, and resistance values, but on switched-capacitors. In these circuits, resistance values are realized with capacitance which is switched at high frequency (higher than the input signal bandwidth).

An example of an IC with a couple of switched capacitor filters is shown in Fig. 6.78; note that it is possible to arbitrarily configure the filters by means of appropriate digital programming lines.

6.11. EXERCISES

Ex. 1

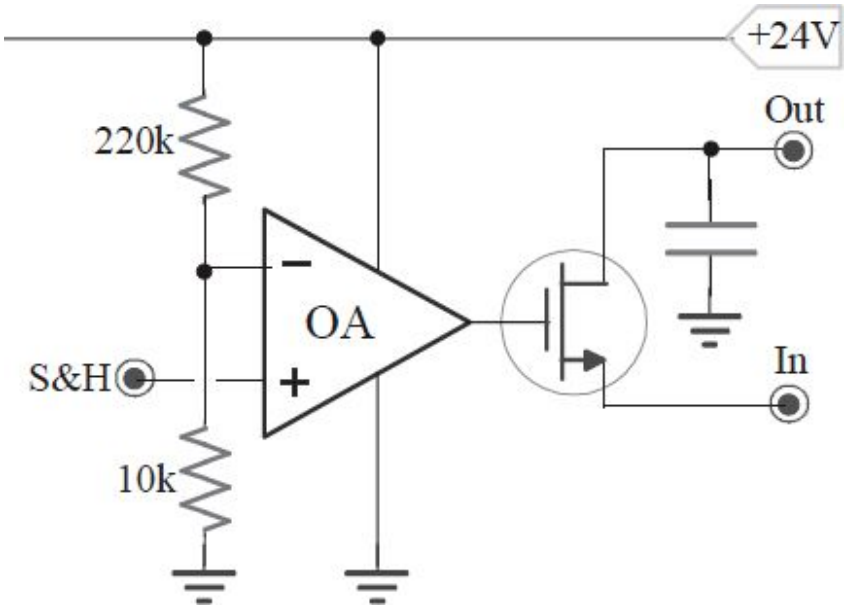
Design the S&H feedback scheme and describe the operations. The acquisition time depends on what? Which errors occur during the Hold time?

Ex. 2

The circuit realizes the gate driving of a Sample&Hold, towards a comparator (*GBWP=30MHz* and *SR=20V/µs*). The command is CMOS logic, with $t_{rise} \approx 10ns$.

a) Compute the opening time jitter due to the resistance tolerance, which is assumed to be *10%*.

b) Calculate the OpAmp and MOS parameters that limit the aperture jitter, providing some numerical examples.

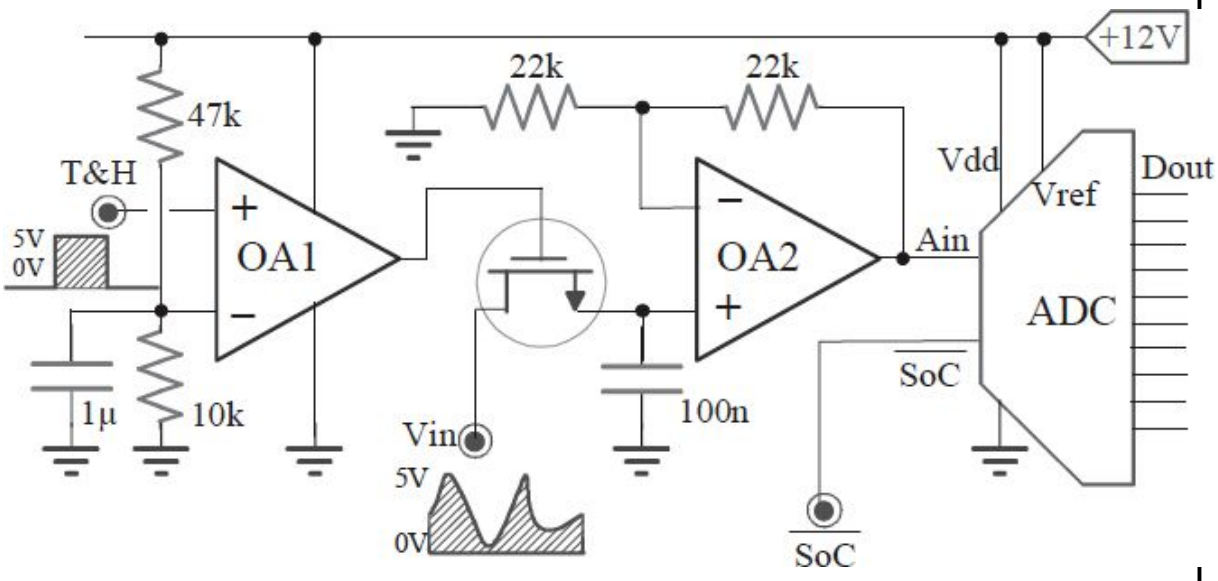


Ex. 3

The command T&H has $t_{rise}=t_{fall}=10ns$. OpAmps are rail-to-rail and have $PSRR=-80dB$, $SR=10V/\mu s$, and $I_{outmax}=\pm 10mA$. The MOS has $V_t=2V$. The input varies between 0 and $+10V$ with $f_{inmax}=20kHz$.

a) Compute the maximum error in terms of bits due to aperture delays.

b) Calculate the maximum error due to the residual ripple of $1V$ on the power supply.

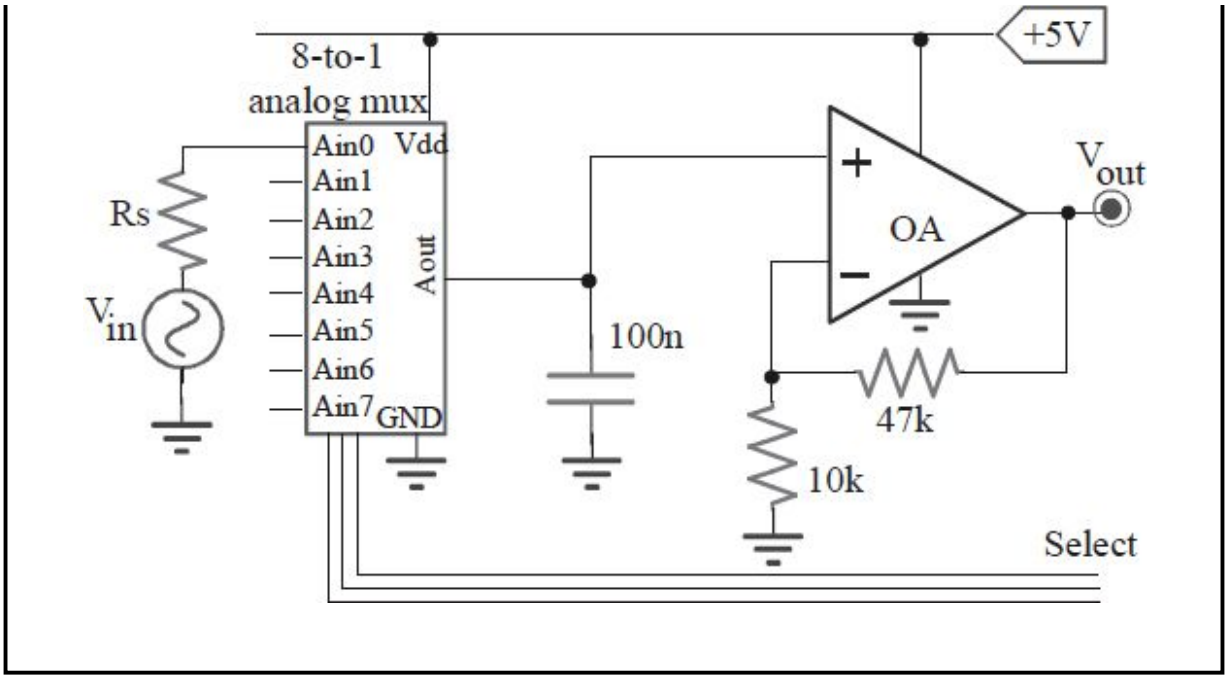


Ex. 4

The circuit acquires signals from eight sensors, characterized by Rs between $1k\Omega$ and $10k\Omega$ and by signals Vin between $0V$ and $400mV$.

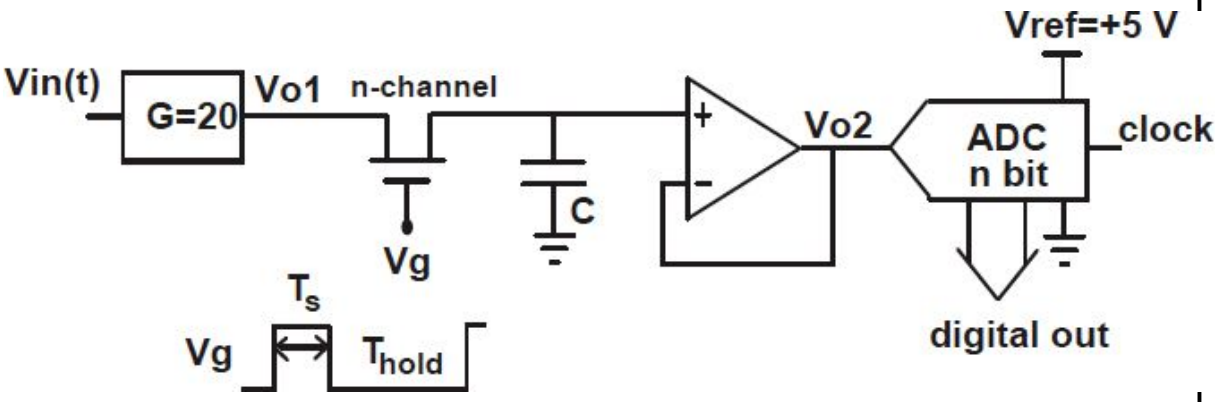
a) Quantify the **static** error (in Volts and bits number), knowing that the mux has $R_{on}=100\Omega$, $R_{off}=10M\Omega$, and $I_{leakage}=10nA$, and the OpAmp has $I_{bias}=50nA$ and $R_{in}=100M\Omega$.

b) Calculate the maximum acquisition time to ensure an accuracy of $10bit$.


+5V
8-to-1
analog mux
Ain0
Ain1
Ain2
Ain3
Ain4
Ain5
Ain6
Ain7
Vdd
Aout
GND
Rs
Vin
100n
OA
Vout
47k
10k
Select

Ex. 5

Consider the circuit (C=10nF) which samples and converts (with an ADC) the signal $V_{in}(t)=125mV\cdot(1+\sin 2\pi ft)$:



a) Compute the ADC *n* bits number which ensures a resolution on the input signal Vin(t) of at least 1/1000. Determine the amplitude of 1LSB (quantization interval) on the circuit input.

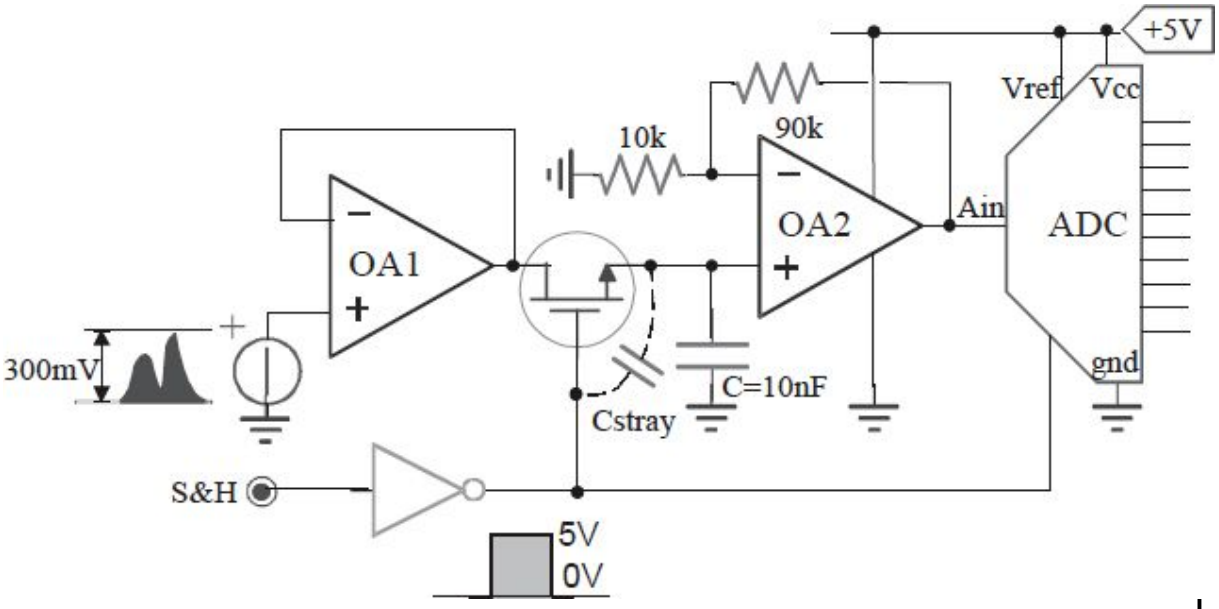
b) Assume that the converter clock frequency is 1MHz. Compute the conversion time of a steps ADC and an SAR-ADC. If T_{hold}=100μs, which of the two ADCs is appropriate in this case?

c) Assume that the OpAmp is characterized by bias currents Ib+=Ib-=100 nA. Determine the maximum value of T_{hold}, which ensures an error less than 1LSB.

d) Assume that the sampling time T_s is equal to 1μs. Determine the maximum MOSFET channel resistance (R_{on}) compatible with 1LSB resolution.

e) Assuming that the OpAmp is used with A_0=100dB, verify that the gain error is less than the ADC resolution.

Ex. 6



a) The input signal has a range of 300mV. Determine the ADC bits number needed for a resolution of 100µV in the input. Note that the ADC full scale is 5V.

b) Using a sampling time T_S=1µs and a hold time of T_H=4µs, estimate the maximum frequency for the input signal that the sampler is able to manage correctly (avoiding aliasing).

c) Considering the MOS parasitic capacitance (C_{stray}=0.4pF), estimate the maximum error due to the charge injection when the switch is open. Express it in LSB.

d) If one wants to follow an input sinusoid with amplitude between 0 and +300mV and with a frequency of 1MHz, what must the first OpAmp's minimum Slew Rate (in V/µs) be?

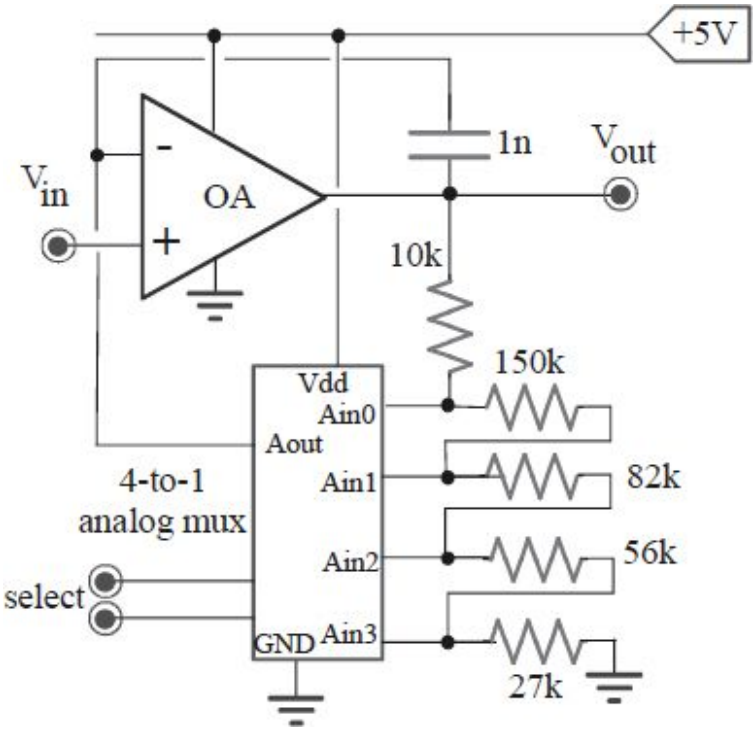
e) Determine the static error during the sampling phase (MOS switch closed with R_{DSon}=0) superimposed to A_{in} due to the outgoing bias currents I_{B+}=I_{B-}=50nA of the second OpAmp.

Ex. 7

The amplifier employs an OpAmp with A_0=*100dB* and *GBWP=50MHz* and an analog mux with $R_{on} \leq 10\Omega$, $R_{off} \geq 1M\Omega$, and $I_{leak} < 300nA$.

a) Explain the stage operations and compute the closed-loop gains, varying the selection.

b) Calculate the closed loop bandwidth, varying the selection.



Ex. 8

The circuit depicted in the figure is used as Sample & Hold for an 8bit ADC, characterized by an input dynamic 0-5 V. The input signal has a dynamic 0-0.8V.

a) Calculate the R_2 value to fully exploit the ADC dynamic to sample and convert the input signal V_{in}.

b) Consider the condition $V_G=V_{GL}$. Determine the output V_{out1} caused by an input step 0→0.8V.

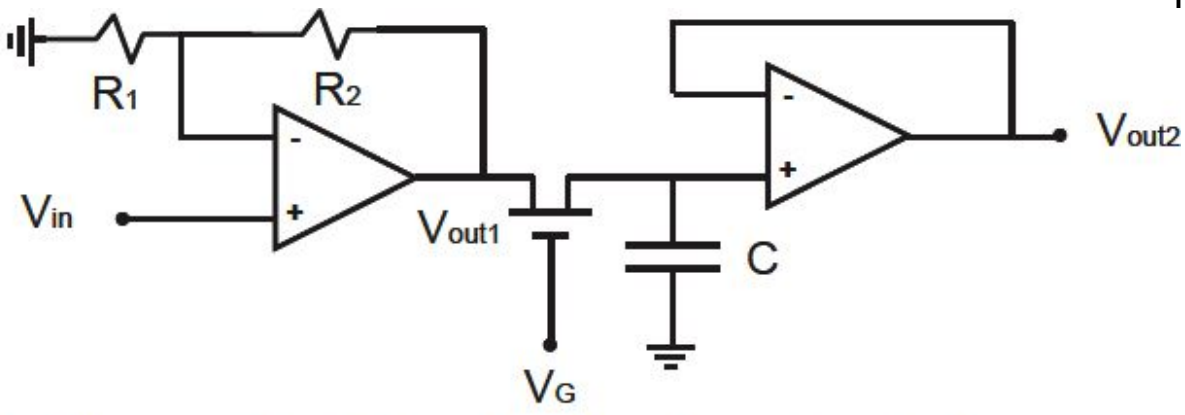
c) Calculate the first OpAmp's minimum SR such that the signal V_{out1} in the preceding point is not SR limited.

d) Compute the gain error introduced by the second OpAmp if it has an open-loop gain A_0=60dB. Express the error in LSB.

e) The hold time is T_H=200μs. Compute the minimum value of the capacitance C to ensure that V_{out2} decreases by a value (due to the bias current) less than LSB/2.

f) Assuming the MOS gate-drain capacitance C_{gd}=1.4pF and the capacitance C=25nF, compute the maximum error due to the charge injection.

g) Assuming to have a clock frequency f_{CK}=1MHz, discuss with quantitative arguments which ADC can convert in 200μs of hold time. Represent the internal scheme of the chosen converter.



R_1=10kΩ, GBWP=1MHz, V_{GL}=-5V, V_{GH}=10V, I_B=100nA.

a) To exploit the whole ADC dynamic, we want to amplify the maximum input signal V_{in} up to 5V. Therefore, the stage non-inverting gain must be equal to $G_{id}=1+R_2/R_1=5V/0.8V$, and so we obtain R_2=52.5kΩ.

b) Because, from the stage analysis, the real gain G_{real} has a unique pole (f_P), the step response has a single pole, and therefore V_{out1} is an exponential.

Ideal gain: $G_{id}=1+R_2/R_1=6.25$

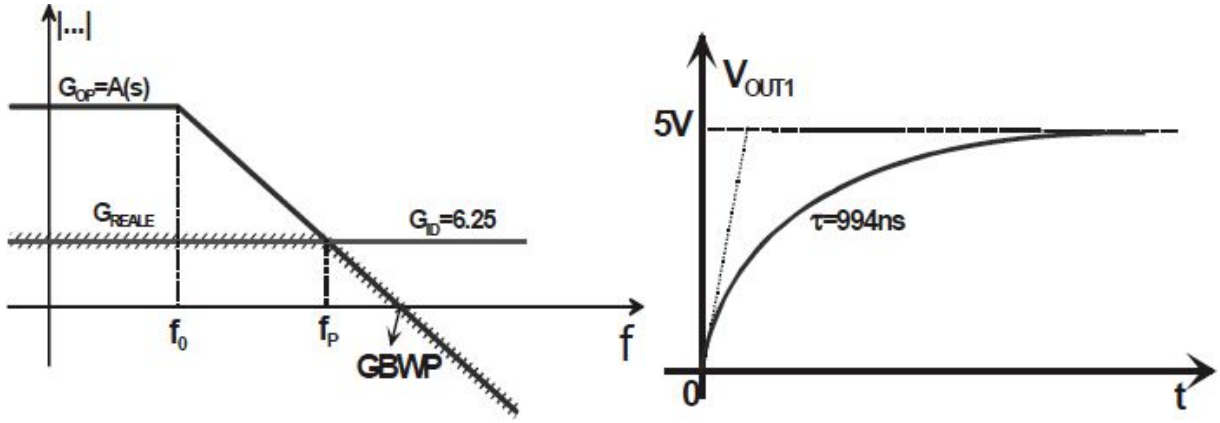
Loop gain: $G_{loop}=-R_1/(R_1+R_2)\cdot A(s)$, where A(s) is the OpAmp transfer function

Open loop gain: $G_{open}=A(s)=-G_{loop}\cdot G_{id}$.

Actual gain: G_{real} is equal to the ideal gain until the G_{loop} is high (at least 1), i.e. up to f_P; for $f>f_P$, G_{real} follows the open-loop gain G_{open}.

$GBWP=f_p\cdot G_{id}$ from which $f_p=160kHz$ and $\tau=1/(2\pi f_p)=994ns$.

The V_{out1} trend is an exponential $V_{out1}(t)=5V\cdot(1-\exp(-t/\tau))$.



c) The maximum slope of $V_{out1}(t)$ occurs when t=0 and is equal to $\Delta V/\tau=5.03V/\mu s$. Therefore, the minimum SR that does not distort the shape of $V_{out1}(t)$ is $SR_{min}=5.03V/\mu s$.

d) The static error ε for a feedback configuration with an OpAmp is the percentage difference between the actual gain and the ideal gain. The general expression $\varepsilon=1/(1-G_{loop}(0))$ which in the case (in which $G_{loop}(0)=-A(0)$) becomes $\varepsilon=1/(1+A(0))\approx 1/A(0)=0.1\%$, which is equal to $0.1\%\cdot 5V=5mV$. Moreover, $LSB=FSR/2^n=5V/2^8=19.5mV$. Therefore, the static error corresponds to $\varepsilon=5mV/LSB=25.6\%$ LSB (i.e. around ¼ of the LSB).

e) During an interval of $T_H=200\mu s$, the bias current $I_B=100nA$ discharges C with a charge $Q=I_BT_H$. The absence of such an amount of charge cannot bring down the voltage across C more than $\Delta V=LSB/2$. Because $Q=C\Delta V$, we have $C_{min}\Delta V= I_BT_H$, from which $C_{min}=2.05nF$.

f) You must evaluate the maximum Charge-Injection Induced Offset (or pedestal error) that happens when V_G switches from -5V to +10V, and the

voltage across C is equal to 0V: ΔV_G=15V. Therefore, $V_{injection}=\Delta V_G \cdot C_{GD}/(C_{GD}+C)$=839μV=4.3% LSB.

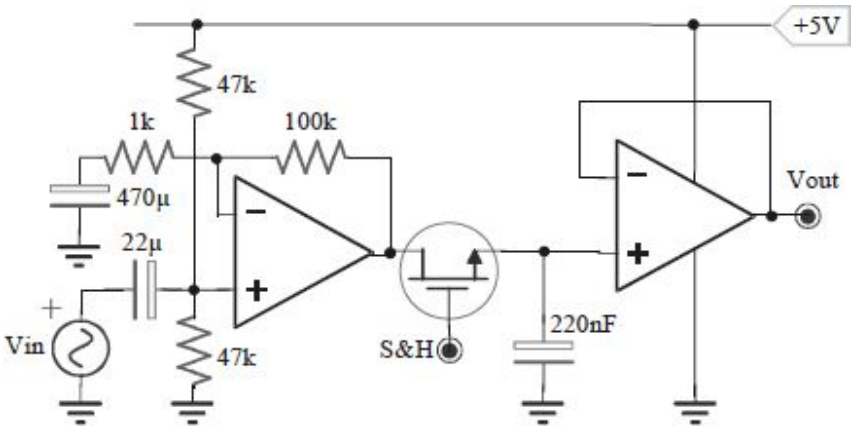
g) A tiered ADC cannot convert the signal because it should require a conversion time of $T_C=2^n/f_{CK}$=256μs. You cannot use an SAR ADC that requires a conversion time of $T_C=(n+1)/f_{CK}$=9μs. For the internal scheme, see the theory part.

Ex. 9

The S&H is upstream a *14bit* ADC with *FSR=5V*. The OpAmps have an *SR=10V/μs* and I_B*=100nA*, where the MOS has R_{on}*<10*Ω.

a) Determine the maximum sampling and hold times.

b) Compute the maximum frequency of the signal with $|v_{in}|$*<20mV*.



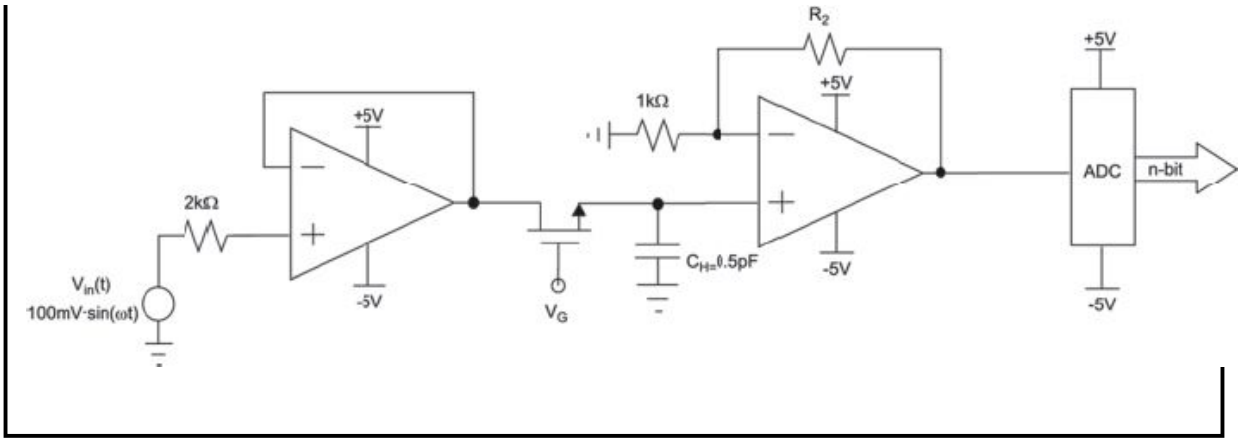
Ex. 10

Consider the following circuit with digital conversion of the input signal $V_{in}(t)$.

a) Size R_2 to exploit the entire ADC dynamic for a given input signal.

b) Knowing that n=10 bit, how much is the maximum tolerable **charge** injected into the switch?

c) If the first stage OpAmp has a bias current equal to $I^+ = I^- = 0.2\mu A$, determine, in LSB, the currents induced error.



a) Because the ADC input dynamic is equal to 5V-(-5V)=10V, the maximum amplitude for a sinusoidal input signal is 10V/2=5V. The signal $V_{in}(t)$, read by a voltage buffer (first stage OpAmp), is amplified by the non-inverting configuration with a gain $G=1+R_2/1k\Omega$ before the ADC. Therefore, the gain that allows exploiting the whole ADC dynamic is that which amplifies V_{in} from 100mV to 5V, i.e. $G=5V/100mV=50$. Thus, $R_2=49k\Omega$.

b) For the ADC, we have that the $LSB=10V/2^{10}=9.76mV$. To generate an error less than ½LSB to the ADC input, the voltage across C_H cannot vary more than $\Delta V=(½\ LSB)/G=97.6\mu V$. Thus, the maximum tolerable variation in the charge of C_H is due to the closure of the switch is $Q=C_H\cdot\Delta V=4.88\cdot10^{-17}C$.

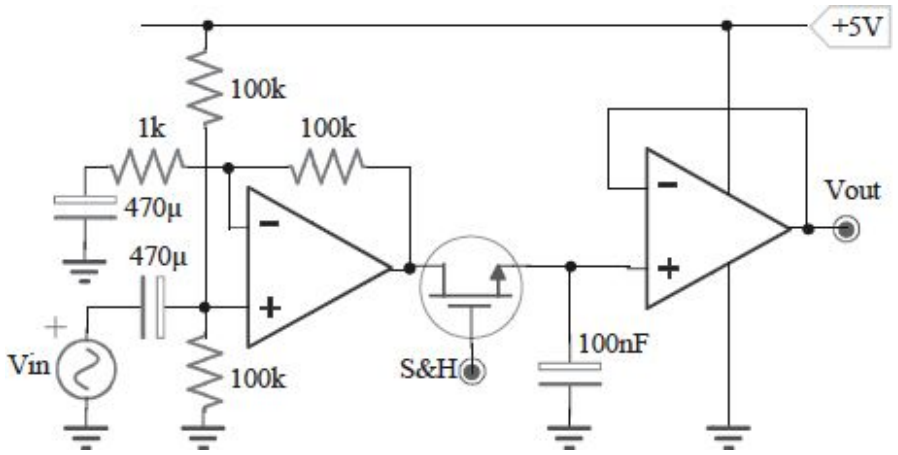
c) First OpAmp's bias current I^+ causes a voltage $I^+\cdot2k\Omega=0.4mV$ across the 2kΩ resistor. Such a voltage determines an ADC input error equal to $0.4mV\cdot G=20mV$, which corresponds to $2.04\cdot LSB$. I^- bias current causes no effect.

Ex. 11

The OpAmps have A_0=*100dB*, *SR=5V/µs*, V_{os}=*10mV*, and I_b=*500nA*. The input is sinusoidal with a peak of *20mV*. The ADC has *16bit* and *FSR=5V*. The MOS has R_{on}<*50*Ω.

a) Determine the maximum $t_{acquisition}$ to ensure an accuracy of *½LSB*.

b) Compute the maximum output errors, after a T_{hold}=*1µs*.



a) The first OpAmp has a medium frequency gain equal to 1+100k/1k=101. The two 100kΩ resistors give a biasing of the OpAmp output in the center dynamic downstream the ADC. The second OpAmp, the MOS, and the 100nF capacitor make a simple Sample&Hold. When the MOS switch is closed, it starts the acquisition phase. The voltage across the Hold capacitance (and therefore V_{out}) should reach the V_{in} voltage except for a negligible error. The time needed for this situation is due to two factors: the OpAmp's finite Slew Rate and the exponential error of the capacitor, through R_{on} of the MOS. The maximum acquisition time is when the signal across the capacitor has the maximum excursion, from -20mV to 20mV, which, at the output, corresponds to a variation from 0.5V to 4.5V with a ΔV equal to 4V. The maximum slope of the C_{HOLD} charge is at the start of the exponential charge equal to

$\left.\frac{dV}{dt}\right|_{max} = \frac{\Delta V}{\tau}$. With $\tau = C_{HOLD} \cdot R_{on=}5\mu s$, you obtain the maximum slope of 0.8V/µs. Because such a value is less than the OpAmp's SR, you can conclude that the acquisition time is limited to the sole exponential charge and can be obtained from $T_{acquisition,max} = \tau \cdot \ln\left(\frac{\Delta V}{\varepsilon}\right)$, assuming an error of ε=½LSB=38µV, you can obtain $T_{acquisition,max}$=58µs.

b) The maximum errors can be both static and dynamic. The static error is due to OpAmp offset voltages, their bias currents, and the finite gain. The dynamic error during the transition between the sampling phase and the hold phase concerns only the charge injection. Finally, the dynamic error during the hold phase concerns also the droop, i.e. the charge/discharge (depending on the current direction) of the Hold capacitor due to the parasitic currents (bias, leakage, etc.).
About the offset, it is very interesting to note that it is brought to the OpAmps' output because it is a DC component. In fact, also the first OpAmp is a buffer configuration because the feedback capacitor behaves as an open circuit. In the worst case, the output is the sum of both contributions:

$$V_{out,OS} = \pm 2 \cdot V_{OS} = \pm 20mV = \pm 262LSB!!$$

The first OpAmp's bias currents provide an output contribution greater than those of the second OpAmp; in fact, I_{B2-} does not generate errors, and I_{B2+} sees only R_{ON}, with a very low value. Indeed, the first OpAmp's currents' effect is equal to:

$$V_{out,bias} = \pm I_{B+} \cdot (100k \,//\, 100k) \mp I_{B-} \cdot 100k = \mp I_B \cdot 50k = 25mV = \pm 328LSB!!$$

The finite gain of the OpAmp should give an important error because it is the gain stage while the second contributes only for 1/101; thus:

$$V\varepsilon = \frac{V_{out,max}}{A_0} = \frac{4.5V}{10^5} = 45\mu V = 0.59LSB$$

Finally, after a Hold time of 1µs, the droop will cause an error equal to $\Delta V = I_{B2+} \cdot T_{Hold}/C_{Hold} = 5\mu V = 0.07LSB$, which is absolutely negligible.

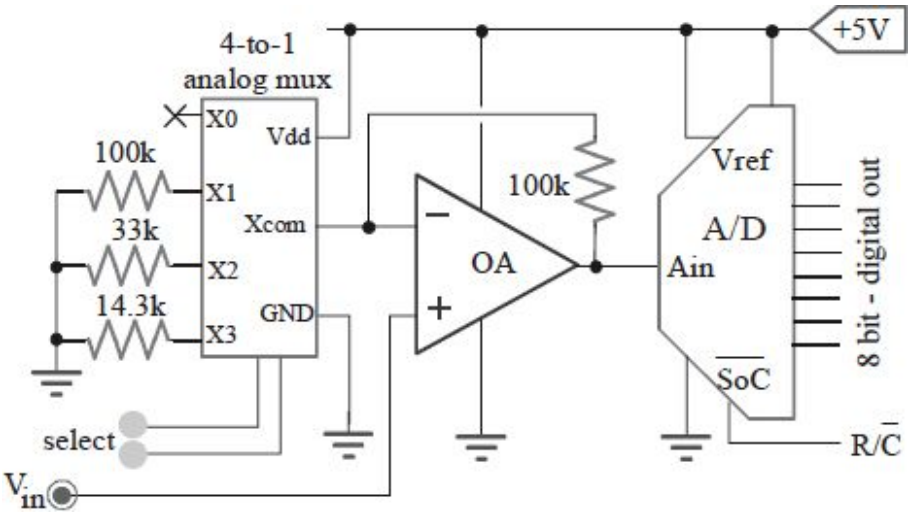
Totally, the static error is extremely high, around 600LSBs (!) for a 16bit ADC. It should be negligible (because less than ½LSB) for an ADC with $\log_2(600\cdot 2)\approx 10$bit less than the first one, i.e. of only 16-10=6bit. You need to reduce the error, for example, by making the impedance values seen in DC by the first OpAmp equal so as to cancel the errors introduced by I_{B1}. Or, the error can be reduced after the ADC by means of a digital post processing of the signal.

Ex. 12

OpAmp (I_B=*100nA and* V_{OS}=*5mV*), mux ($R_{on}\leq 100\Omega$ *and* $R_{off}\geq 100k\Omega$) and ADC (*8bit*, V_{ref}=*5V and* I_{in}=*1μA*).

a) With *select=11*, determine the static error in LSB at the ADC input, which is caused by the non-idealities.

b) Compute the gains for various selections.



a) With the selection 11, the ideal gain for the amplifying stage is equal to $\left(1+\frac{100k\Omega}{14{,}3k\Omega}\right)=8.$ Actually, because of the presence of the resistance values R_{ON} and R_{OFF}, the actual gain is equal to:

$$\left(1+\frac{100k\Omega}{(14{,}3k\Omega+100\Omega)\,//\,(33k\Omega+100k\Omega)\,//\,(100k\Omega+100k\Omega)\,//\,(\infty+100k\Omega)}\right)=9{,}2.$$

R_{ON} R_{OFF} R_{OFF} R_{OFF}

Therefore, against a maximum input of 5V/8=625mV, the code of the ADC saturates at 5V for an input voltage equal to 5V/9.2=543mV. The error is computed as (625mV–543mV)/625mV=0.13, i.e. equal to 13% of the input dynamic! In terms of LSB, because the quantization levels are 256 (2^8), this equals to 0.13·256=33LSB. The bias current of the inverting terminal flows to the feedback resistor with a voltage drop equal to 100nA·100kΩ=10mV=½LSB; those outgoing from the non-inverting input gives no effect. The offset voltage, ideally applied to the non-inverting terminal, is amplified with a factor of a gain of 9.2 and, therefore, causes a voltage variation equal to 5mV·9.2=46mV=2.4LSB.

b) With the code 00, the ideal gain should be 1; in the real case, it is:

$$\left(1+\frac{100k\Omega}{(14{,}3k\Omega+100k\Omega)\,//\,(33k\Omega+100k\Omega)\,//\,(100k\Omega+100k\Omega)\,//\,(\infty+100\Omega)}\right)=3{,}1.$$

With the code 01, the ideal gain should be 2; in real case, it is:

$$\left(1+\frac{100k\Omega}{(14{,}3k\Omega+100k\Omega)\,//\,(33k\Omega+100k\Omega)\,//\,(100k\Omega+100\Omega)\,//\,(\infty+100k\Omega)}\right)=3{,}6.$$

With the code 10, the ideal gain should be 4; in real case, it is:

$$\left(1+\frac{100k\Omega}{(14{,}3k\Omega+100k\Omega)\,//\,(33k\Omega+100\Omega)\,//\,(100k\Omega+100k\Omega)\,//\,(\infty+100k\Omega)}\right)=5{,}4.$$

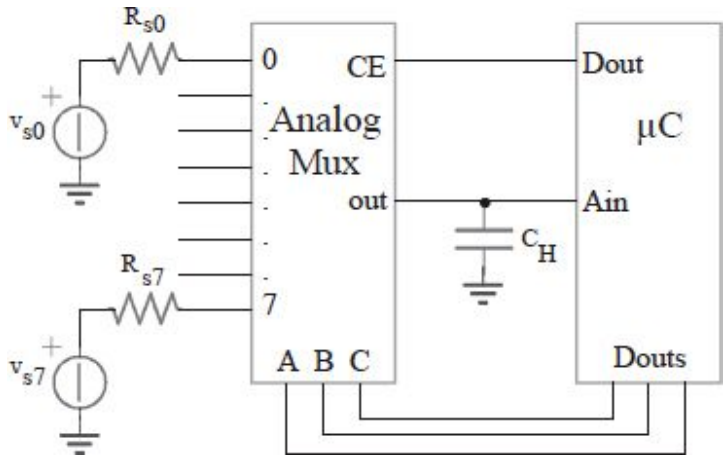
For the code 11, such as the point a), the ideal gain should be 8, and the real one is 9.2.

Ex. 13

The 8 inputs analog-mux (*Ron<100Ω* and *Ileak<50nA*) is controlled by the µC with an 8 bit internal ADC (*Vref=5V*, *IleakAin<50nA*, and *RAin>8MΩ*) without S&H.

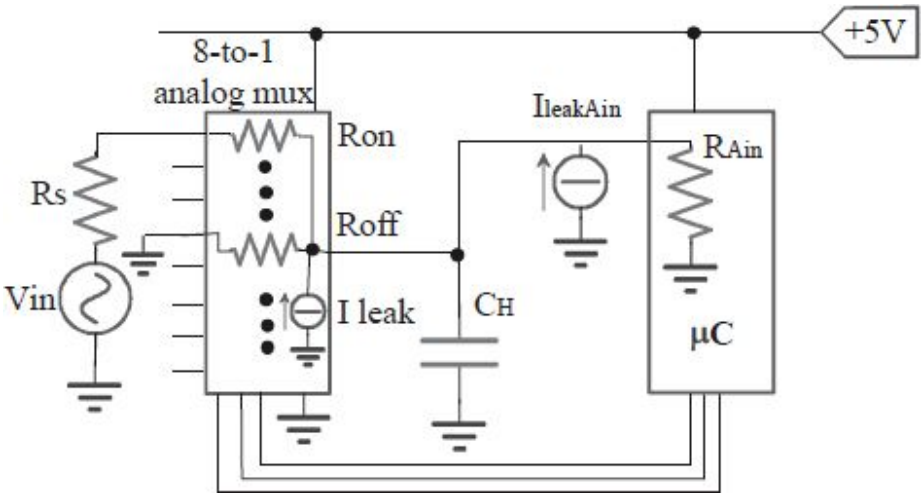
a) Compute the maximum value of the *RSx* to have a precision of *½LSB*.

b) Estimate the acquisition time for an accuracy of *100µV* when *CH=10nF*.



a) The circuit is designed with an 8-to-1 analog mux and a microcontroller. Such a stage can be used, for example, to sample eight different signals from many sensors. The µC must be opportunely programmed to obtain the desired timing. The lines employed are only five: four digital lines to select the input and the mux enabling and the analog line for the internal ADC.

Every source $V_{s0} \div V_{s7}$ has a series resistor R_{sx} across which there is an undesired voltage drop that causes an error. Other error causes are due to mux leakage currents and the ADC, to the R_{on} mux resistance, and to the R_{off} mux resistors and non-infinite R_{in} ADC resistance values. The error sources are reported here.



The ADC input maximum static error due to the resistive drop is given by:

$$V_{\varepsilon,\max} = V_{in,\max} \cdot \frac{R_{sx} + R_{on}}{(R_{sx} + R_{on}) + (R_{off}/7 \,//\, R_{in,ADC})} \cong V_{in,\max} \cdot \frac{R_{sx} + 100}{R_{sx} + 100 + R_{in,ADC}}$$

$$\cong V_{in,\max} \cdot \frac{R_{sx} + 100}{R_{in,ADC}} \le \frac{1}{2} LSB$$

where the hypothesis is $R_{sx} << R_{in,ADC}$. Supposing moreover that $V_{in,max}=V_{ref}$, we obtain:

$$R_{sx} \le \frac{1}{2} LSB \cdot \frac{R_{in,ADC}}{V_{in,\max}} - 100 = \frac{1}{2}\frac{5V}{2^8} \cdot \frac{8M\Omega}{5V} - 100 = 15525\Omega$$

The ADC maximum input static error due to the parasitic currents is given by:

$$V_{\varepsilon,\max} = (8 \cdot I_{leak} + I_{leakAin}) \cdot \left[(R_{sx} + R_{on}) // R_{off}/7 \,//\, R_{in,ADC} \right] \cong (8 \cdot I_{leak} + I_{leakAin}) \cdot (R_{sx} + R_{on})$$

This error must be less than ½LSB, and we must choose R_{sx}<9.5mV/450nA=21kΩ. Between these two conditions, the more pressing is R_{sx}<15kΩ.

b) When the microcontroller selects one of the eight input signals, which is sent to the internal ADC for the sampling, some time is needed to have a settled tension on the hold capacitance to the correct value V_{in}. To estimate

the acquisition time, we need to consider the worst case, i.e. when the internal input is from 0 to 5V. Note the capacitance charge law, the acquisition time can then be estimated as:

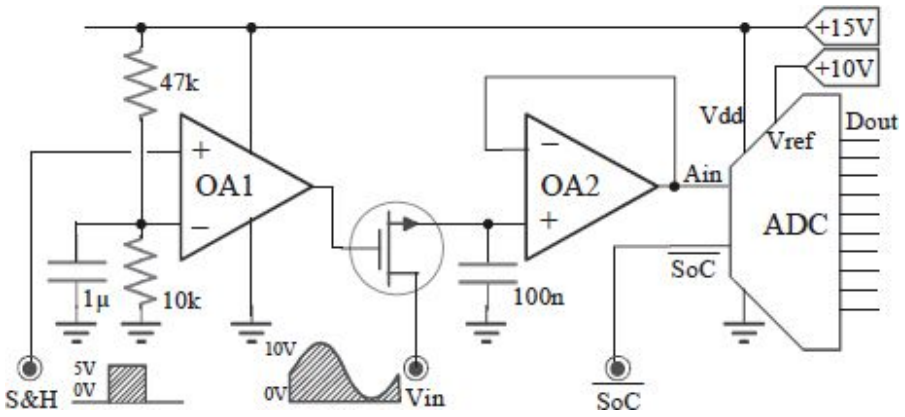
$$T_{acquisition} = \tau \cdot \ln \frac{\Delta_{\max}}{\varepsilon} = (R_{sx} + R_{on}) \cdot C_H \cdot \ln \frac{5V}{100\mu V} = 1.7ms$$

Ex. 14

The S&H command has $t_{rise}=t_{fall}=20ns$. The rail-to-rail OpAmps have $SR=5V/\mu s$ and $I_{leak}\leq 100nA$. The MOS has $V_t=2V$ and $C_{GD}=C_{GS}=10pF$. The input between $0\div 10V$ has $f_{inmax}=20kHz$.

a) Compute the maximum hold time to ensure an accuracy of *1LSB* with a *10bit* ADC.

b) Compute the maximum error for the opening delay.



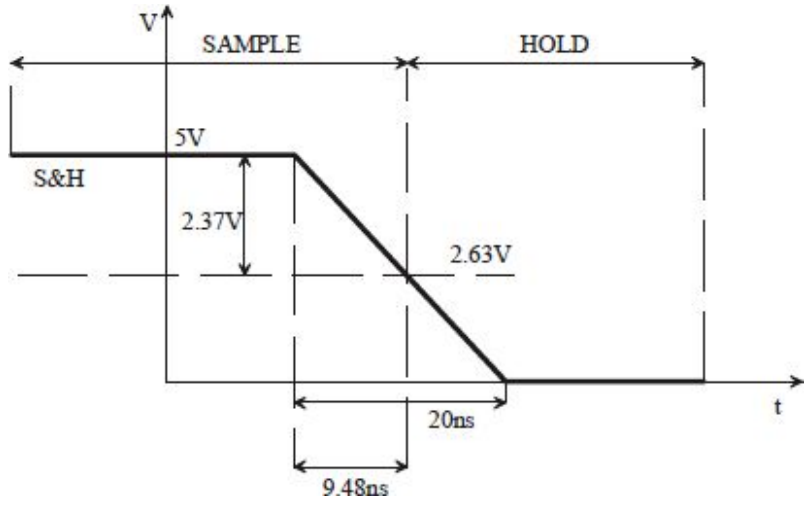
a) The circuit realizes a Sample & Hold to sample a signal between 0÷10V. The first OA1 is used as a comparator; its output is always saturated to the power supplies, to +15V or 0V (GND), because the OpAmps are rail-to-rail. When the S&H input is low (0V), the output of OA1 is 0V, and, consequently, the n-MOS is open (Hold phase). When S&H is brought to 5V, the situation is reversed: the output is 15V. Whatever the voltage stored on the capacitance is, the MOS is closed because $V_G>>V_{in}+V_T$ (Sample phase).

During the hold phase, the leakage currents discharge the capacitance, causing an error known as the droop. Employing a 10bit ADC with FSR equal to 10V, the LSB is equal to $LSB=FSR/2^n=10V/2^{10}=9.76mV$. This is the value of the LSB also at the capacitance level because the OA2 is a buffer that has a gain equal to 1. Because the leakage discharges the capacitor with a speed

$\frac{\Delta V}{\Delta T} = \frac{I_{leakage}}{C_H}$ and assuming ΔV_{max}<1LSB, we obtain:

$$T_{HOLD,MAX} = \Delta T_{MAX} = \frac{LSB \cdot C_H}{I_{leakage}} = \frac{9.76mV \cdot 100nF}{100nA} = 9.76ms \cdot$$

b) The aperture error gives an inaccuracy on the sampled signal due to the MOS aperture non-coincident with the command S&H signal. A first cause is represented by nonzero rising and falling edges of the S&H command. Assuming that t_{rise} and t_{fall} are referred to the whole signal range from 0V to 5V and vice-versa, the delay between the starting of the S&H command and the OA1 comparator switching is as shown in the following figure. In fact, the comparator threshold is 2.63V; therefore, the signal must decrease by 5-2.63=2.37V before the inverting pin voltage exceeds that of the non-inverting pin. The time the signal takes to cover the whole range is equal to $T'_{aperture} = 20ns \cdot \frac{2.37V}{5V} = 9.48ns \cdot$



A second effect which delays the MOS opening is represented by the OA1 Slew-Rate, defined as the maximum speed with which the OpAmp output can vary. The time required for OA1 to bring its tension below the source voltage (V_{in}) plus the required V_T depends on the voltage stored on the capacitor (precisely V_{in}). If the capacitor stores a voltage of 10V, the OA1 output must start from 15V during the sampling phase to 12V to open the switch and goes in the hold phase: the excursion is only 3V. Indeed, in the worst case, the excursion will be 13V when on the capacitor is stored the minimum voltage V_{in}=0V. You can obtain the corresponding opening times:

$$T''_{aperture} = \frac{\Delta V}{SR} = \begin{cases} \dfrac{15V - 12V}{5\dfrac{V}{\mu s}} = 0.6\mu s \\ \dfrac{15V - 2V}{5\dfrac{V}{\mu s}} = 2.6\mu s \end{cases}$$

We can say that the opening delay has an average value of (2.6+0.6)/2=1.6μs and a jitter of (2.6-0.6)/2=±1μs, which, although it is not perfectly correlated with the input signal, can be treated as a casual fluctuation, to simplify the calculations.
The maximum input variation speed is equal to:

$$\left.\frac{dV_{in}}{dT}\right|_{\max} = 2\pi \cdot f_{in,\max} \cdot V_{peak} = 2\pi \cdot 20kHz \cdot 5V = 0.63\frac{V}{\mu s}$$

The error due to the opening delay is obtainable as:

$$\varepsilon_{medio} = \left.\frac{dV_{in}}{dT}\right|_{\max} \cdot T_{aperture,mean} = 0.63\frac{V}{\mu s} \cdot 1.6\mu s = 1V$$

while the error due to the opening jitter is:

$$\varepsilon_{\max} = \left.\frac{dV_{in}}{dT}\right|_{\max} \cdot T_{aperture,jitter} = 0.63\frac{V}{\mu s} \cdot (\pm 1\mu s) = \pm 0.63V \cdot$$

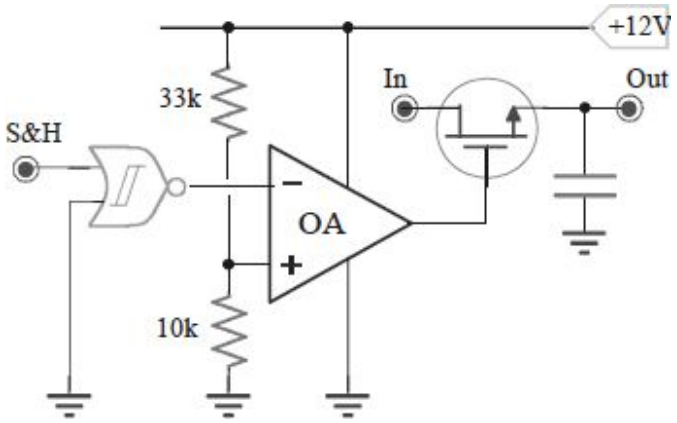
The error due to the mean delay can be acceptable in the case of continuous free-running acquisition (corresponding to a constant delay of the sampling comb) while that due to the jitter is non-negligible.

Ex. 15

The comparator has an *SR=120V/µs* while the *5V* input CMOS command has t_{rise}*=20ns*. The input is variable in the range *0÷5V*. The MOS has the threshold voltage equal to *2V*.

a) Compute the minimum and maximum opening times (different because of the input signal variations).

b) Assuming that the opening delay is *100ns* and that the input analog signal *In* is *1kHz*, determine the output *aperture-time error*.



a) The circuit realizes a driver for a Sample&Hold MOSFET. The S&H command is a CMOS type (therefore can assume values 0V and 5V). Because the input is variable between 0 and 5V, it is not enough to allow the switch closure (Sampling phase) for every input value. You need to translate the logic level *High* to a greater voltage to ensure that, with the maximum V_{in} equal to 5V, the MOS is “well closed”.

When the S&H command is high, the output of the NAND logic gate is low, and the comparator output is saturated to 12V (assuming that it is rail-to-rail); the MOS conducts for every input value. When the S&H command is low, the output of the NAND gate is high, and the inverting terminal presents a voltage greater than that of the non-inverting terminal, equal to

12·10k/(10k+33k)=2.8V; consequently, the comparator output is saturated to the ground, and the MOS is open for every input value.
The transition from Sampling to Hold is non-instantaneous because of the inevitable switching times of the S&H command and the finite OpAmp slew rate. This causes a delay for the switching opening and a great aperture-time error. Because it is not specified, we can suppose that the logic gate switches in a negligible time. We suppose moreover that t_{rise} is referred to the whole excursion of the command, from 5V to 0V and vice-versa.
The aperture delay depends on the input signal; it can be computed by adding the two delays, the first of which is fixed and the second variable:

$$T_{aperture,min} = T_{rise} \cdot \frac{2.8V}{5V} + \frac{(V_{G,sampling} - V_{in,max}) - V_T}{SR} = 11.2ns + \frac{12V - 5V - 2V}{0.12V/ns} = 53ns$$

$$T_{aperture,max} = T_{rise} \cdot \frac{2.8V}{5V} + \frac{(V_{G,sampling} - V_{in,min}) - V_T}{SR} = 11.2ns + \frac{12V - 0V - 2V}{0.12V/ns} = 94ns$$

b) A switching aperture delay involves an error in the voltage stored in the capacitor. This aspect has a little influence for the free-running sampling because we have a sampling comb uniformly delayed (except the jitter). More problematic is the case of a single-shot sampling, for random and non-period instants, in which this error must be restrained. The worst case for the aperture error is when the signal is sampled in the maximum variation, i.e. where its derivative is the maximum. In this case, the ΔV (error) that we have is equal to:

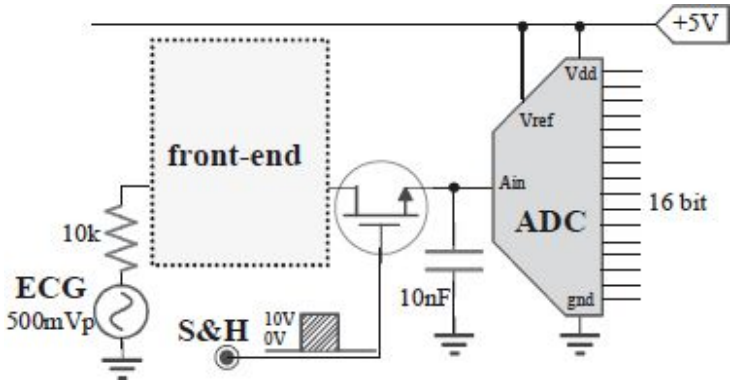
$$\Delta V_{max} = \left.\frac{dV}{dt}\right|_{max} \cdot T_{aperture} = \left.\frac{d(V_p \sin(2\pi ft))}{dt}\right|_{max} \cdot T_{aperture} = 2\pi \cdot f \cdot V_p \cdot T_{aperture} = 2\pi \cdot 1kHz \cdot 2.5V \cdot 100ns = 1.57mV$$

Ex. 16

The MOS has $R_{on}<200\Omega$, $C_{GD}=C_{GS}=C_{DS}=3pF$, and $I_{leakage}=1nA$. The input is an ECG signal with an excursion of $\pm 500mV$ and bandwidth of *0.2Hz-200Hz*.

a) Design the front-end stage to exploit the whole FSR.

b) Without the front-end and with the generator switched off (short-circuit), calculate the total error in *LSB* due to the charge-injection, leakage, and R_{on} after *2ms* of MOS aperture.



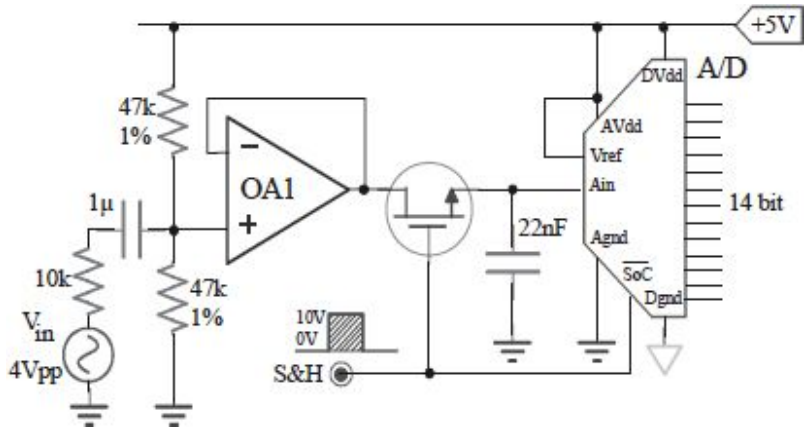
Ex. 17

The OpAmp has *IB=100nA, A0=100dB,* and *CMRR=80dB.* The ADC has *14bit* with *Vref=5V*. The MOS has $R_{on} \leq 200\Omega$.

a) Compute the static errors at the ADC output, in *LSB*, caused by *IB* and the resistive tolerance degrees.

b) Compute the errors due to the finite value of *A0* and *CMRR*.

c) Determine the **minimum** input frequency which ensures an accuracy of *1LSB*.



a) The 14bit ADC with V_{REF}=5V has LSB=5V/2^{14}=305µV.

The outgoing bias current is divided by two: half on the ground branch and half on the power supply branch (the input capacitor is an open circuit in DC). Considering the branch towards the ground, the current produces a voltage drop across the non-inverting pin, which is equal to:

$$\frac{100nA}{2} \cdot 47k\Omega = 2{,}35mV = 8LSB$$

As we can see, the voltage on the non-inverting input is buffered to the output with unity gain. The inverting pin's outgoing bias current does not cause modifications at the output voltage because there are not any feedback resistors.

The 1% resistance tolerance on the 47kΩ resistor means that the actual value can vary between 47470Ω and 46530Ω. In the worst case, i.e. when each has the extreme values, the static error is equal to:

$$2{,}5V - 5V \cdot \frac{47470}{47470 + 46530} = 25mV \qquad \text{i.e. 25mV / 305µV = 82LSB.}$$

b) Because of the finite gain (A_0), the output voltage is $V_O=(V_{IN}–V_O)\cdot A_0$ (because V_O is brought to the input by the buffer feedback branch). The error ε, i.e. the nonzero difference ($V_{IN}–V_O$) due to the finite gain is equal to:

$$(V_{IN} - V_O) = \varepsilon = {V_O}/{A_O}$$

in the worst case, i.e. with an input voltage equal to $2V_p$ is ε=2V/100000=20µV=LSB/10.
The CMRR is the ratio (expressed in dB) between the OpAmp differential gain (A_0) and the common-mode gain A_{CM}. The common mode gain is therefore:

$$CMRR = 20Log_{10}\left({A_0}/{A_{CM}}\right) \rightarrow A_{CM} = {A_0}/{10^{CMRR/20}} = 10$$

i.e., **if the OpAmp were used in an open-loop configuration**, a common-mode signal simultaneously applied to both the input terminals should be amplified 10 times. Actually, the OpAmp is used in a closed-loop configuration; therefore, it is necessary to evaluate the impact of the common-mode gain. To work properly, the OpAmp always has an error voltage ε between its two inputs:

$$V_O = V_{diff} \cdot A_{diff} + V_{CM} \cdot A_{CM} = \varepsilon \cdot A_0 + \frac{V_+ - V_-}{2} \cdot A_{CM} \approx \varepsilon \cdot A_0 + V_{in} \cdot A_{CM} = \varepsilon \cdot A_0 + V_O \cdot A_{CM}$$

The last inequality holds because, in this case, we have $V_{in}=V_o$ because the stage is a buffer. We can obtain that

$$\varepsilon = \frac{V_O}{A_0} - \frac{V_O}{CMRR}$$

Actually, because the common-mode gain sign is unknown, you must consider the worst case, which is summing two different contributions, actually, subtracting them from each other. Therefore, in respect of the ideal case with an infinite CMRR, the error at the OpAmp input is not given by V_O/A_O=20μV, but must be added to the term V_0/CMRR=2V/10000=200μV, absolutely predominant in this case and equal to 1LSB.

c) The passive input network is a high-pass filter. At high frequencies, the gain is 1 (0dB); therefore, with 2V peak as the maximum input, we should have an output of a peak of 2V. The pole has a frequency 1/(2π·1μF·[(47kΩ//47kΩ)+10kΩ])=4.8Hz.

If the input were a sinusoid with the same frequency as that of the pole, the -3dB attenuation should determine an ADC input signal of 1.41V, indeed of the expected 2V, with a consequent drop approximately of 1LSB in the case of a trivial 4bit ADC or 9666LSB in the case of a 14bit ADC! Therefore, we need to use a frequency far from the pole or, vice-versa, to use a pole with a frequency lower than the minimum bandwidth of interest. To be precise, requiring a precision of ½LSB at the minimum input frequency $f_{in,min}$, we need that the pole's frequency respects this equation:

$$\frac{f_{in,min}}{f_{pole}} \cong \sqrt{2 \cdot 2^n \frac{V_{inpp}}{FSR}} = (\text{if } V_{inpp} = FSR) = \sqrt{2 \cdot 2^n}$$

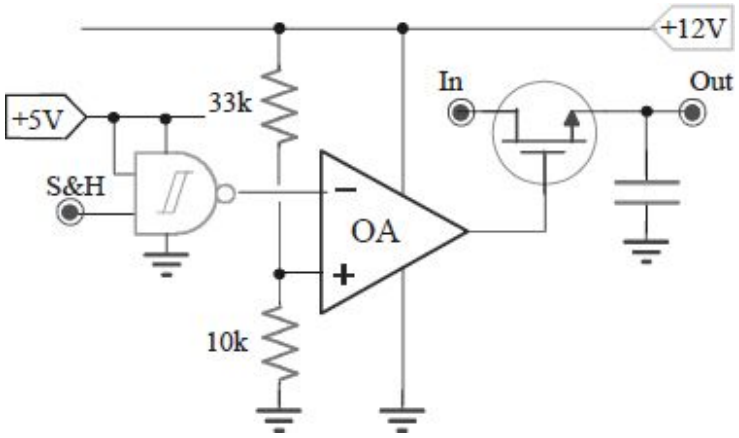
In the case of 14bit, we need to use a $f_{in,min} > f_{pole} \cdot 2^{15/2} = f_{pole} \cdot 181 = 870Hz$, more than two orders of magnitude above.

Ex. 18

The comparator has an *SR=20V/µs*, and the NAND has a negligible delay. The input varies between *0÷5V*. The MOS has V_t*=2V*.

a) Compute the minimum and maximum aperture times due to the variability of the input signal.

b) Assuming that the jitter of the S&H signal is *100ns* and that the *In* signal is at *20kHz*, determine the error due to the timing jitter.



a) 0.25µs, 0.5µs.

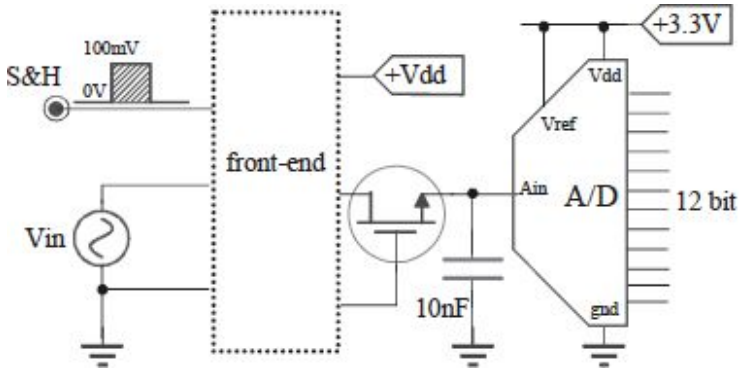
b) Error=31.4mV.

Ex. 19

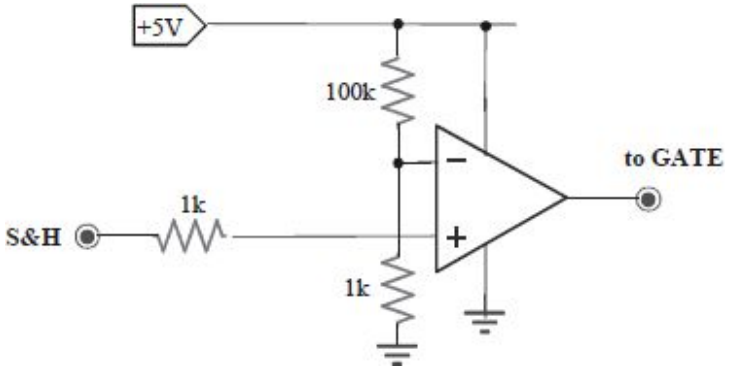
The input V_{in} is sinusoidal in the range *±24V* and has bandwidth of *20kHz*. The S&H command has only a dynamic of only *100mV*.

a) Design the driver stage for the MOS gate and choose, if required, the appropriate V_{dd}.

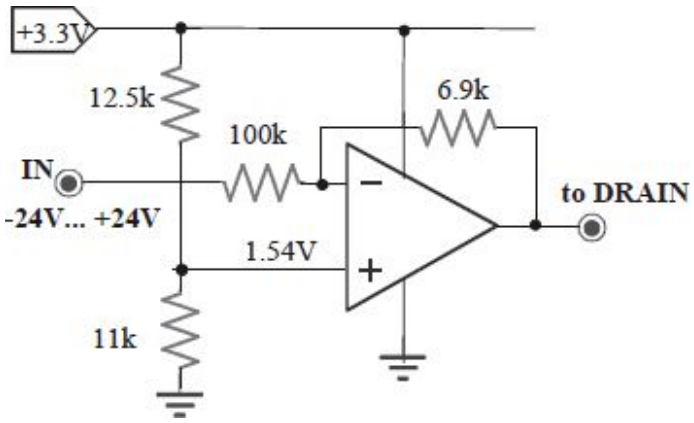
b) Design the conditioning front-end circuit for the Vin input signal to exploit the whole ADC *FSR*.



a) It should be enough to switch the MOS on, also for signals applied to its drain of 3.3V, corresponding to the whole FSR of the ADC downstream. It is better to use a rail-to-rail OpAmp connected as a comparator, supplied with Vdd=+5V if the MOSFET has a threshold voltage less than 5V-3.3V=1.7V. Otherwise, you must use a greater Vdd. The circuit is shown in the following figure. The switching threshold can be set around 50mV.



b) You need to attenuate the ±24V signal to bring it within the FSR, between 0÷3.3V, of the ADC. We can use a normal inverting stage with R2/R1=3.3V/2·24V=0.069, for example R1=100kΩ and R2=6.9kΩ. The OpAmp can be safely supplied with a +3V single-power supply because the voltage of its input terminals is always the virtual ground provided that it is rail-to-rail. If you want to avoid that Vin=+24V does not coincide with a negative voltage, but 0V, it is enough to translate the input voltage to V*, such as V*·(1+R2/R1)=3.3V/2, i.e. V*=1.54V.

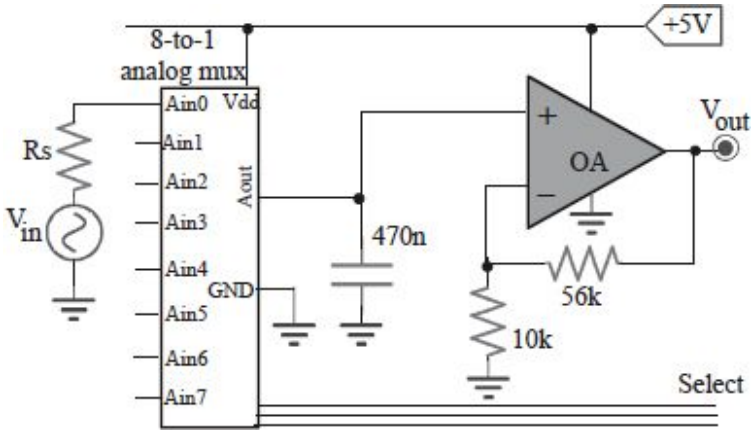


Ex. 20

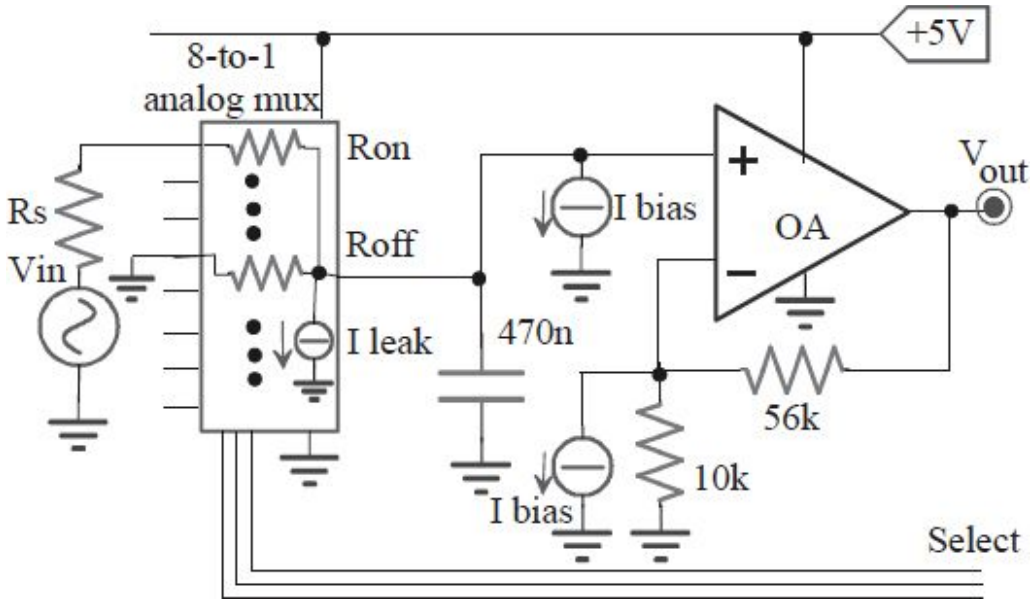
The circuit acquires signals from eight sensors, with $Rs \leq 10k\Omega$ and signals V_{in} in the range $0 \div 700mV$. The mux has $R_{on}=10 \div 500\Omega$, $R_{off} \geq 800k\Omega$ and $I_{leak}=100nA$. The OpAmp has $I_b=100nA$.

a) Compute the maximum **static** error.

b) Compute the **minimum** acquisition time and the **maximum** hold time to ensure a *14bit* precision.



a) The errors can be evaluated starting from the following scheme.



The error across the capacitance can be found by applying the superimposition principle: $\Delta V = \pm Vin \cdot \frac{Rs + Ron}{Rs + Ron + Roff/7} \pm (Ileak \cdot 8 + Ibias) \cdot [(Rs + Ron) // Roff/7]$ which, combined in the worst case, give 66mV. The OpAmp's output total error is, including also I_{bias-}, equal to $66mV \cdot \left(1 + \frac{56k\Omega}{10k\Omega}\right) + I_{bias-} \cdot 56k\Omega = 430mV$. It is extremely high if compared with 1LSB (with a 14bit ADC, it is only 305µV!)

b) To obtain an error less than ½LSB (153µV), taking into account the input signal maximum excursion and assuming no limitation by the OpAmp's Slew-Rate, the minimum time required for the acquisition can be obtained as follows:

$$Tacq = \tau \max \cdot \ln \frac{Vin, \max}{0.5 \cdot LSB} = 41.6ms$$

The hold time can be limited because of the droop phenomenon, i.e. the discharge of the hold capacitance due to the parasitic currents (bias, leakage ...). The maximum time required to ensure the precision of ½LSB can be obtained from $I = C \cdot \Delta V / T_{hold}$ and is equal to $T_{hold} = 470nF \cdot 153\mu V / 100nA = 719\mu s$.

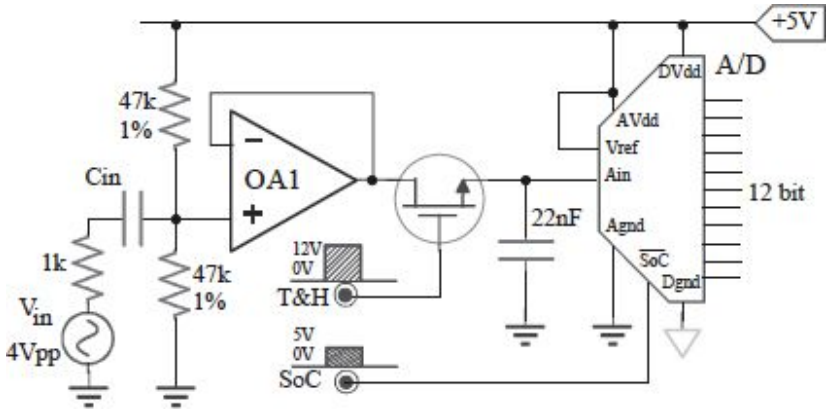
Ex. 21

The OpAmp has *Iout,max=10mA* and *SR=10V/µs*, The MOS has $R_{on} \leq 100\Omega$, and the ADC converts in $T_{conv}=2\mu s$.

a) Compute $f_{in,max}$ which reaches the OpAmp limits.

b) Estimate $f_{S,max}$ which ensures the accuracy of *±1LSB*.

c) With $C_{in}=2.2\mu F$, determine the error in LSB due to a ripple of *100Hz* and *500mV* on the power supply.



a) 36.4kHz.

b) 50kHz.

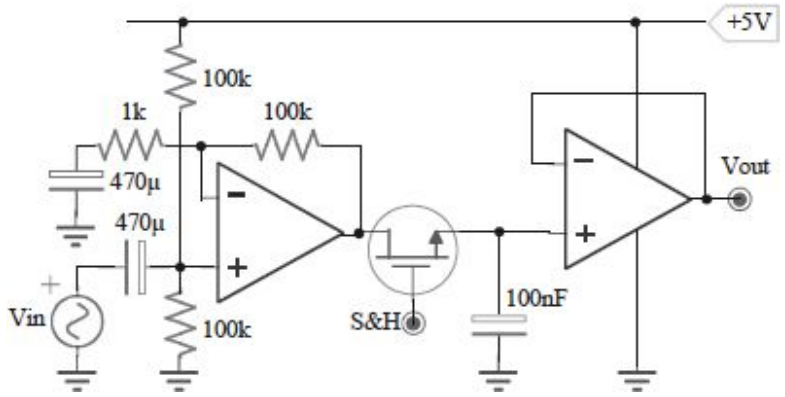
c) 8.5LSB.

Ex. 22

The OpAmps have A_0=*100dB*, *SR=5V/µs*, V_{os}=*10mV*, and I_b=*500nA*. The input is sinusoidal with a peak of *20mV*. The *16bit* ADC has *FSR=5V*. The MOS has R_{on}<*50*Ω.

a) Determine the maximum S&H acquisition time to ensure a precision of *½LSB*.

b) Compute the maximum output errors due to the OpAmp non-idealities, after a T_{hold}=*1µs*.



a) T=58µs

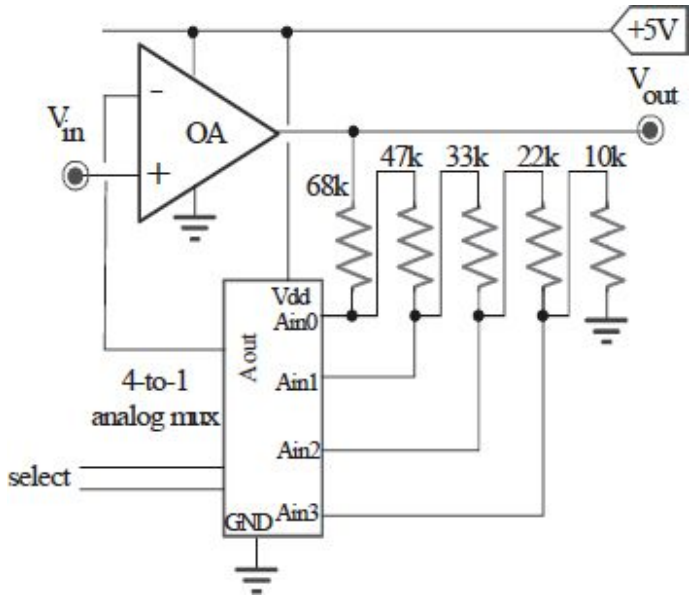
b) $V_{out,OS}$=20mV, $V_{out,bias}$=25mV, $error_{finite\ A0}$=45µV, $error_{droop}$=5µV.

Ex. 23

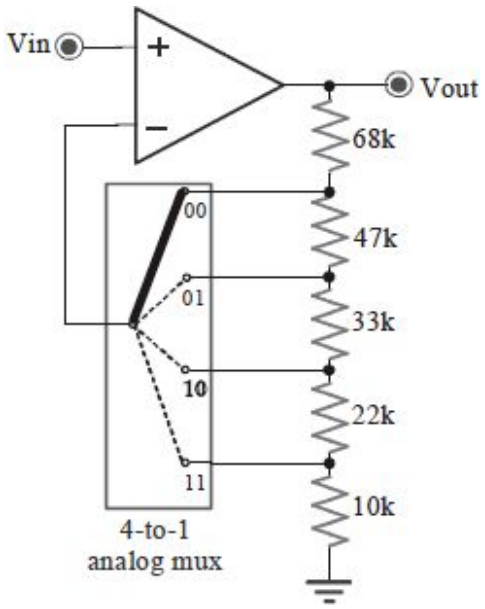
The OpAmp has I_B=*200nA*, V_{OS}=*2mV*, A_0=*100dB*, and *GBWP=50MHz*, and the mux has R_{on}≤*10Ω*, R_{off}≥*1MΩ*, and I_{leak}<*300nA*.

a) Explain the stage operations and compute the gains and close-loop bandwidth values for various selections.

b) Compute the output voltage error when the input V_{in} is connected to *GND*, and the selection is equal to *11*.



a) After a simple analysis, we can say that the circuit is a non-inverting amplifier, as shown in the following figure.



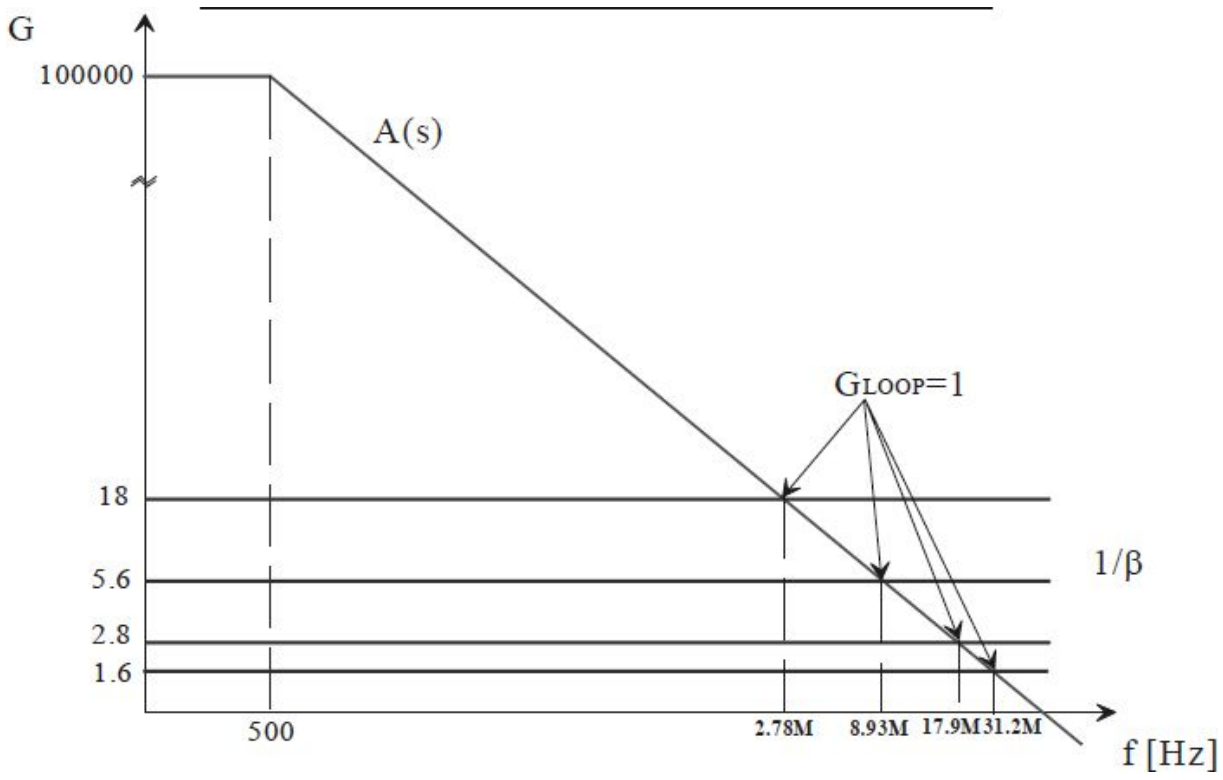
The stage gain, with a good approximation, is that of a normal non-inverting amplifier: $A_v = \frac{V_{out}}{V_{in}} = 1 + \frac{R2}{R1}$. The values of R1 and R2 change with the selection of the analog mux. Remember that a mux connects the output with one single input, in relation to the choice with the select lines. Unlike the digital mux which brings only the logic level selected between the *n* available, an analog mux should bring the input signal to the output, introducing the minimum error.

You can initially compute the ideal gains, neglecting R_{on}, i.e. the maximum resistance of the path selected by the mux and R_{off} i.e. the minimum resistance values with remaining lines unconnected. In this case, the analog mux non-idealities are negligible; in fact, R_{on} is always in series with a larger resistor (the resistor R_{on} sees always R1//R2 towards the ground; the worst case is the selection 10 for which R1//R2 is the minimum and is equal to ~26kΩ), and R_{off} makes a path parallel to the selected line, but, because the line has an order of magnitude greater than the external resistance values, they do not appreciably modify the stage gain. In the following, there are the gains for various selections:

SELECTION	GAIN
00	$Av = 1 + \frac{68k\Omega}{47k\Omega + 33k\Omega + 22k\Omega + 10k\Omega} = 1.6$
01	$Av = 1 + \frac{68k\Omega + 47k\Omega}{33k\Omega + 22k\Omega + 10k\Omega} = 2.8$
10	$Av = 1 + \frac{68k\Omega + 47k\Omega + 33k\Omega}{22k\Omega + 10k\Omega} = 5.6$
11	$Av = 1 + \frac{68k\Omega + 47k\Omega + 33k\Omega + 22k\Omega}{10k\Omega} = 18$

To calculate the bandwidth, we can note that the circuit does not have capacitors which can add poles or zeroes to the stage transfer function. The ideal stage transfer function is 1+R2/R1 until the feedback is good, i.e. the loop gain G_{LOOP} is greater than the unity. The following graph can help to compute the bandwidth values; A(s) is the OpAmp stage, 1/β is the inverse of the feedback transfer function, equal to 1/[R1/(R1+R2)]=1+R2/R1=A_v. Finally, we can remember that the open-loop pole for a compensated OpAmp is simply computed as GBWP/A_0=50MHz/10^5=500Hz. We can obtain the bandwidth values for the various selections:

SELECITON	BANDWIDTH
00	$BP = \frac{GBWP}{A_v} = \frac{50MHz}{1.6} = 31.25MHz$
01	$BP = \frac{GBWP}{A_v} = \frac{50MHz}{2.8} = 17.86MHz$
10	$BP = \frac{GBWP}{A_v} = \frac{50MHz}{5.6} = 8.93MHz$
11	$BP = \frac{GBWP}{A_v} = \frac{50MHz}{18} = 2.78MHz$



b) Because of the non-idealities of the analog mux and the OpAmp non-idealities, the output voltage is not exactly the input voltage, but it is affected by an error due to various contributions. We must consider the bias currents, the OpAmp offset voltages, and the leakage currents of the analog mux. Strictly speaking, we should consider also the voltage drop across R_{on} and the gain change due to the R_{off}, but, as we said, such effects are negligible. To compute the error, we can refer to the following circuit which shows the stage in the required configuration with the abovementioned non-idealities. We can proceed with using the superimposition effect and assuming the worst case, i.e. when all the contributions are added on top of each other with the same sign.

Vos: $Vout = Vos \cdot \left(1 + \frac{R2}{R1}\right) = 2mV \cdot 18 = \pm 36mV$

I_B^+: $Vout = 0$ (actually it depends on the source resistance)

I_B^-: $Vout = I_B - \cdot (68k + 47k + 33k + 22k) = \pm 34mV$ (the current does not flow through 10kΩ)

I_{L1}: $Vout = I_{L1} \cdot 68k = 20.4mV$ (the current does not flow through 47kΩ+33kΩ+22kΩ+10kΩ)
I_{L2}: $Vout = I_{L2} \cdot (68k + 47k) = 34.5mV$
I_{L3}: $Vout = I_{L3} \cdot (68k + 47k + 33k) = 44.4mV$
I_{L4}: $Vout = I_{L4} \cdot (68k + 47k + 33k + 22k) = 51mV$
The output static error is:
$V_{out\ ,\ errore} = 36mV + 34mV + 20.4mV + 34.5mV + 44.4mV + 51mV = 220.3mV$! !

Ex. 24

The circuit must follow the input signal, which ranges in *0÷2V*, with a maximum slope of *10mV/µs*, employing a *dig-pot* with *64taps*.

a) Design the whole circuit.

b) Implement the two digital outputs which provide the input signal slope (Slope=high if increasing, low=decreasing) and the information about the correct work of the circuit (input following), i.e. Tracking=high if the difference is less than *200mV*).

a) A possible circuit implementation is the following: Assuming to use a power supply of +5V, every tap of the dig-pot is equal to 5V/64=78mV. Because the maximum speed of variation in the INPUT signal is 10mV/µs, therefore varies of 78mV in 7.8µs, we chose a CLK to control the DIGPOT with a maximum period of 1µs. We can realize with a NAND astable, resistance of 1kΩ, and 1nF of capacitance). The potentiometer counting direction is set by the input comparator.

b) The digital SLOPE, as we can see, is realized by the comparator output.
So as to make the Tracking signal, we can use a difference amplifier connected to the comparator inputs, i.e. to INPUT and the DIGPOT output. The gain is chosen so that, with a difference of 200mV, the comparator output provides, for example 2.5V, equal to the threshold of the CMOS Schmidt trigger downstream. For these reasons, we can choose the resistors as: R2=R4=330kΩ and R1=R2=27kΩ.
There is the problem of the single power supply +5V. The difference amplifier can switch for positive difference (depending on the inputs connection) while, for negative difference, the output is saturated to 0V. To avoid this problem, we can realize another twin differential stage with inverted inputs. And

instead of a single inverter, you can connect one Schmidt trigger OR to the OpAmps' outputs.
Desiring a threshold switch more precise, instead of a logic gate, we can use a fourth OpAmp connected as a comparator, which exploits an accurate Zener diode as threshold, for example with 3.9V of Zener voltage.

7

DAC converters

"Trouble is her only friend and he's back again
Makes her body older than it really is
And she says it's high time she went away
No ones got much to say in this town
Trouble is the only way is down, down, down
As strong as you were
Tender you go
I'm watching you breathing
For the last time
A song for your heart
But when it is quiet
I know what it means
And I'll carry you home
I'll carry you home

"Carry you home", James Blunt

7.1. General Characteristics

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.

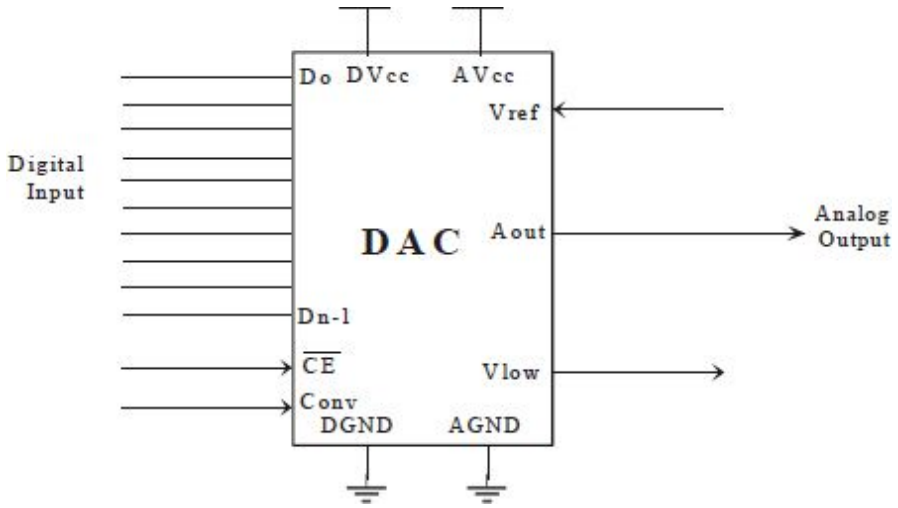


Fig. 7.1: General scheme of a digital/analog converter.

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.

7.2. Quality Factors

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.

n Bits	States 2ⁿ	LSB Weight 1/2ⁿ	LSB Weight in ppm	LSB Weight in % of Full Scale Range	Bit Weight or LSB for 10V FSR	Theoretical Signal/ Quantitization Noise Ratio in dB	Dynamic Range in dB
0	1	1.000	1,000,000	100	10.000 V	1.76	0.00
1	2	0.500	500,000	50	5.000 V	7.78	6.02
2	4	0.250	250,000	25	2.500 V	13.80	12.04
3	8	0.125	125,000	12.5	1.250 V	19.82	18.06
4	16	0.062500	62,500	6.25	625.000 mV	25.84	24.08
5	32	0.031250	31,250	3.125	312.500 mV	31.86	30.10
6	64	0.015625	15,625	1.5625	156.250 mV	37.88	36.12
7	128	0.007812500	7,812.500	0.781250	78.125 mV	43.90	42.14
8	256	0.003906250	3,906.250	0.390625	39.063 mV	49.92	48.16
9	512	0.001953125	1,953.125	0.195312	19.531 mV	55.95	54.19
10	1,024	0.000976562500	976.562	0.097656	9.766 mV	61.97	60.21
11	2,048	0.000488281250	488.281	0.048828	4.883 mV	67.99	66.23
12	4,096	0.000244140625	244.141	0.024414	2.441 mV	74.01	72.25
13	8,192	0.000122070313	122.070	0.012207	1.221 mV	80.03	78.27
14	16,384	0.000061035156	61.035	0.006104	610.352 µV	86.05	84.29
15	32,768	0.000030517578	30.518	0.003052	305.176 µV	92.07	90.31
16	65,536	0.000015258789	15.259	0.001526	152.588 µV	98.09	96.33
17	131,072	0.000007629395	7.629	0.000763	76.294 µV	104.11	102.35
18	262,144	0.000003814697	3.815	0.000381	38.147 µV	110.13	108.37
19	524,288	0.000001907349	1.907	0.000191	19.073 µV	116.15	114.39
20	1,048,576	0.000000953674	0.954	0.000095	9.537 µV	122.17	120.41

Tab. 7.1: DAC converter resolution as the number of bits changes.

Scale	*Offset binary*	*1s complement*	*2s complement*	*Sign magnitude*
+FSR −LSB	11111111	01111111	01111111	11111111
$+\frac{3}{4}$ FSR	11100000	01100000	01100000	11100000
$+\frac{1}{2}$ FSR	11000000	01000000	01000000	11000000
$+\frac{1}{4}$ FSR	10100000	00100000	00100000	10100000
+0	00000000	00000000	00000000	10000000
−0	—	11111111	—	00000000
$-\frac{1}{4}$ FSR	01100000	11011111	11100000	00100000
$-\frac{1}{2}$ FSR	01000000	10111111	11000000	01000000
$-\frac{3}{4}$ FSR	00100000	10011111	10100000	01100000
−FSR + LSB	00000001	10000000	10000001	01111111
−FSR	00000000	—	10000000	—

Offset Binary: natural binary code in which the code represents analog values between –FS and +FS. The code 000...000 corresponds to –FS.

Complementary binary: binary code which is the logical complement of straight binary. All 1s became 0s and vice versa.

Two's complement: bipolar binary code in which positive and negative codes of the same magnitude sum to all zeros plus a carry. The two's complement is obtained by complementing the MSB of the offset binary code.

One's complement: bipolar binary code in which positive and negative codes of the same magnitude sum to all ones.

Binary coded decimal: binary code used to represent decimal numbers in which each digit, from 0 to 9, is represented by four bits weighted 8-4-2-1. Only the lowest 10 of the possible 16 values are used.

3795=011 0111 1001 0101

Sign magnitude: code where the MSB denotes positive (1) and negative (0) polarities. Natural binary or BCD code may be used to represent the magnitude.

Tab. 7.2: Examples of digital codes adopted in converters.

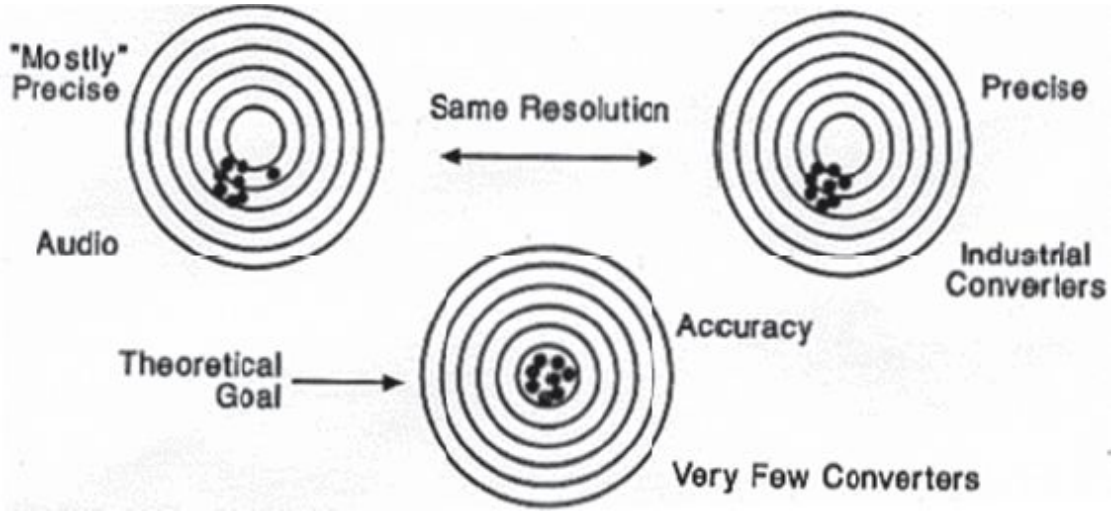


Fig. 7.2: Different meanings of accuracy, resolution, and precision.

Traditionally, there are two classes of converters: those that require a high accuracy (i.e. the DAC must represent the actual value of the input) and those that require a precision within a certain range of values, namely the “repeatability” of the output in response to a given input. DACs are generally “precise” within a value specified by the manufacturer; sometimes, there is an additional adjustment to make the desired accuracy last long.

As you can see from this quick analysis, resolution is an insufficient indicator to define the performance of converters since converters may have applications in various fields. An example of a very specific market is that of converters for audio applications; in this field, it is important that the converters are, first of all, precise. In fact, since the analog signal has to undergo many digital processing steps, the output will be acceptable only if all the conversions have been precise enough. In other areas, instead, the request will be to be accurate within a certain error along the whole transfer function. The ultimate desire is still to design a converter in which the absolute accuracy tends to the theoretical one. In the audio field, for example, where the resolution required is typically of 16 bits, “perfection” is not reachable. Moreover, the continuous need to reduce costs has led to

choices always more penalizing for the qualities of DACs, such as smaller packages that do not allow offset and gain adjustments and faster testing that measure only the main variables and only some of the frequencies and temperatures (without warranties or remedies for any drift). That is why these audio DACs do not have the specifications required for most demanding applications, such as those in industrial field and those for precision instrumentation. To this end, different manufacturers warn the customer to “use audio products for industrial purposes at your own risk!”

7.3. ARCHITECTURES

We now present the architectures of some commercial DACs. The aim is to provide a comprehensive overview of how these devices work and what their limitations in static terms (typically the accuracy) and dynamic terms (conversion rate) are.

7.3.1 Weighted R DAC

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 $2^{n-1}\cdot$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;

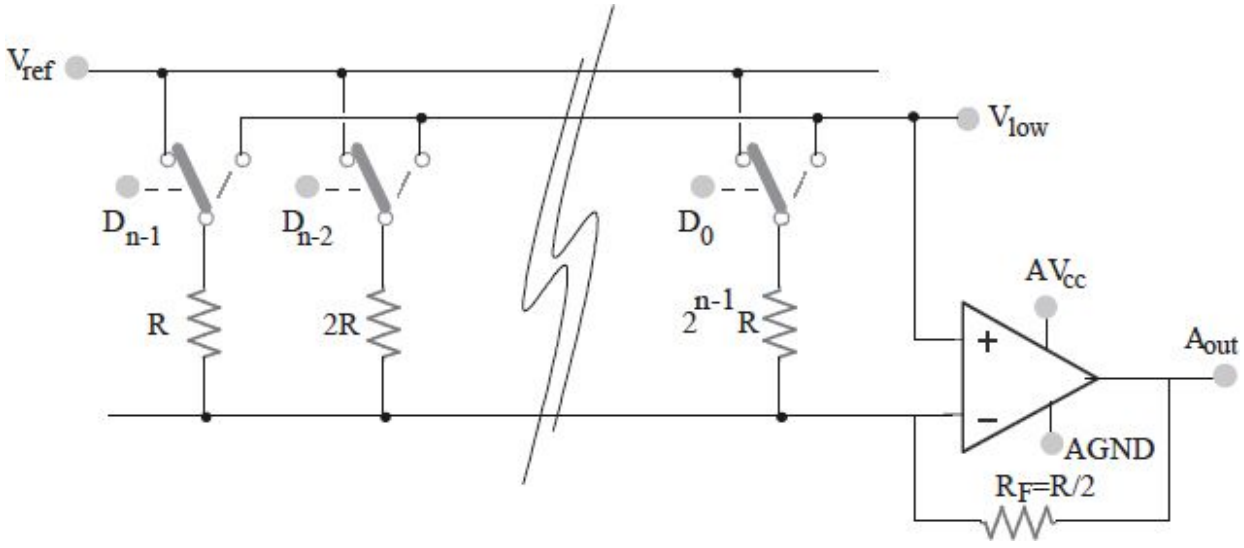


Fig. 7.3: Weighted R DAC and its schematics.

• R_{ON}: each R_{ON} of the MOS switches weighs differently when in series with R rather than with $2^{n-1}\cdot R$;
• 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.

7.3.2 DAC voltage scaling

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.

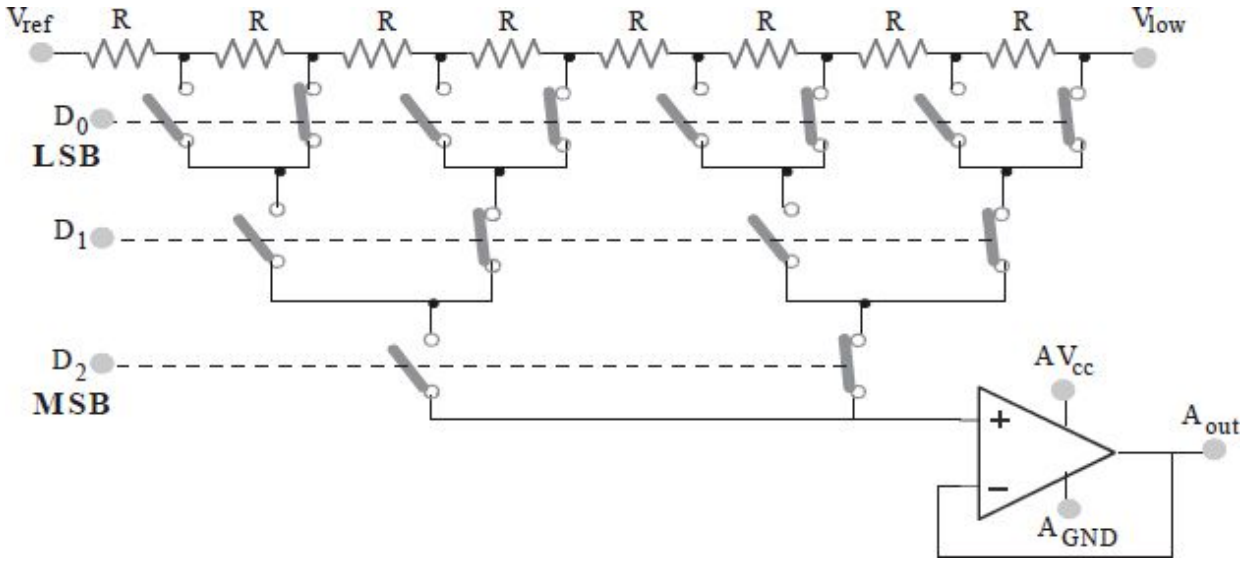


Fig. 7.4: DAC voltage scaling schematics.

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.

7.3.3 Serial DAC

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:

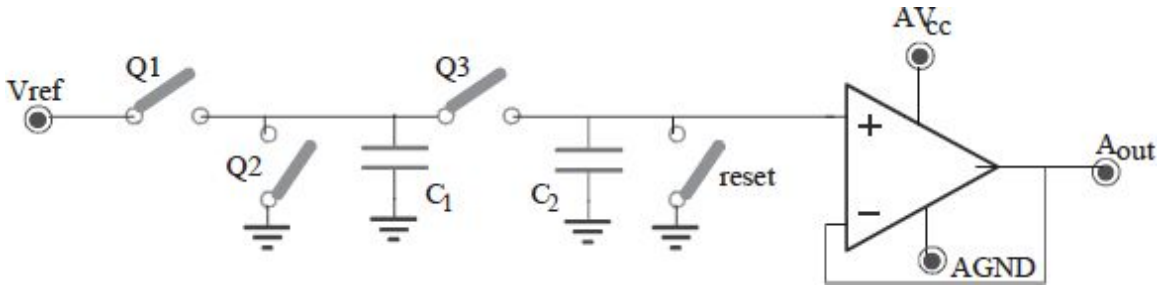


Fig. 7.5: Serial DAC schematics.

7.4. Static Errors and Non-Linearity

7.4.1 Offset and Gain Errors

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.

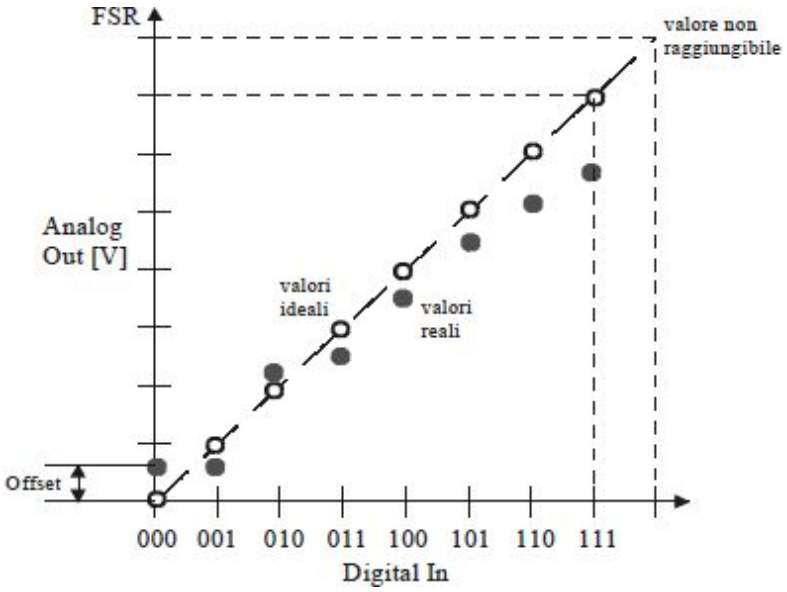


Fig. 7.6: Characteristic of a real DAC with gain and offset errors; empty dots represent the ideal values of the conversion while full dots represent the real values.

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.

7.4.2 Integral Non-Linearity INL and Differential Non-Linearity DNL

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.

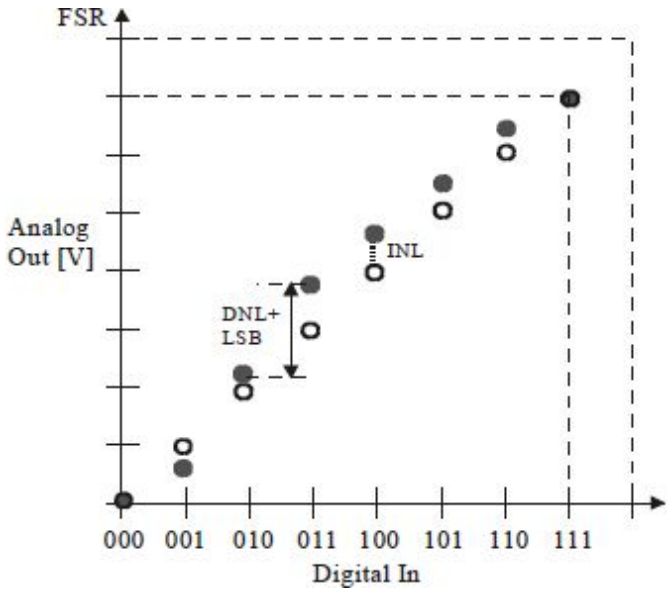


Fig. 7.7: Characteristic of a real DAC, including integral and differential non-linearity. Gain and offset errors have already been corrected.

7.4.3 Monotonicity

It is important that the differential non-linearity does not exceed the value of 1LSB; otherwise, we may run into cases of non-monotonicity of the characteristic of the converter (i.e. the 'step' goes down instead of going up). This can be particularly harmful for control theory applications since it can cause the triggering of oscillatory patterns. In the following, we can see a non-monotonic characteristic of a DAC (Fig. 7.8) and three diagrams (Fig. 7.9) representing possible errors due to non-linearity and the relationships between these and the concept of monotonicity just expressed. In the first diagram of Fig. 7.9, we have, in terms of LSBs, the difference between the obtained characteristic with respect to the ideal for each digital input. The DNL, shown in the second diagram, is obtained by subtracting 1LSB from the values of the first. Note that a value of DNL lower than –1 implies non-

monotonicity. Finally, adding together DNL values obtained up to that code, we get the third diagram, representing the INL.

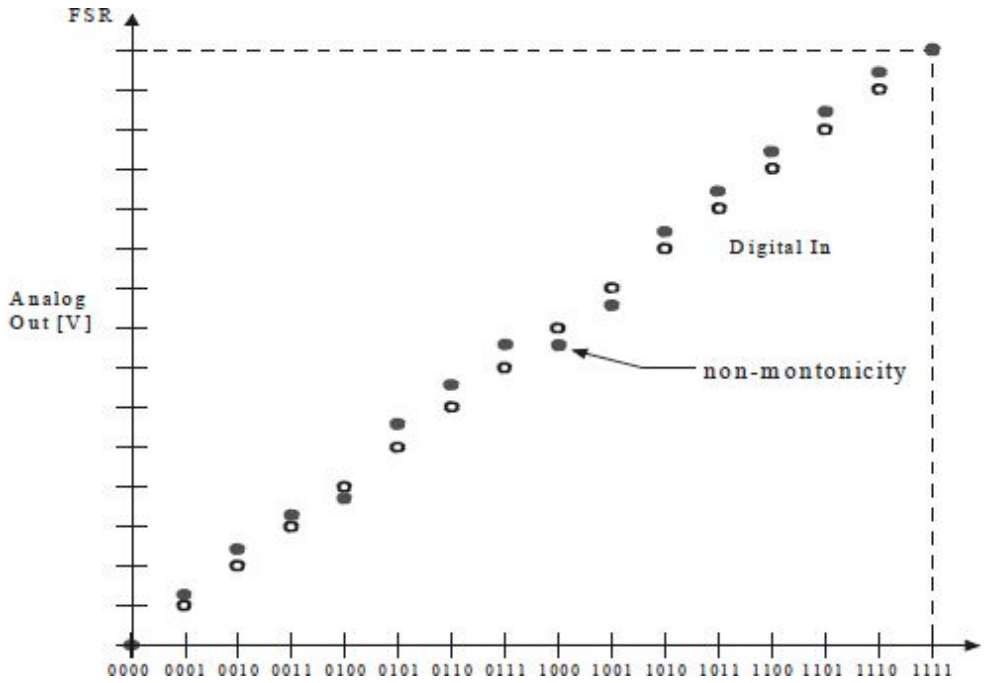


Fig. 7.8: Example of non-monotonicity in a DAC.

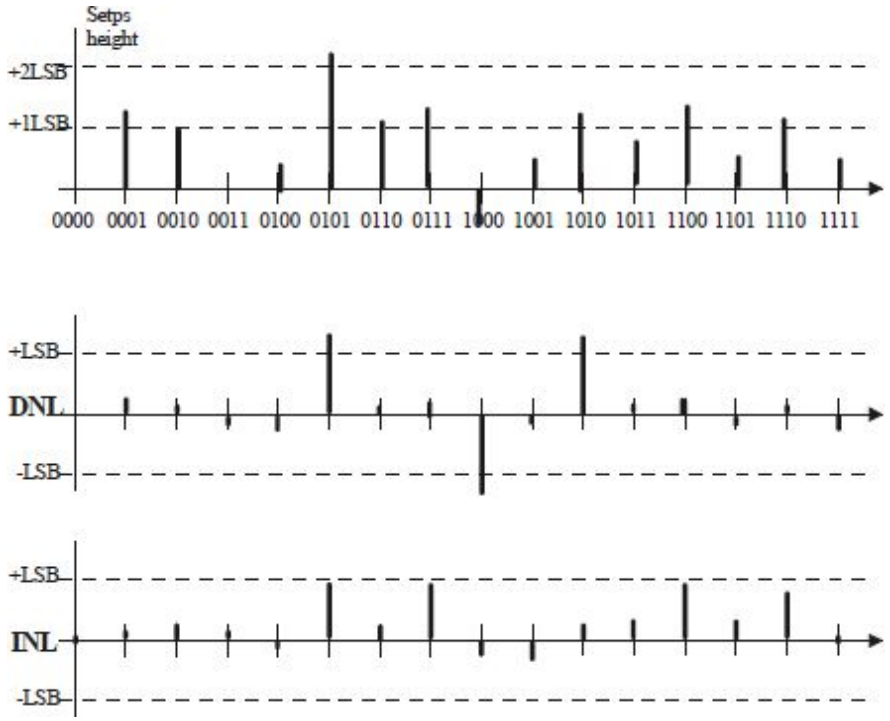


Fig. 7.9: Non-linearity error diagrams.

7.4.4 Thermal Drifts

The dependence of some parameters on the operating temperature of the circuit may cause a clear deterioration in the conversion. Both the DNL and the INL are sensitive to changes in temperature with typical coefficients of 5ppm/°C (each 25°C the error reaches ½LSB in a 12bit DAC!). Temperature Coefficients (T.C.) are specified also for V_{ref} and the output resistance of the DAC. The stability of V_{ref} is a parameter of great importance since the precision of the DAC directly depends on it. In fact, even the LSB of a DAC with only 8bits is equal to 0.39% of V_{ref}; therefore, a variation of 2% (e.g. due to the ripple of the power supply to which V_{ref} is connected) would cause an error of 5LSB (2%/0.39%)! Therefore, it is common to find DACs with an integrated bang-gap reference circuit which internally generates a reference voltage of about 2.5V with a temperature coefficient of only about 50ppm/^{o}C. Despite this high precision of the reference, a change in temperature of 50^{o}C (for example, going from a night temperature of –10°C to a day temperature of +40°C) causes a full scale error of about 2500ppm, similar to the resolution of an 8-bit DAC (1LSB = 3906ppm).

7.4.5 Output Impedance

In each DAC, the maximum current absorption must be fixed to ensure a voltage drop across the output resistor lower than, for instance, 1LSB. The DAC ZN558D (Fig. 7.10) has output resistance of 4kΩ with V_{ref}=5V. To ensure an error compatible with the resolution of the DAC, it is necessary that the voltage drop across this resistance is less than ½LSB. In this case, the current should be kept less than 5µA (this requirement is not easy to achieve, even for high input impedance buffers particularly if in inverting configuration); so, if we had a load R_L lower than V_{ref}/5µA=1MΩ, it would be necessary to interpose a suitable decoupling OpAmp between the DAC output and the load. Note that the current requirement is extremely stringent, even for a high impedance general-purpose buffer.

Another problem is due to the fact that R_{out} has a temperature coefficient that amounts to around 0.2%/^{o}C. Assuming that the load R_L has T.C.=0, i.e.

it is invariant with temperature, the thermal drift causes a change in the output equal to:

$$\Delta V_{out} = V_{ref} \cdot \frac{\Delta R_{out}}{R_L + R_{out}} \cong V_{ref} \cdot \frac{R_{out}}{R_L} \cdot \frac{0.2}{100} \cdot \Delta T$$

With 8 bits and R_{out}=4kΩ, with a temperature variation of 50°C, we have to choose a minimum value of 20kΩ for R_L to ensure an error of no more than 1LSB with V_{ref}=5V.

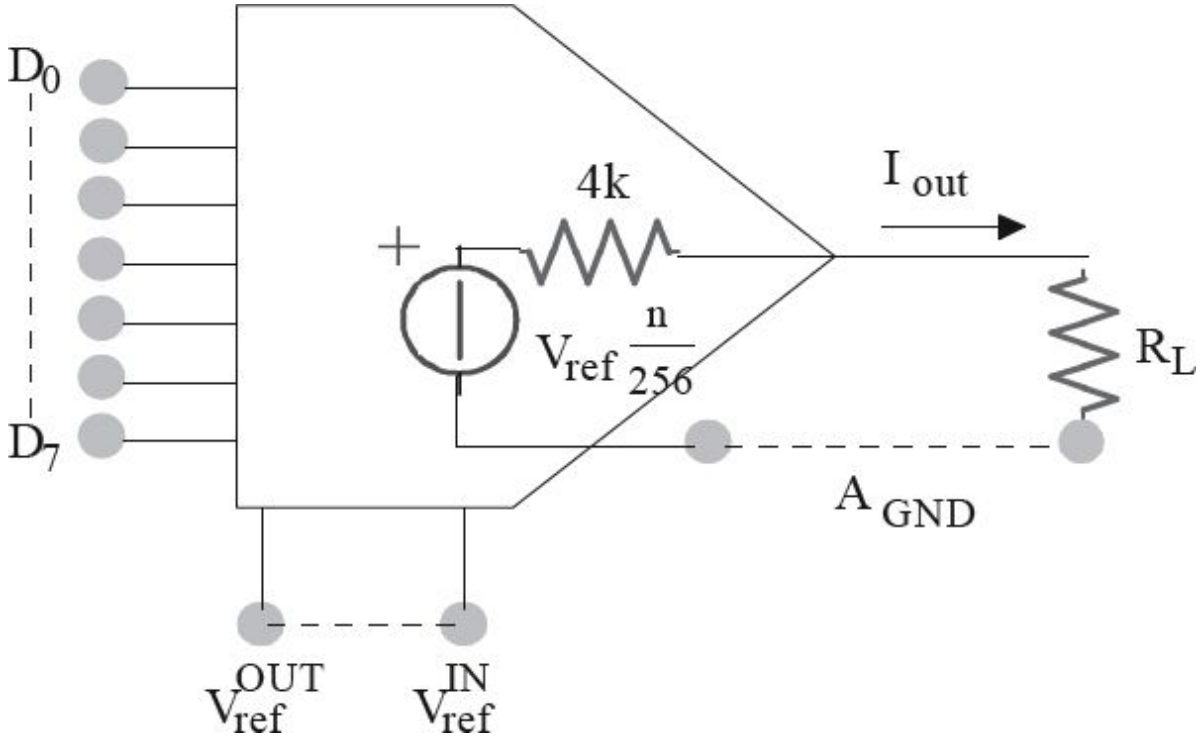


Fig. 7.10: Scheme for the evaluation of the minimum load resistance for a DAC.

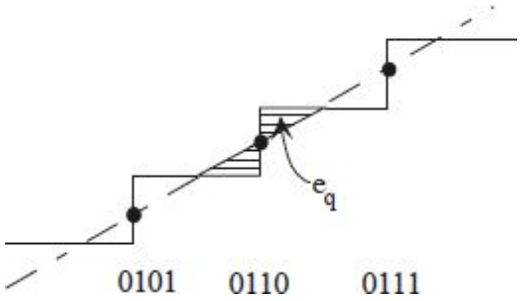


Fig. 7.11: Quantization error.

7.5. Signal to Noise Ratio

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.

7.5.1 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:

$$\sigma^2 = \int_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} \varepsilon_q{}^2 \cdot \frac{1}{LSB} \cdot d\varepsilon_q = \frac{1}{LSB} \cdot \left[\frac{\varepsilon_q^3}{3} \right]_{-\frac{LSB}{2}}^{+\frac{LSB}{2}} = \frac{1}{LSB} \cdot \frac{2}{3} \cdot \frac{LSB^3}{8} = \frac{LSB^2}{12}$$

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).

7.5.2 Theoretical 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:

$$SNR\big|_{theoretical} = 10\log\frac{\text{Signal Power}}{\text{Quantization Noise Power}} = 10\log\frac{\left(\frac{FSR/2}{\sqrt{2}}\right)^2}{\frac{LSB^2}{12}} = 10\log\left(\frac{12}{8}\cdot 2^{2n}\right) = 6.02\cdot n + 1.76 \quad [dB]$$

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.

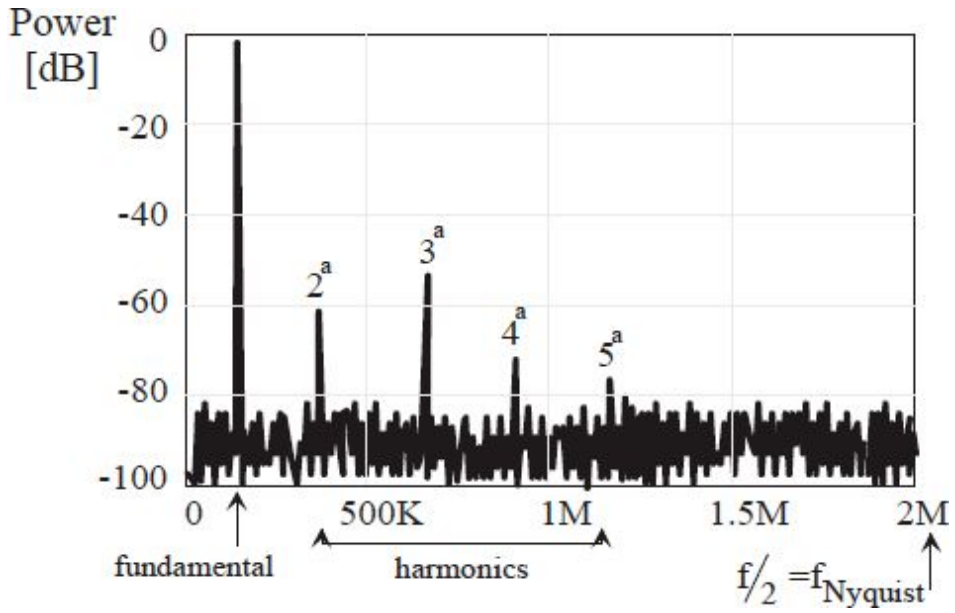


Fig. 7.12: Spectrum of the output signal of the DAC: we can see the sine wave of interest, the spurious harmonics, and the background noise; each histogram has a width that is called bin width.

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 $SNR_{theoretical}=6.02\cdot 8+1.76=50dB$, 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.

7.5.3 Signal to Noise and Distortion 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:

$$\text{Noise Floor} = -\{6.02\cdot n + 1.76 + 10\cdot \log N\} = -\text{SNR}\big|_{\text{theoretical}} - 10\cdot \log N$$

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.

7.5.4 Harmonic Distortion

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:

$$THD = \frac{\sqrt{\sum_{i=2}^{9} V_i^2}}{V_1} = 20 \cdot \log \sqrt{10^{\left(\frac{2^a \text{ harmonic (dB)}}{20}\right)^2} + 10^{\left(\frac{3^a \text{ harmonic (dB)}}{20}\right)_2} + \ldots} \quad [\text{dB}]$$

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 $V \cdot \sin[2\pi(f_1+f_2)t]$.

7.6. Dynamic Parameters

Similar to any other digital device, for DACs, it is crucial to give the performance in terms of conversion speed and settling time and consideration about the quality and stability of the data provided at the output.

7.6.1 Settling Time

Among the dynamic parameters that characterize the functioning of a DAC, there is the *settling-time*, which is the time needed to have a stable value at the output after a maximum amplitude transition (e.g. from 0 to FSR-LSB) inside a specified error band which is usually set at ±LSB/2 (Fig. 7.13).

In the transition between 000…0 and 111…1, a time t_s is taken, also related to the bandwidth of the used OpAmp. Note that after t_s, the FSR must settle within ±LSB/2.

7.6.2 Glitch

A phenomenon that affects the dynamic behavior of a DAC is the presence of *glitches* in the value of the output. They are caused by the non-simultaneous switching of the input bits or of the ones inside the DAC. In this way, in fact, completely meaningless "transient" codes can be generated, which the DAC equally converts with disastrous effects. For instance, changing from the word 1000 to the word 0111, it is possible that the three 0s and the 1 switch with a slight delay producing, for an instant, the word 1111 with a subsequent impulse at the output, as shown in Fig. 7.14. The glitch is all the more likely, the higher the number of bits that change their value from 0 to 1 or vice versa is. A quantity to assess the disturbance introduced by the glitch is the area of the impulse, which is called "energy" (even if it does not have its dimensions).

To avoid the appearance of glitches at the output of the circuit, it is possible to introduce a sampler, such as an S&H, at the output of the DAC, as shown in Fig. 7.15.

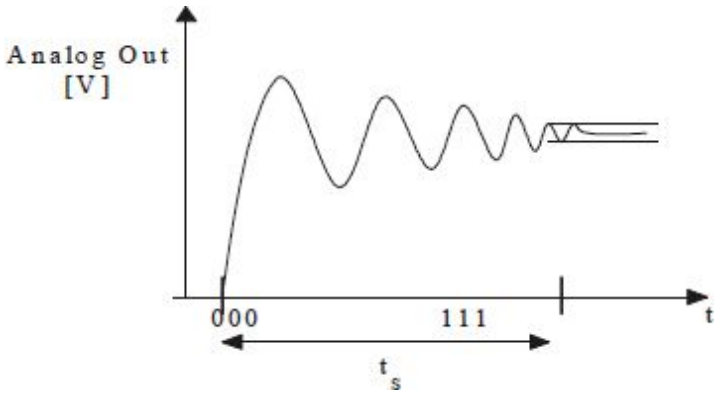


Fig. 7.13: Settling-time.

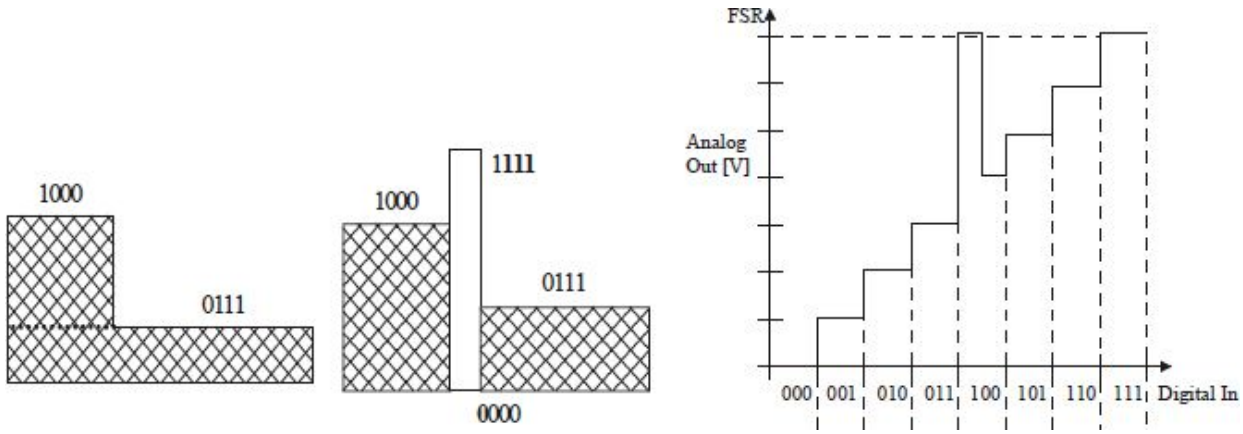


Fig. 7.14: Generation of a spurious code during the switching (indicated in bold are the bits that switch first) and transient effect on the dynamic characteristic of the DAC (where the individual codes are made to increase by a counter placed at the DAC input).

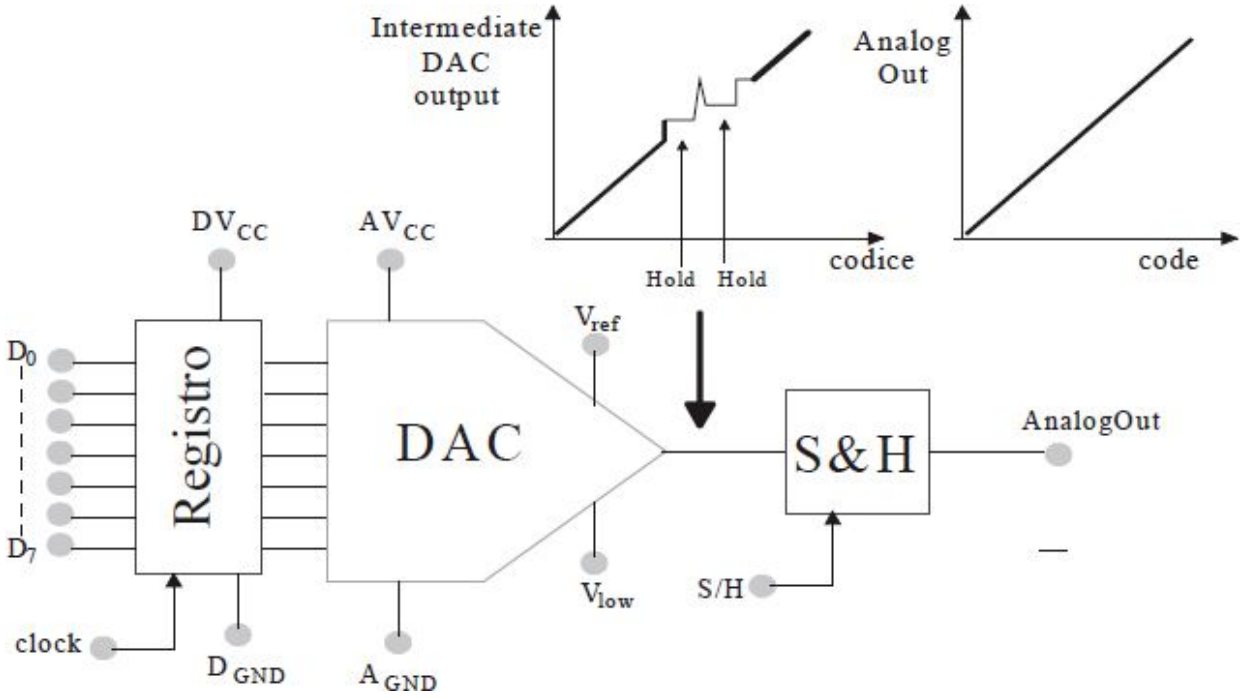


Fig. 7.15: The introduction of an S&H eliminates the problem of glitches.

A circuit able to generate arbitrary functions stored in a ROM memory is proposed in Fig. 7.16. The first counter scans the waveform at the desired sample rate (f_{clock}) while the second counter selects the memory bank (i.e. the upper part of the address) from which the waveform is taken. In this way,

it is possible to recall one of the different waveforms stored in a ROM simply by pressing the selection button.

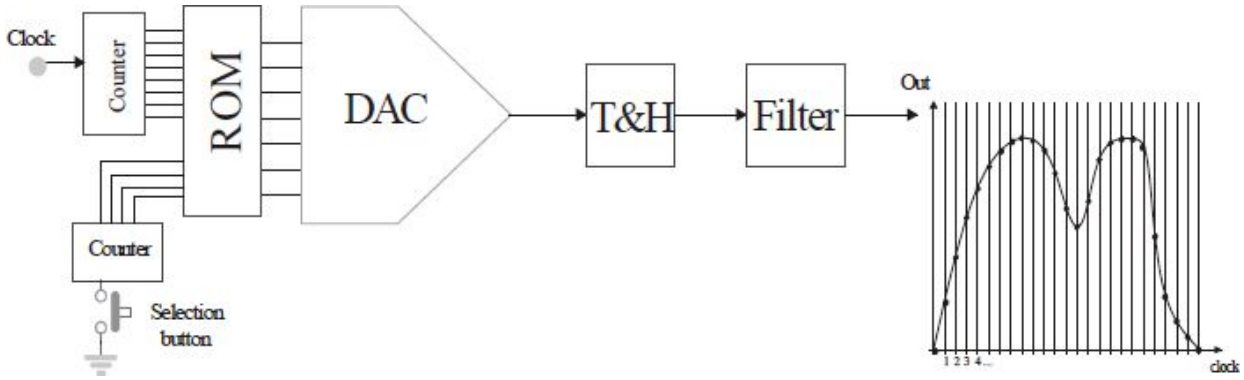


Fig. 7.16: Example of waveform generators, with waveforms stored in a ROM memory.

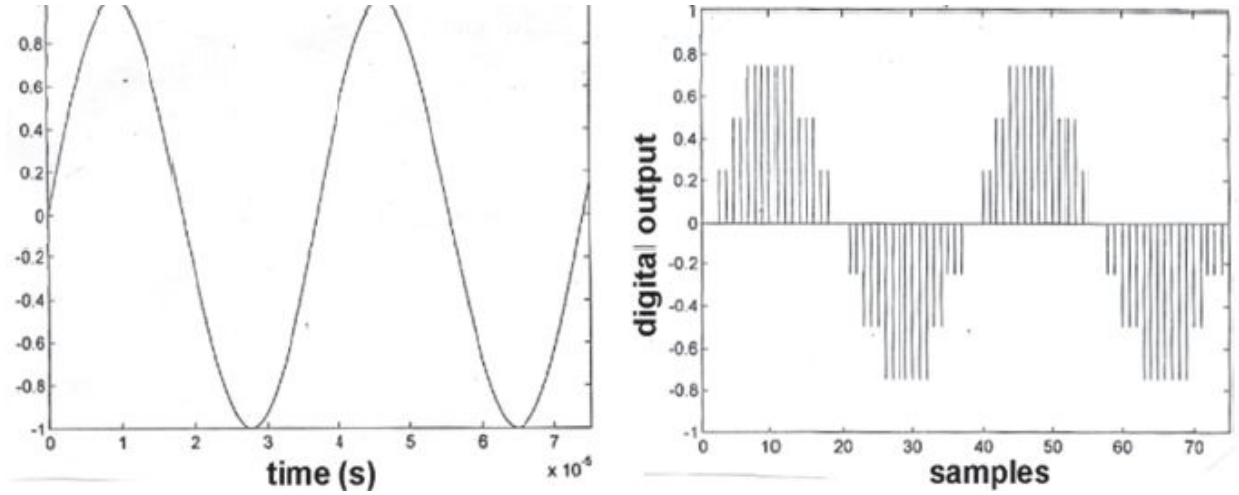


Fig. 7.17: Analog input sine wave (on the left) and sampled with three bits (on the right).

7.7. EXAMPLE

Consider a sine wave of 1V peak at a frequency of 27kHz (Fig. 7.17, on the left), which is converted to digital with only three bits and which generates a numerical code ranging from 000 to 111 (Fig. 7.17, on the right). The input sine wave has a period T=37μs with V_p=1V and f_{CK}=27kHz. Below, we can see the digital sine wave having T_s=1μs and f_s=1MHz; note the 8 levels given by the 3 sampling bits (2^3=8). What is done is to send the analog sine wave to a spectrum analyzer and the digital sine wave to the FFT

and DFT. To perform the FFT of these samples, we have to choose a time period long enough to acquire N digital samples, which, in our case, is 75.

Using a sampling frequency of 1MHz, the samples will follow one another every microsecond. Using only 512 samples in the time domain, 512 samples are obtained in the transform domain, as shown by the FFT in Fig. 7.18, on the left. The bin-width is $f_s/N=10^6/512=1953Hz$. In this way, the resolution is very poor, as is visible in the following graph where we have N=512 samples, SNR=-20dB (with S=0dB normalized), and a noise-floor of –44dB (not 20dB since we go down by 256 histograms, i.e. –24dB).

To increase the frequency resolution, one has to increase the number of frequency samples of the FFT; using, for example, N=4096 samples in time, corresponding to a time window of about 4ms, the spectrum in Fig. 7.18, on the right, is obtained. Now the width of each histogram is only 244Hz, and the level of the background noise is located at $-(6.02\cdot3+1.76+10\cdot\log N)=-56dB$.

To lower the noise level, it is necessary to increase the number of samples or to use a higher number of bits if we sample the same sine wave. With 10 bits, for example, the digital version would appear at the input of the DAC, as in Fig. 7.19.

Looking at the spectrum in Fig. 7.19, the improvement between three and ten bits is equal to $7\cdot6.02=42dB$; in fact, the noise-floor has now been considerably lowered, and it is equal to $6.02\cdot10+1.76+10\cdot\log4096=98dB$. Note that the power of the signal is given by $(FSR/2\cdot\sqrt{2})^2$ if we have a sine wave with amplitude of FSR; in the histogram, that power is put at 0dBc. Moreover, the SNR is given by the sum of the noise histograms (with the signal always at 0dBc).

By performing a spectral analysis of the analog signal, at the output of the DAC, new peaks can be seen, which are at frequencies multiple of the fundamental at its input. These are due to non-linearity of the DAC. An example is shown in Fig. 7.20, where the output density of a DAC with n=10bits, f_s=10MHz, and N=4096 is represented.

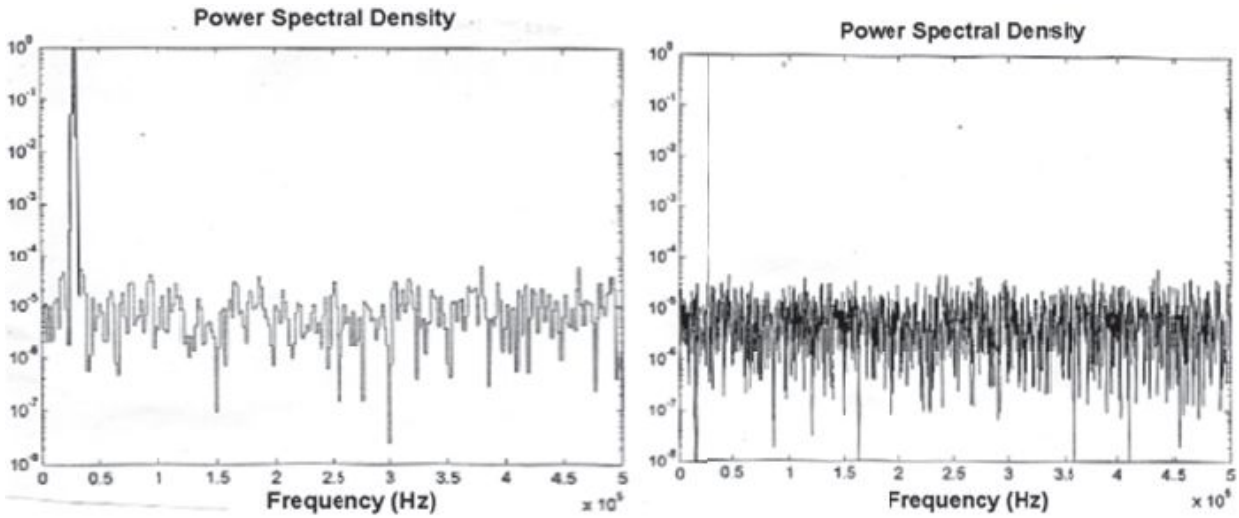


Fig. 7.18: Power spectral density of the sine wave sampled with three bits, which is calculated by computing the FFT over N=512 samples (on the left) and N=4096 samples (on the right).

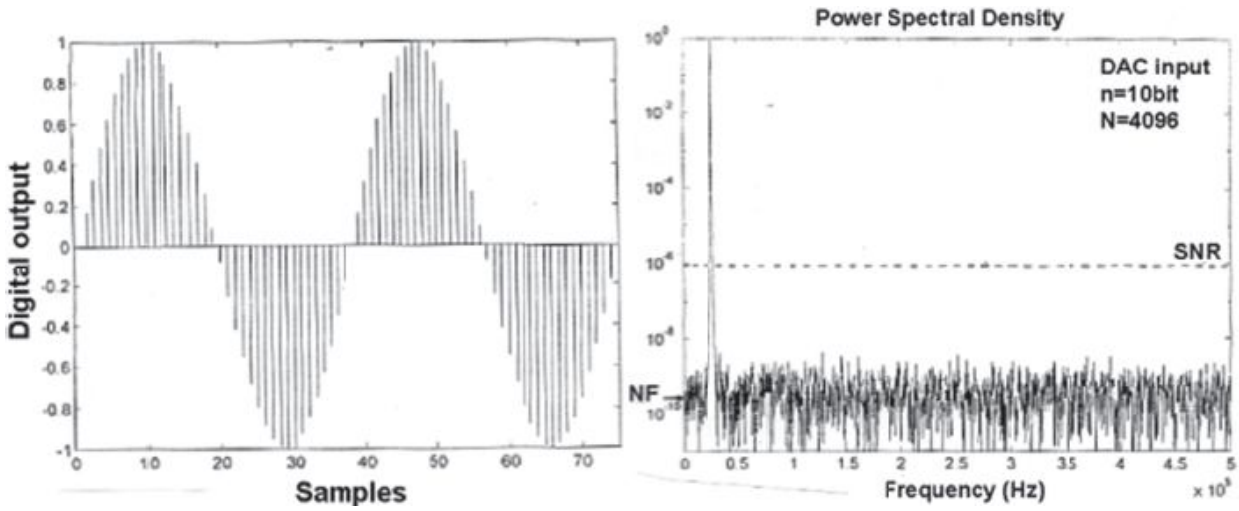


Fig. 7.19: Sine wave at f_{in}=27kHz, sampled with 10 bits and f_s=1MHz (on the left) and the corresponding spectrum, calculated over N=4096 samples (on the right).

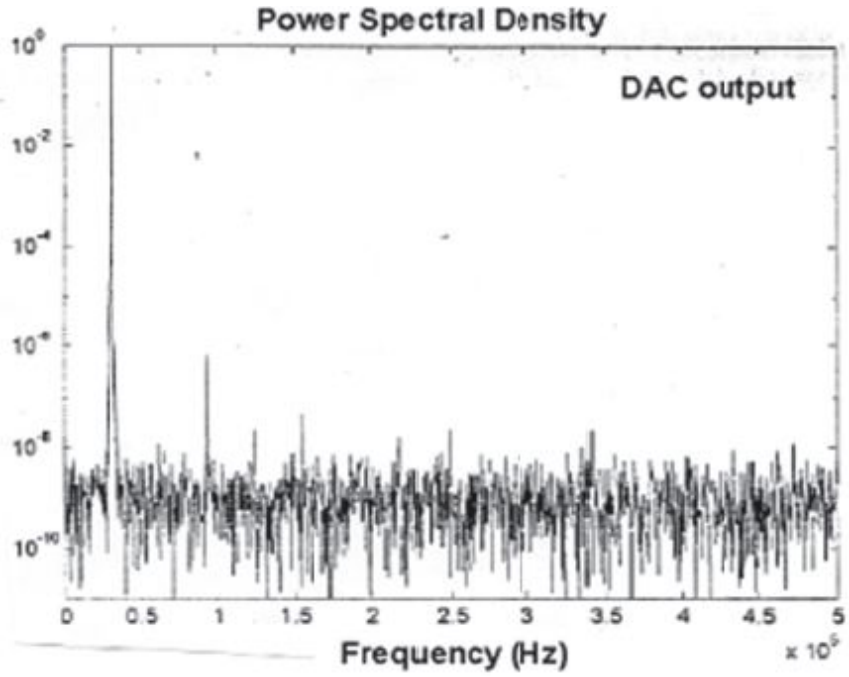


Fig. 7.20: Spectral density of the sine wave at the output of the DAC

The noise-floor is now placed at –90dB, although, at the input of the DAC, there are only 10 bits (with a theoretical SNR of 6.02n+1.76=62dB), and the spectrum has been computed with 4096 samples (for a theoretical achievable noise-floor of -62dB-33dB=-95dB). Therefore, the real SNR at the DAC output is 95dB-90dB=5dB worse.

Assuming that the FSR is 2V with V_p=1V, the output noise power will not be:

$$\sigma_{noise.in} = \sqrt{\sigma^2} = \sqrt{\frac{LSB^2}{12}} = \sqrt{\frac{\left(\frac{FSR}{2^{2n}}\right)^2}{12}} = 0.56mV_{rms}$$

(where σ^2=LSB^2/12 is the quantization noise power), but 5dB worse, that is:

$$\sigma_{noise.out} = \sigma_{noise.in} \cdot 10^{\frac{5dB}{20}} = \sigma_{noise.in} \cdot 1.78 \cong 1mV_{rms}$$

It is important to remember that the power is quadratic, and, with N=1mV^2 and σ^2=(0.56mV)^2, a noise of 0.5mV was added since (0.56mV)^2+

$(0.5mV)^2=1mV^2$.

The same result can be obtained by observing the spectrum: the power of the fundamental is 0dB, equal to $1V/\sqrt{2}=707mV_{rms}$ (where the noise of the DAC is $N=10^{90/10}\cdot 4096\cdot(FSR/2\cdot\sqrt{2})^2$), while the noise floor placed at -90dB reveals that each histogram brings a power contribution of $(707mV)^2/10^{90/10}=(22.3\mu V_{rms})^2$. Using 4096 histograms, the total noise will be $22.3\mu V\cdot\sqrt{4096}=1mV_{rms}$.

Nonetheless, we have to be very careful because, with few samples, the background is high, and there is the risk that it may hide other harmonics. However, the only SNR is not significant because it does not provide information about the harmonic distortion; to this end, it is necessary to calculate also the THD (Total Harmonic Distortion). From the graph, we can note that the second harmonic is at –80dB, the fourth at –77dB, the fifth at –73dB, the sixth at –80dB, the seventh at –78dB, and the eighth at –75dB. This results in a THD=-60dB. The total power of the harmonics is $\sigma=(1mV)^2$, found by adding all the various contributions. The SiNAD is then:

$$SiNAD = 10\log\left(\frac{\left(\frac{V_p}{\sqrt{2}}\right)^2}{\sigma_{noise}^2+\sigma_{thd}^2}\right) = 54dB$$

corresponding to 8.7 “effective” bits (in fact $6.02\cdot n_{eff}+1.76=6.02\cdot 8.7+1.76=54dB$, see the chapter concerning ADCs). Hence, it is as if we lose 1.3 bits because of noise superimposed to the signal (which increases the noise-floor) and of harmonic distortion.

7.8. EXERCISES

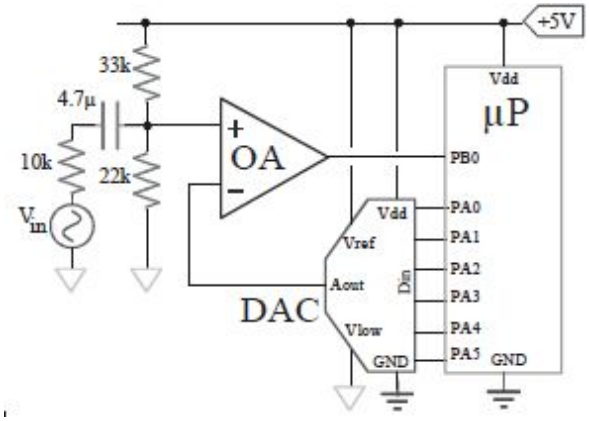
Ex. 1

The *6bit* DAC is driven by the port A of the µC.

a) Plot the transfer function of the input passive network.

b) With no input signal, determine the code range which maintains the input PB0 at the level *high*.

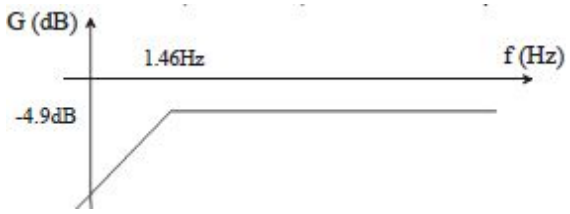
c) When the µC applies PA5:PA0=100000_2=32_{10}, determine the intensity of the input signal V_{in}, which manages to make the pin PB0 switch (assume the capacitor "closed").



a) Due to the presence of a decoupling input capacitor, the transfer function has a dc gain equal to zero (zero at the origin) and a high frequency transfer equal to:

$$\frac{V_{OUT}}{V_{IN}} = \frac{13.2k\Omega}{10k\Omega + 13.2k\Omega} = 0.57 = -4.9dB$$

Remember, in fact, that calculating the frequency behavior of a node connected to the ground and that connected to the power supply (both voltages at zero frequency) is equivalent. From what has been said about the general trend of the transfer function, we can deduce the presence of a pole. This pole is given by the series of the 10kΩ resistance and the parallel of the two resistance values (13.2kΩ) and C=4.7µF, so f_{POLE}=1/2πRC=1.46Hz.



b) With no applied input, at the non-inverting terminal of the OpAmp, there is a voltage given by the resistive partition between 33kΩ and 22kΩ. V_{IN} is therefore $5V \cdot \frac{22k\Omega}{22k\Omega + 33k\Omega} = 2\text{V}$. Since the dynamic range of the DAC is 0-5V, 2V represents 2/5 of the dynamics. The DAC has 6 bits, then 64 codes; thus, 2V represents the code number 64·2/5=25,6. This means that PB0 remains high until the code 25 (011001_{BIN}) is included.

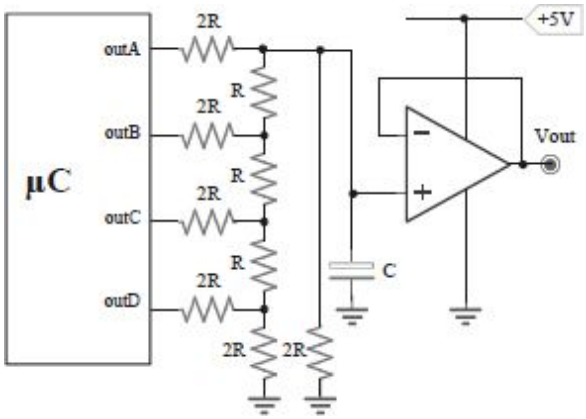
c) The code 100000_{BIN} is the center of the dynamics of the DAC, which is a voltage equal to 2.5V. Since, as seen, the resistive partition moves the input signal around 2V, at the non-inverting terminal, a voltage of 0.5V is sufficient. Since the input network at the mid-frequency has a gain of 0.57, to get at the terminal a voltage equal to 0.5V, it is necessary to apply a voltage with a peak equal to 0.5V/0.57=+8.77V.

Ex. 2

Realize the R-2R DAC with *4 bits*, using a µC, with I_{outmax}*=1mA*.

a) Size the resistors and approximately calculate the tolerance that ensures *±0.5LSB*.

b) Desiring to convert *1,000 nibble/s*, size *C* to smooth the output, ensuring a *4bit* precision.

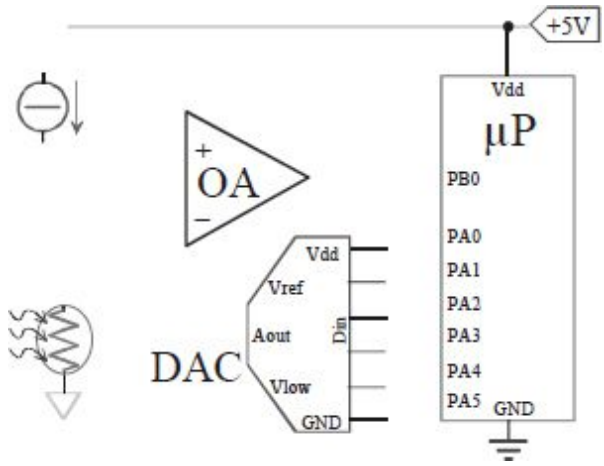


Ex. 3

The µC has to control the excitation current of a resistive sensor through a *6bit* DAC, from *0.5mA* to *32mA*.

a) Design the programmable generator, capable of delivering a current to the sensor connected to the ground with a nominal value of *10÷100Ω*.

b) Amplify the signal in order to send it to the DAC inside the µC through PB0, exploiting the whole *0÷5V* dynamics.



8

ADC Converters

“You’re just too good to be true
can’t take my eyes off of you
you’d be like heaven to touch
I wanna hold you so much
at long last love has arrived
and I thank God I’m alive
you’re just too good to be true
can’t take my eyes off of you
Pardon the way that I stare
there’s nothing else to compare
the sight of you leaves me weak
there are no words left to speak
but if you feel like I feel
please let me know that it’s real
you’re just too good to be true
can’t take my eyes off of you
I need you baby and if it’s quite alright
I need you baby to warm the lonely nights
I love you baby, trust in me when I say: ok
Oh pretty baby, don’t let me down, I pray
oh pretty baby, now that I’ve found you, stay
and let me love you, o baby, let me love you, o baby”

“Too good to be true”, Lauryn Hill

8.1. GENERALITIES

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).

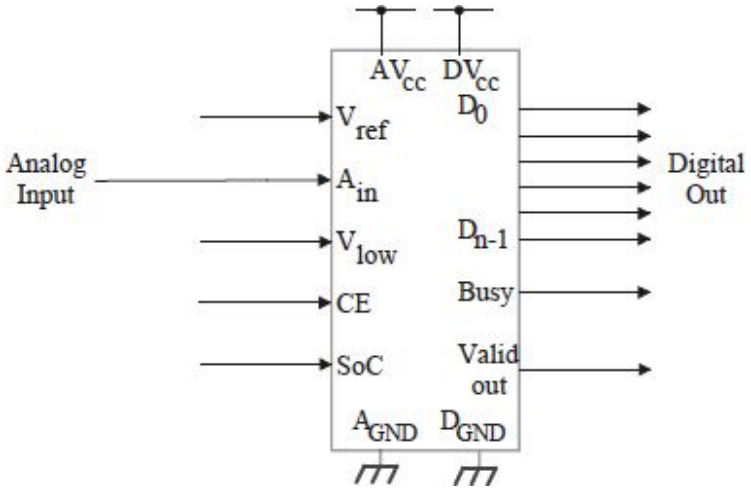


Fig. 8.1: Scheme of a generic ADC converter.

Scale	Offset Binary	Two's Compl.	One's Compl.	Sign Mag.
+FS−1LSB	1111 ... 1111	0111 ... 1111	0111 ... 1111	1111 ...1111
+3/4 FS	1110 ... 0000	0110 ... 0000	0110 ... 0000	1110 ...0000
+1/2 FS	1100 ... 0000	0100 ... 0000	0100 ... 0000	1100 ...0000
+1/4 FS	1010 ... 0000	0010 ... 0000	0010 ... 0000	1010 ...0000
+0	1000 ... 0000	0000 ... 0000	0000 ... 0000	1000 ...0000
-0	- - - - ... - - - -	- - - - ... - - - -	1111 ... 1111	0000 ...0000
−1/4 FS	0110 ... 0000	1110 ... 0000	1101 ... 1111	0010 ...0000
−1/2 FS	0100 ... 0000	1100 ... 0000	1011 ... 1111	0100 ...0000
−3/4 FS	0010 ... 0000	1010 ... 0000	1001 ... 1111	0110 ...0000
−FS+1LSB	0000 ... 0001	1000 ... 0001	1000 ... 0000	0111 ...1111
-FS	0000 ... 0000	1000 ... 0000	- - - - ... - - - -	- - - - ...- - - -

Tab. 8.1: Most common digital codes.

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.

8.2. Static Characteristics

8.2.1 Quantization Error

In Fig. 8.2 is shown the characteristics of a generic 3bit ADC converter. As can be seen from the graph, the striking feature of analog to digital conversion is the fact that each output code corresponds to a range of values of the analog input, which, in fact, can vary continuously. The width of these intervals is $V_{ref}/2^n$ and takes the name of Least Significant Bit (LSB). Precisely, for this reason, the conversion will be affected by an error which cannot be eliminated, with maximum amplitude of ±½LSB, due to the discretization (quantization) of the signal through the steps of the ADC, rather than an ideal straight line.

This is definitely an error correlated with the input signal, as can be seen from Fig. 8.3. Even if the input were constant, the quantization error would also be constant, ϵ_q=cost. Nevertheless, it is convenient to treat quantization error analytically as if it were a white noise uncorrelated with the input signal. This approximation becomes more realistic at increasing resolution of the converter, since step widths get reduced. The two processes will have the same “effect” if they will possess the same power, i.e. the same effective value.

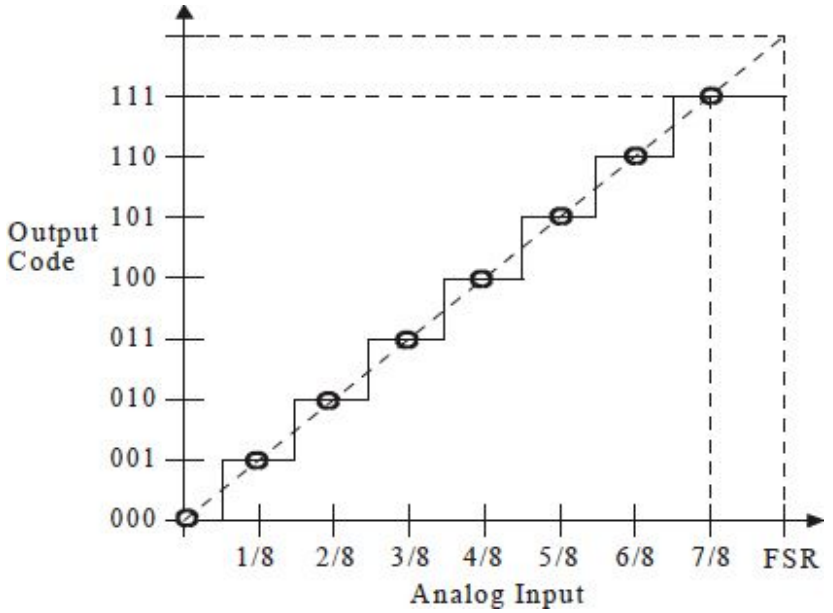


Fig. 8.2: Ideal characteristics of a three bit ADC.

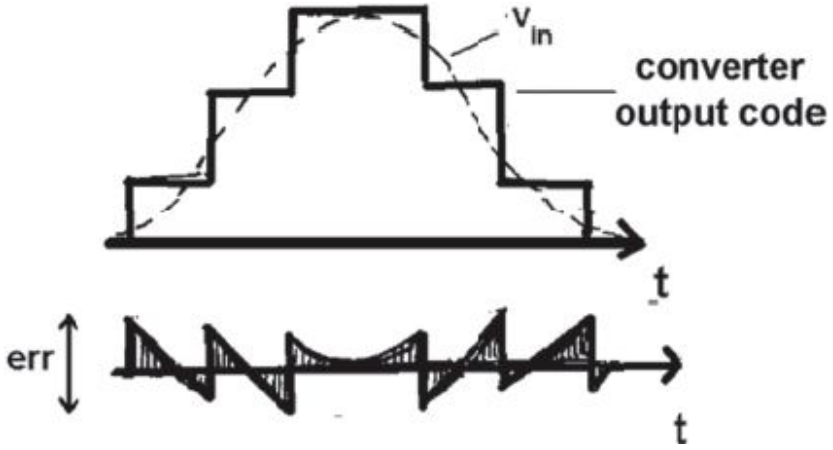


Fig. 8.3: Quantization error due to the difference between the ideal signal and that obtained after the ADC quantization. Note that, although it has a shape perfectly correlated with V_{in}, it could be also imagined as a random and correlated white noise as long as the two processes have the same effective value σ_q.

Assuming that the error is perfectly 'white' and uniformly distributed between -½LSB and +½LSB (i.e. with its mean value equal to zero), it is possible to calculate its mean square value, i.e. the power of such noise (error):

$$\sigma_q^2 = \frac{1}{T}\int_0^T \varepsilon_q^2 dt = \frac{LSB^2}{12} = \frac{FSR^2}{12 \cdot 2^{2 \cdot n}}$$

And σ_q=LSB/√12 represents the effective value of this **deterministic** error. As the number of bits increases, particularly in the case of variable (not too slowly) signals, the quantization error (shown in Fig. 8.3), though deterministic, becomes more and more difficult to understand, ending up appearing similar to an error correlated with the input and, therefore, assuming a **statistical**, random behavior (the typical "grass" effect of the white noise).

It is convenient to consider σ_q as if it were the effective value of a statistical white noise with Gaussian distribution so that it can be compared with that of other electronic noises superimposed on the signal or created by the electronics, being truly random. For example, a sine wave of 2.5V peak converted with a 3bit ADC (which provides an LSB=625mV) will have a 'step' representation that deviates from the real signal of the beginning as if a noise of 180mV$_{rms}$ were superimposed on the sine wave.

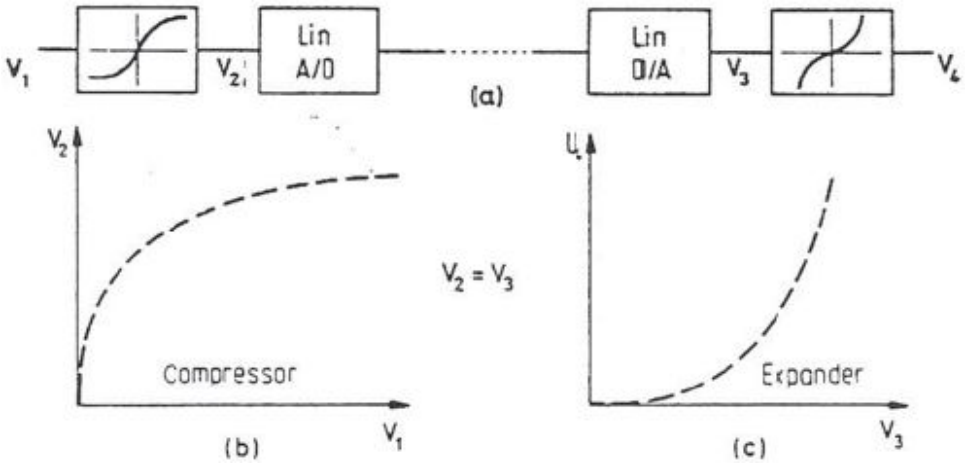


Fig. 8.4: Block scheme of a possible non-linear quantizator.

8.2.2 Non-Linear Quantizator

In some applications, a non-linear quantization is adopted. For example, in the case of voice signals, in instrumentation, and in control systems, it may become impossible for signals with large amplitude to reach the same resolution necessary for small signals. For this reason, one has to try to make

the quantization intervals become greater as the signal increases; in this way, we get a constant SNR, independent of the level converted. This is important, for example, in voice transmission with PCM (Pulsed Code Modulation) so that the reproduction fidelity is independent of the level of the input signal. In fact, the sensitivity characteristics of the hearing system are such that the increase in noise due to the lower resolution is “masked” by a proportional increase in signal intensity.

Fig. 8.4 shows a possible PCM transmission system in which an analog compression stage precedes a linear ADC and another follows the DAC, which is also linear. The combined effect of these two stages is just to make a non-linear quantization with an LSB whose amplitude grows in an exponential way. A quantization as such allows using the transmission channel more efficiently. It can be demonstrated that a code of only 8 bits with non-linear quantization leads to the same transmission quality of a 12bit linear code.

8.2.3 Offset and Gain Errors

There are some types of errors that can be adjusted without changing the external features of the device. To evaluate them, the real characteristics of the ADC are compared with the ideal steps, having the midpoints of each step laid out along the bisector of the quadrant. In doing so, the first step (corresponding to the output code with all bits to 0) should be only ½LSB wide while the last should be 1.5LSB wide. In fact, the centre of the step corresponding to the Full Scale Range FSR (i.e. when $V_{in}=V_{ref}$) does not match the code with all bits to 1, but that immediately following (which, however, is not actually available since it would require an additional bit at the output of the ADC). In other words, the code with all bits to 1 would correspond to an input not equal to V_{ref}, but to $V_{in}=V_{ref}$-1LSB. This fact is clearly visible in Fig. 8.2.

The o*ffset error* is the deviation of the width of the first step from the ideal value of ½LSB. The *gain error* is the difference between the width of the last step and the ideal value of 1.5LSB. Otherwise, it is said that the gain error encloses the effect of a slope not equal to 45° of the ADC characteristics. Often, both errors can be easily cancelled externally to the ADC by the careful calibration of the circuit, acting on V_{ref} or on the pin of the ADC for offset adjustment (if present).

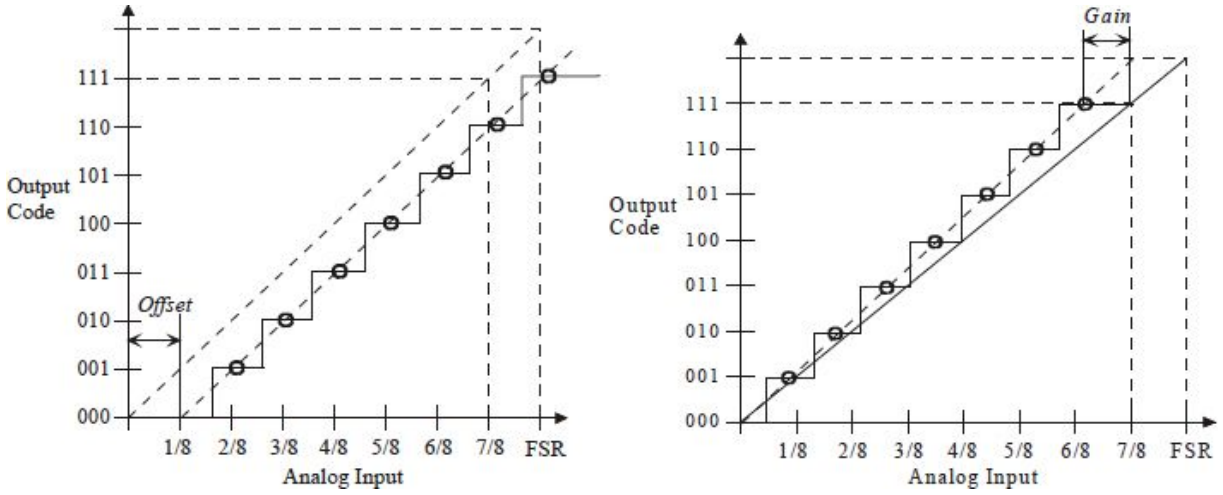


Fig. 8.5: Offset error (on the left) and gain error (on the right).

8.2.4 Non-Linearity Errors

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.

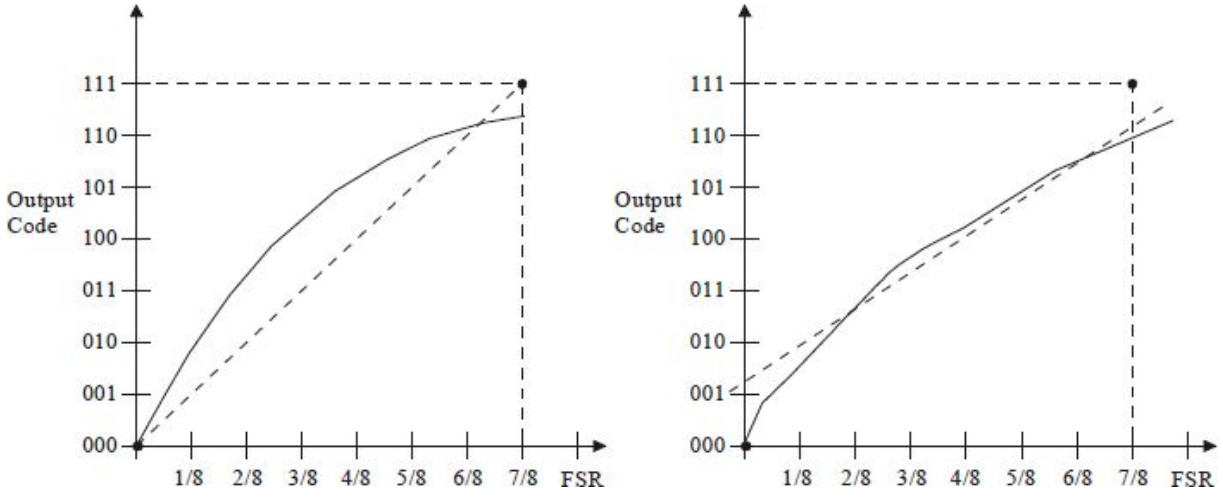


Fig. 8.6: Non-linearity error of the ADC evaluated in relation to the ‘best-fitting’ curve.

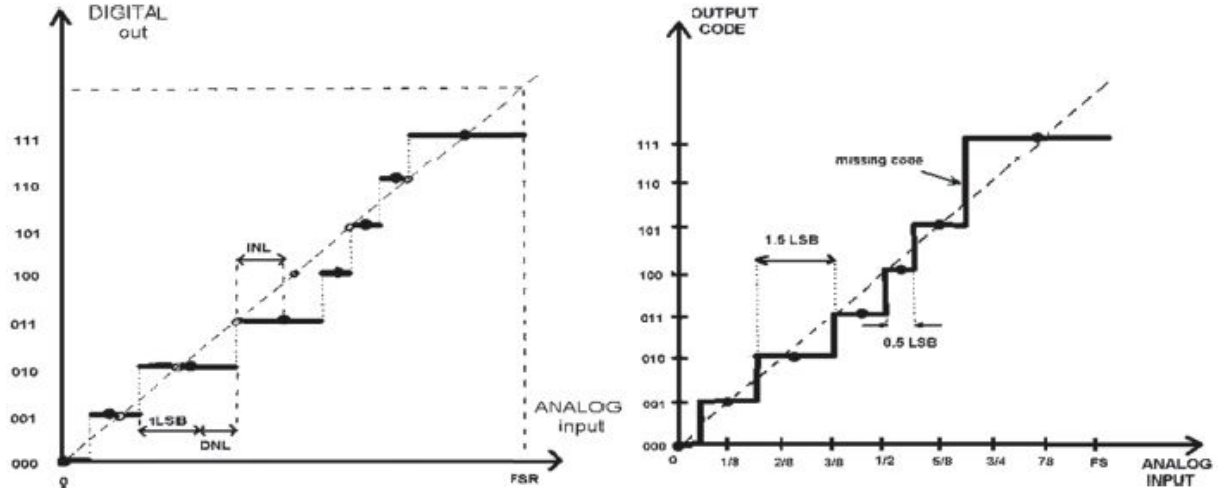


Fig. 8.7: Differential DNL and integral INL non-linearities (left) and esample of missing code 110 (right).

8.3. ADC CLASSIFICATION

ADCs can be classified according to the technique they exploit for the conversion of the analog input signal. A possible subdivision is the following:

- *direct conversion*, in which the output is obtained by direct comparison between the analog input signal and, for example, a series of “weighed” voltages; this can be done simultaneously for all bits or with more stages in sequence (pipeline);
- *indirect conversion*, where the analog signal is converted into an intermediate quantity (e.g. time or frequency), from which the digital code is derived.

Another possible subdivision is as follows:

- *conversion based on the coupling between components (component matching)*, in which the ultimate factor limiting the performance is just the “matching” between the values (resistance values, capacitance values, currents, …), as is understandable in techniques based on comparison;
- *conversion based on the count (counting algorithm)*, in which the precision required for the components is not very difficult to obtain, thanks to the fact that the conversion is performed indirectly.

Further classification brings to differentiate between:

- *single-shot conversion* in which, following a single command of Start of Conversion (SoC) and after a certain conversion time, the output reaches the final value, the End of Conversion (EoC) output is stated, and the converter is ready for a new conversion. The sample is converted within a single SoC cycle;
- *free-running conversion* in which, in order to correctly proceed in the conversion, the converter has to continue receiving SoC commands, thus converts samples in sequence. It is typical of pipeline architectures, in which the sample is converted in more SoC cycles.

Perhaps an even more useful classification is the following one:

- *instantaneous conversion*, when the instantaneous value of the analog input is converted (usually held by a S&H circuit), independent of its waveform; if noise is present on the input signal, its instantaneous value will be converted as well, and it would add “noise” to the conversion.
- *integration conversion*, when the analog input signal is first integrated for a well-defined integration time, then the final converted digital code is obtained; possible added noise will be integrated and hence eventually reduced in the final converted value. If disturbances (i.e. spurious tones at well-defined frequencies) are present on the input signal, by setting an integration time equal to a multiple of all those periods, the conversion can

be independent of those contributions since their integral would eventually vanish

These classifications are convenient because what distinguishes the two classes is not only the conversion philosophy, but also the kind of performance achievable (speed, resolution) and, consequently, the field of use. For example, the direct approach is usually faster than the indirect one. The approach based on counting can also be implemented with discrete components while the matching approach requires an integrated implementation. Free-running converters are faster than single-shot ones as long as we acquire samples in sequence, and the latency between SoC and the corresponding EoC is not a problem in the application. Finally, integration converters are preferable when we have a signal whose average value we are interested in converting within a certain time-frame, rather than its actual value; the typical case is when on the signal is superimposed a noise or a disturbance with a higher frequency.

It is now worth offering a quick overview of the various types of ADCs on the market to appreciate their different performance and select them according to the specific application.

8.4. Architecture of Various ADCs

8.4.1 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).

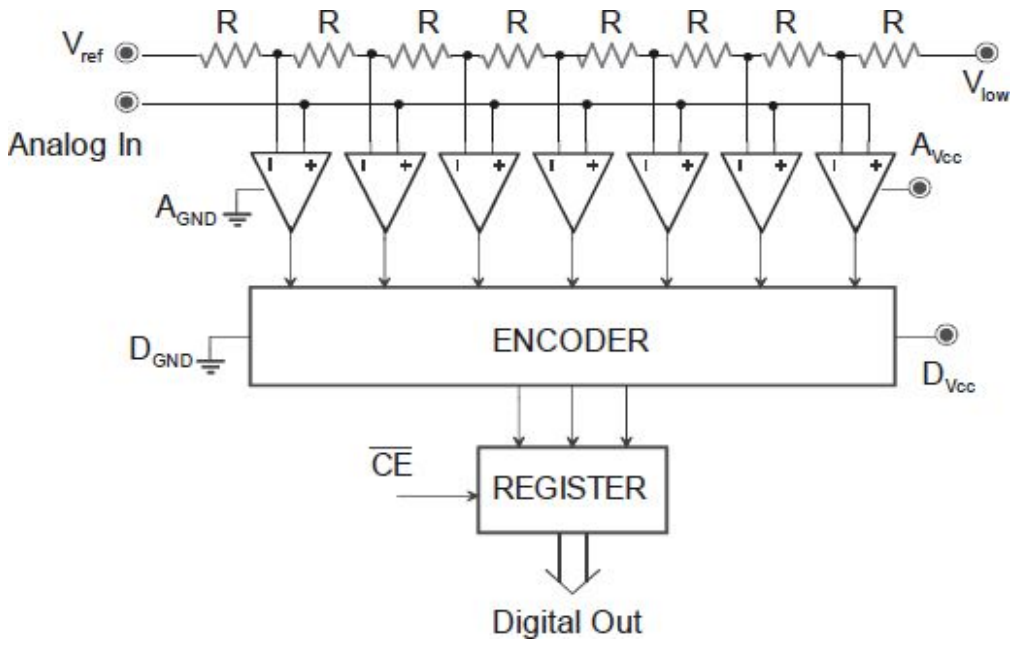


Fig. 8.8: Scheme of a flash converter.

8.4.2 Staircase Tracking 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}.

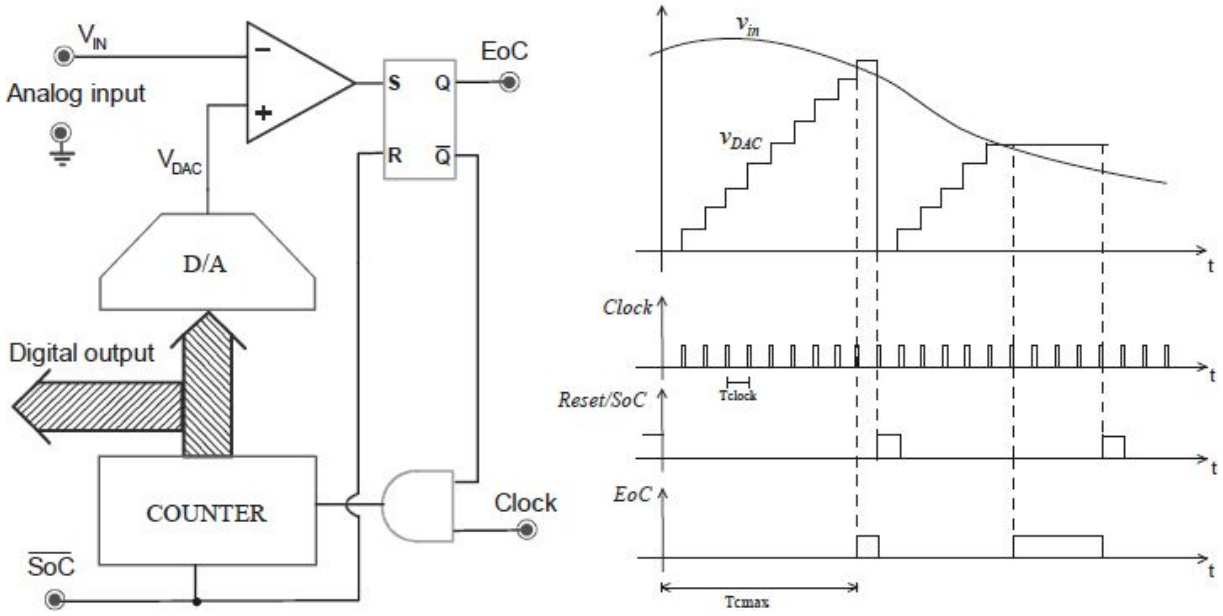


Fig. 8.9: Staircase tracking ADC (on the left) and timing of the signals (on the right).

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.

8.4.3 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:

$$f_{max} \leq \frac{f_{clock}}{2^n \pi}$$

For example, a 10bit converter with f_{clock}=10MHz could follow a sine wave at full dynamics with $f_{max}\leq$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.

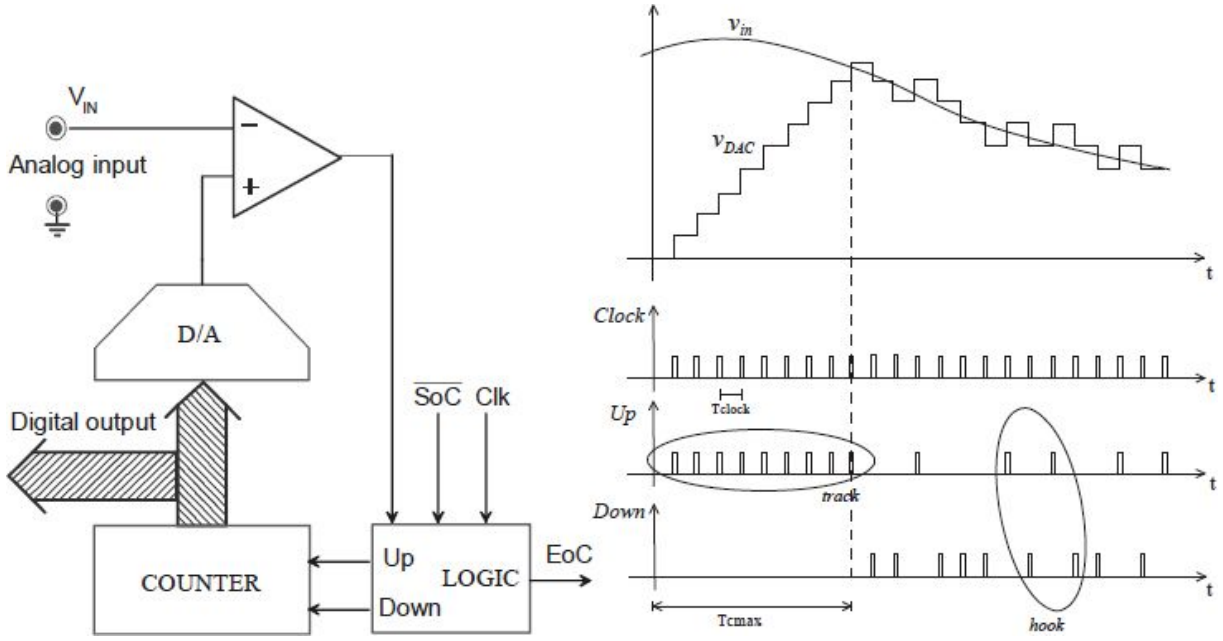


Fig. 8.10: Tracking ADC (on the left) and timing of the signals (on the right).

8.4.4 Single-Slope ADC, Single Ramp

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}.

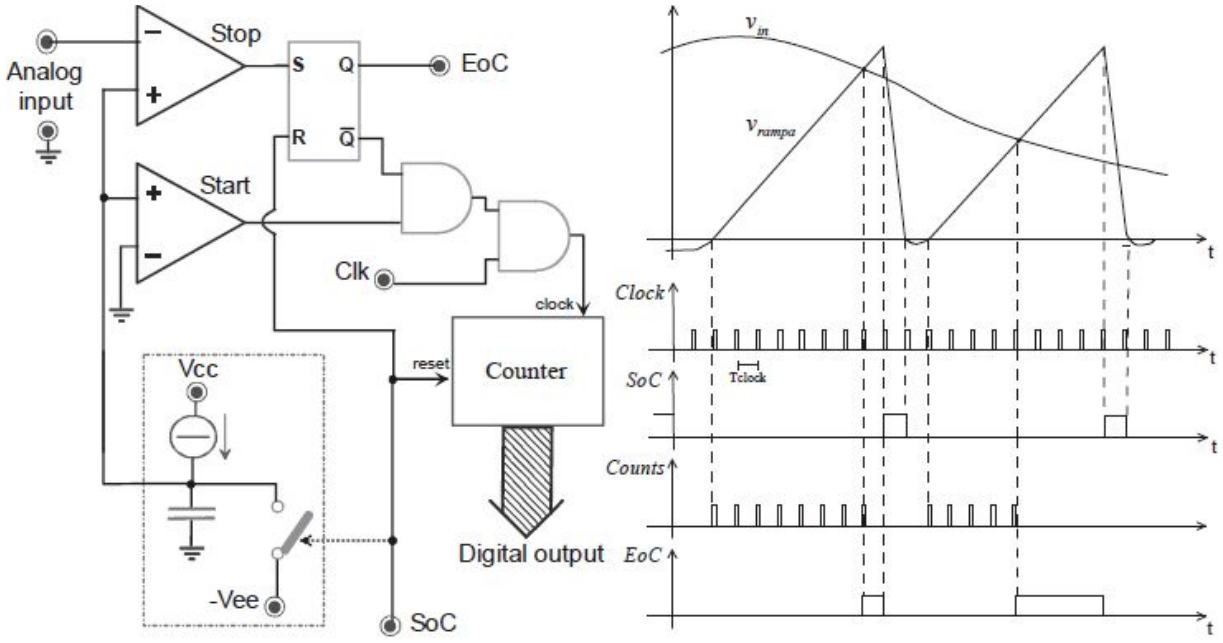


Fig. 8.11: Single slope ADC (on the left) and trends of the signals (on the right).

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.

8.4.5 Double Ramp 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

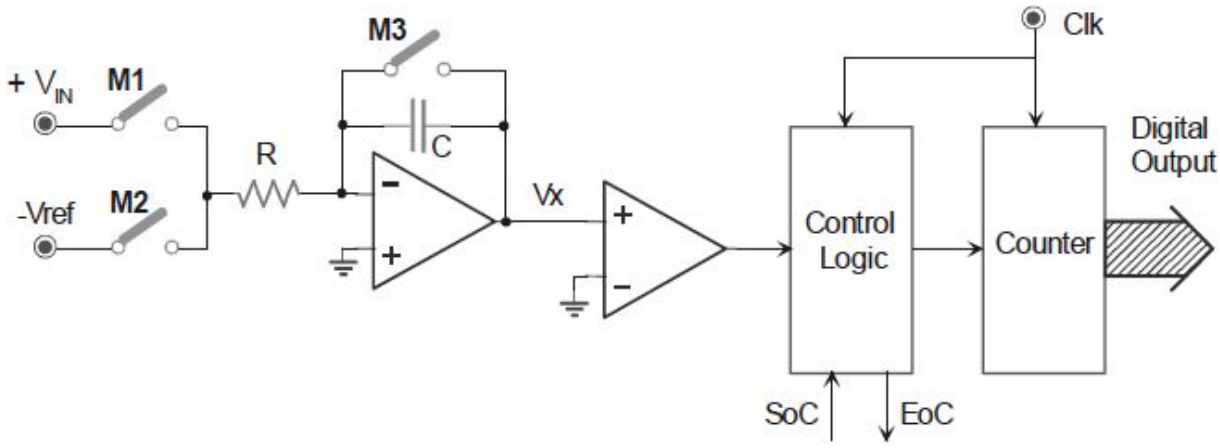


Fig. 8.12: Double ramp ADC.

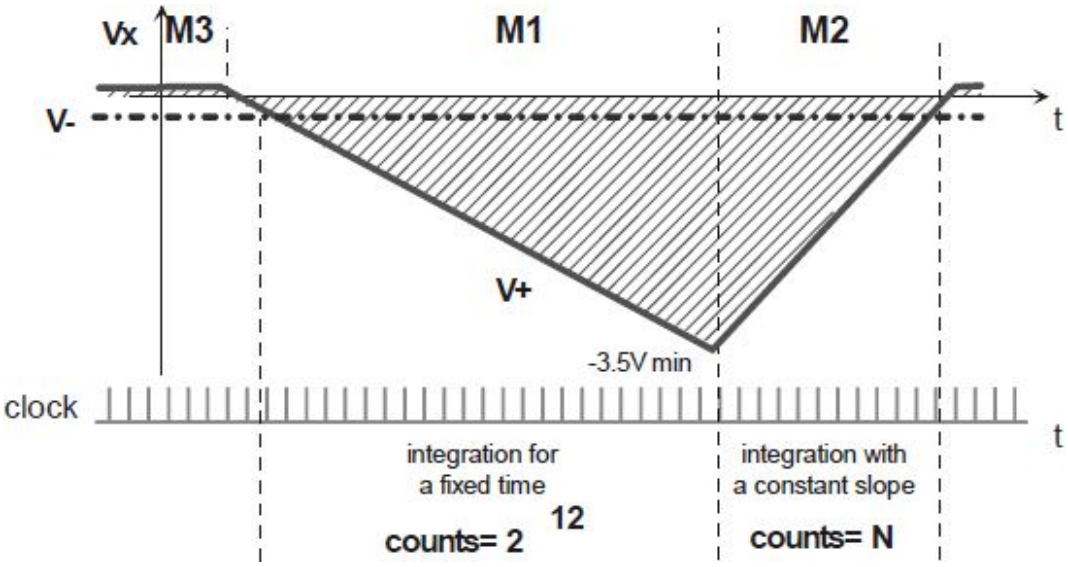


Fig. 8.13: Trend of the voltage across the integration capacitor.

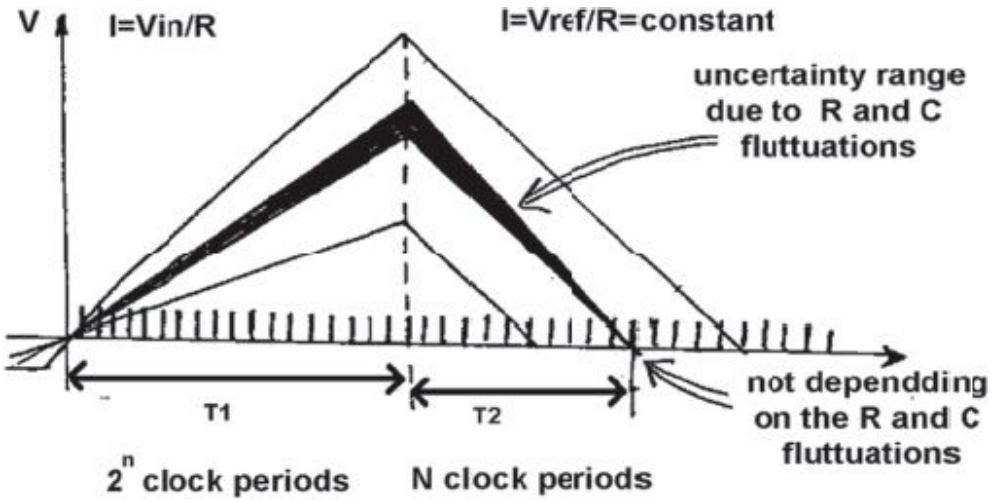


Fig. 8.14: Charge and discharge conversions.

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 $n+1^{th}$ 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.

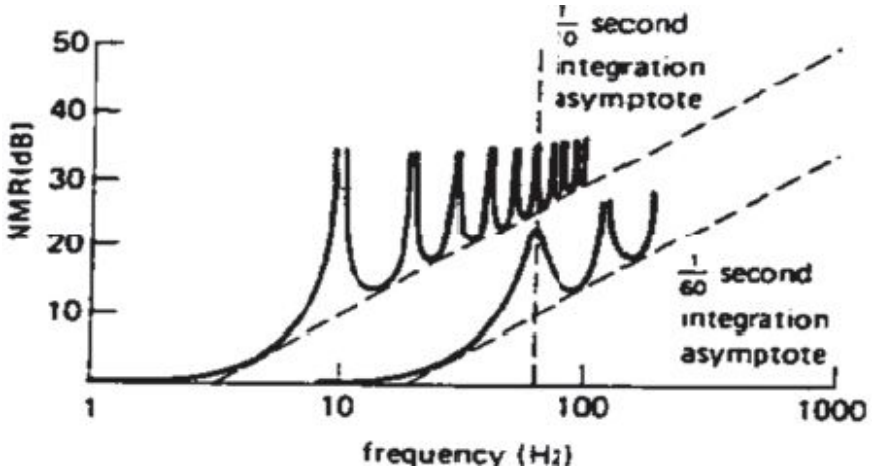


Fig. 8.15: Normal Mode Rejection ratio in a Dual-Slope ADC.

8.4.6 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}=1.1µs (instead of 100µs of a ramp ADC), and it becomes possible to handle signals with a maximum frequency of

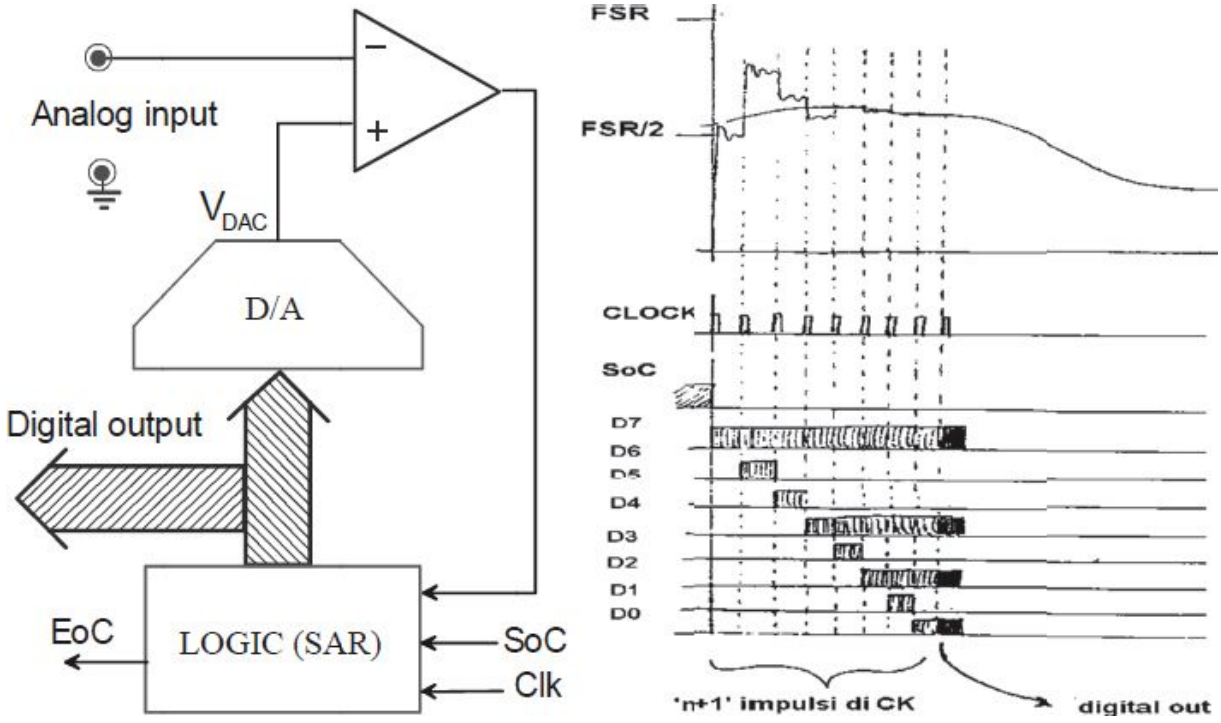


Fig. 8.16: Successive approximation ADC (on the left) and respective timings (on the right).

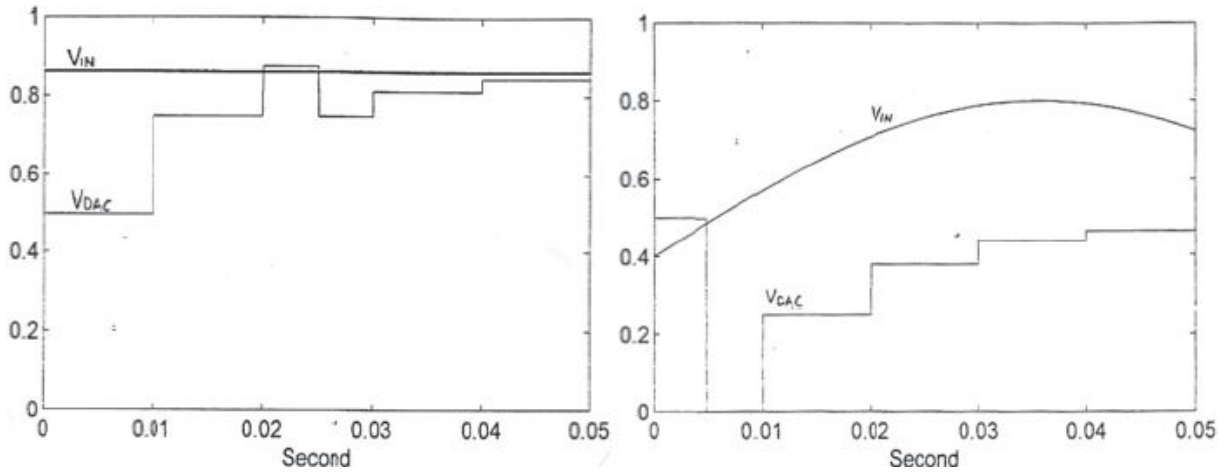


Fig. 8.17: Operation of an SAR in the presence of a constant (on the left) or a variable input (on the right).

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.:

$$f_{in,max} \leq \frac{f_{clock}}{2\pi \cdot 2^{n+1} \cdot (n+1)}$$

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.

8.5. Quality Factors

Some of the main quality factors are accuracy, resolution, and precision. Since DACs and ADCs are partly analog circuits, they suffer from inaccuracies due to mismatches between components, electronic noise, and thermal drifts, all of which can degrade their performance in terms of achievable resolution.

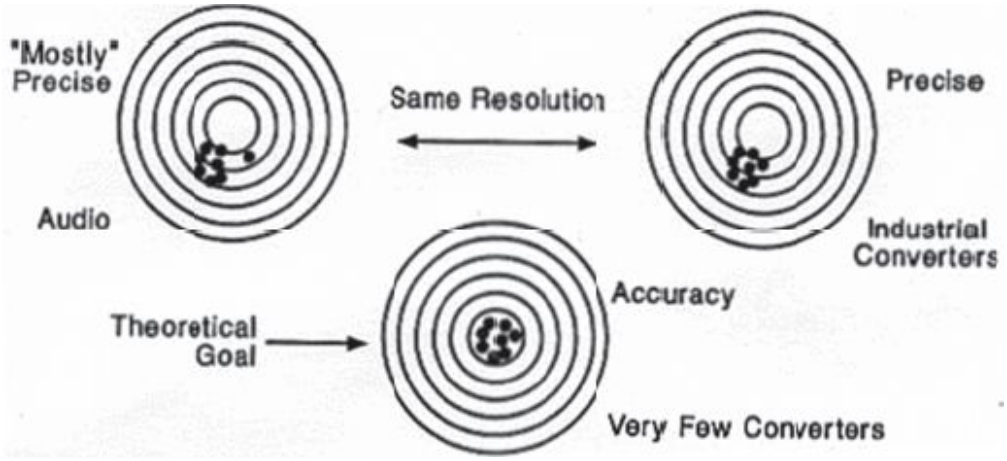


Fig. 8.18: Different meanings of accuracy, resolution, and precision.

One indicator of ADC performance is the absolute **accuracy**, defined as the maximum error between the analog output value and the theoretical value expected. As for the other figures of merit of the converter, also this is usually expressed as a fraction of the LSB. Ideally, it should not exceed ±0.5LSB, which is what would give the only quantization error.

In addition to accuracy, also the other quality factors define performance, as pictorially illustrated in Fig. 8.18. The **resolution** is the largest number of subdivision of the output dynamics, which the converter is able to distinguish (it is the digital fineness with which the analog quantity can be described). Denoting n as the number of input bits, the ADC will have 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 quality of the ADC to provide the same digital value (repeatable) with the same analog input applied. Note that, while accuracy implies precision, the opposite is not true.

Traditionally, there are two classes of converters: those that require a high accuracy (i.e. the ADC must represent the actual value of the input) and those that require precision within a certain range of values, namely the

“repeatability” of the output in response to a given input. As can be seen from this quick analysis, the resolution is an insufficient indicator to determine the performance of converters since converters may have applications in various fields. An example of a very specific market is that of converters for audio applications; in this field, it is important that the converters are, first of all, precise. In fact, since the analog signal must undergo many digital processing, the output will be acceptable only if all the conversions have been precise enough. In other areas, however, the request will be precision within a certain error along the whole transfer function. The ultimate desire is still to design a converter in which the absolute accuracy tends to the theoretical accuracy. In the audio field, for example, where the resolution required is typically 16 bits, “perfection” is not reachable. Moreover, the continuous need to reduce costs has led to increasingly burdensome choices for the ADC qualities: smaller packages that do not allow offset and gain adjustments, faster testing that measures only the main quantities and only at certain frequencies and temperatures (without warranties or remedies for possible drifts). That is why these audio ADCs often do not have the specifications required for more demanding applications, such as those in industrial area and those for precision instrumentation.

8.6. 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 …).

ADC	Type	Resolution (bits)	Conversion time (ns)	Linearity error (LSB)	Power dissipation (mW)
TL 507	SA	7	1000	± 1/2	25
ADC 302	F	8	20	± 1/2	550
ZN 439	SA	8	5000	± 1/2	150
ZN 433	T	10	1000	± 1/2	500

Tab. 8.2: Some features of commercial ADCs: Successive-Approximation (SA), Flash (F), and Tracking (T).

Parallel non-pipelined ADC (single-shot)

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: $SpS<1/T_C$.

Parallel pipelined ADC (free running)

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.

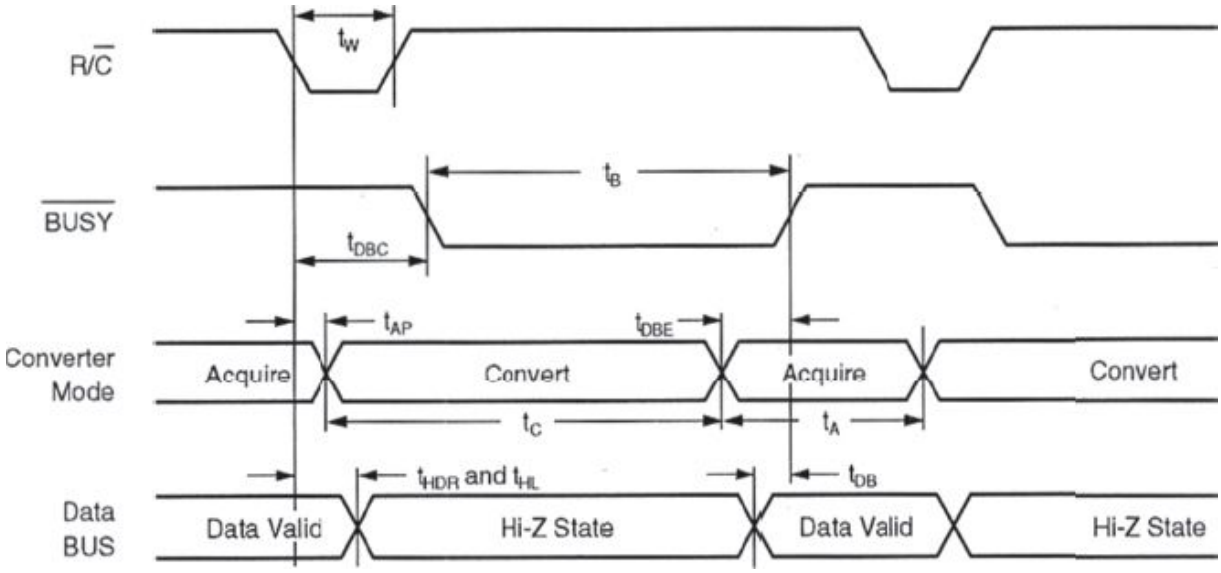


Fig. 8.19: Timing diagram of a parallel single-shot ADC.

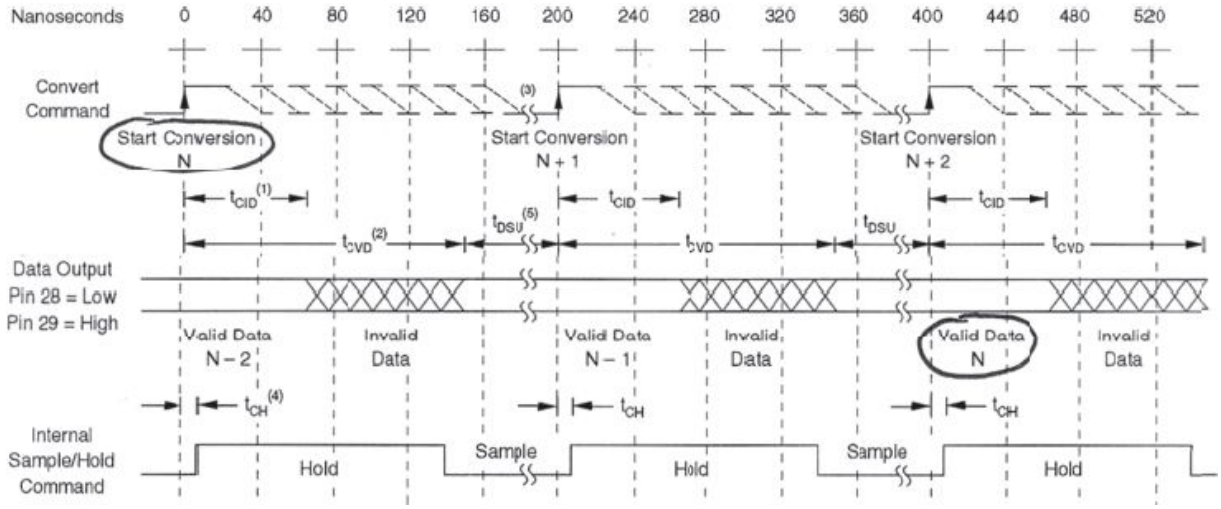


Fig. 8.20: Timing diagram of a parallel pipelined ADC.

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.

Serial pipelined ADC

Serial ADCs (as opposed to the parallel ones) provide at the output the bits of the result in a serial way, thereby allowing a significant reduction in the number of pins of the integrated circuit and, therefore, in its cost and size. Such architecture is suitable for a pipelined approach to make up for the time that is necessary for the parallel to serial conversion. In Fig. 8.21, the control of the operations is handled by a clock external to the ADC and by the pins of ChipSelect and Ready/Convert. The sync signal informs the external world that the ADC is ready to serially provide at the output the bits of the result of the conversion, once released the busy line. Serial ADCs have SpS>1/T_C if pipelined while they have SpS<1/T_C if non-pipelined.

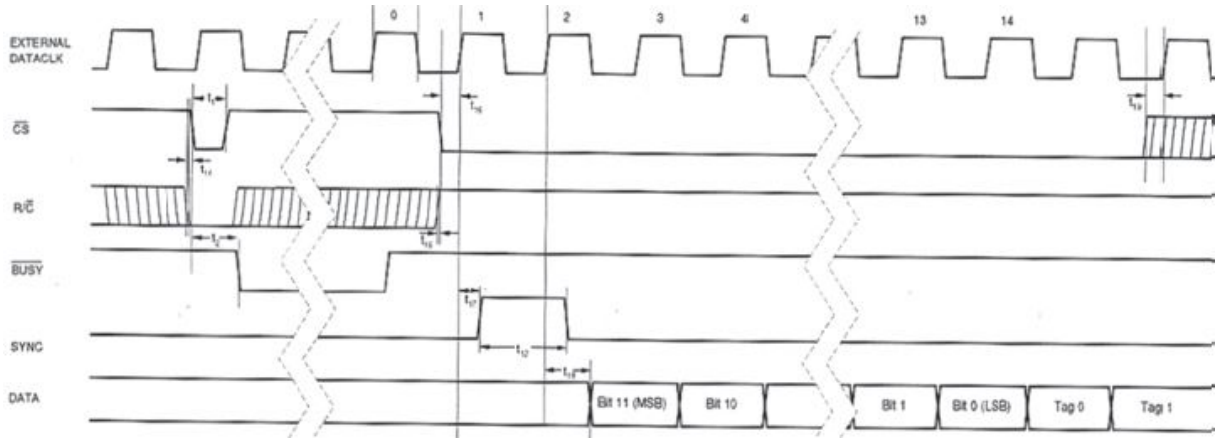


Fig. 8.21: Timing diagram of a serial pipelined ADC.

n Bits	States 2^n	LSB Weight $1/2^n$	LSB Weight in ppm	LSB Weight in % of Full Scale Range	Bit Weight or LSB for 10V FSR	Theoretical Signal/ Quantitization Noise Ratio in dB	Dynamic Range in dB
0	1	1.000	1,000,000	100	10.000 V	1.76	0.00
1	2	0.500	500,000	50	5.000 V	7.78	6.02
2	4	0.250	250,000	25	2.500 V	13.80	12.04
3	8	0.125	125,000	12.5	1.250 V	19.82	18.06
4	16	0.062500	62,500	6.25	625.000 mV	25.84	24.08
5	32	0.031250	31,250	3.125	312.500 mV	31.86	30.10
6	64	0.015625	15,625	1.5625	156.250 mV	37.88	36.12
7	128	0.007812500	7,812.500	0.781250	78.125 mV	43.90	42.14
8	256	0.003906250	3,906.250	0.390625	39.063 mV	49.92	48.16
9	512	0.001953125	1,953.125	0.195312	19.531 mV	55.95	54.19
10	1,024	0.000976562500	976.562	0.097656	9.766 mV	61.97	60.21
11	2,048	0.000488281250	488.281	0.048828	4.883 mV	67.99	66.23
12	4,096	0.000244140625	244.141	0.024414	2.441 mV	74.01	72.25
13	8,192	0.000122070313	122.070	0.012207	1.221 mV	80.03	78.27
14	16,384	0.000061035156	61.035	0.006104	610.352 µV	86.05	84.29
15	32,768	0.000030517578	30.518	0.003052	305.176 µV	92.07	90.31
16	65,536	0.000015258789	15.259	0.001526	152.588 µV	98.09	96.33
17	131,072	0.000007629395	7.629	0.000763	76.294 µV	104.11	102.35
18	262,144	0.000003814697	3.815	0.000381	38.147 µV	110.13	108.37
19	524,288	0.000001907349	1.907	0.000191	19.073 µV	116.15	114.39
20	1,048,576	0.000000953674	0.954	0.000095	9.537 µV	122.17	120.41

Tab. 8.3: Comparison between bits and resolution.

8.7. Dynamic Characteristics

With the term *dynamic characteristics of a converter*, we often refer to that converter's performance in terms of dynamics, harmonic distortion, noise, and distortions introduced, hence parameters different from the timings described above. Tab. 8.3 links the parameters resolution and LSB as the number of bits changes.

8.7.1 Dynamics

It is defined as the ratio between the maximum value supplied by an ADC and the minimum quantum (LSB). It is therefore trivial to find that it is equal (in dB) to:

$$D = 20 \cdot \log \frac{FSR}{LSB} = 20 \cdot \log 2^n = 6.02 \cdot n$$

i.e. an 8bit ADC with FSR=5V can give a minimum output equal to 0000'0001, corresponding to 19mV and a maximum one 1111'1111, corresponding to 5V, with a dynamics of just 6.02·8=48dB, just equal to 5V/19mV.

8.7.2 Signal to Noise Ratio

A feature of the utmost importance to define the performance of an ADC is the Signal to Noise Ratio (SNR). As seen above, the maximum theoretical limit of this parameter is obtained by applying to the input of the ADC a sine wave with the maximum amplitude and assuming that the ADC itself does not introduce any noise if not only the quantization noise (which depends on its resolution in terms of bits). Describing the quantization error as a white noise uniformly distributed between -½LSB and +½LSB, we can derive the theoretical SNR:

$$\mathrm{SNR}\Big|_{\mathrm{ideal}} = \frac{\text{maximum signal power}}{\text{only quantization power}}\Bigg|_{\mathrm{dB}} = 10\cdot\log\frac{\left(\mathrm{FSR}/2\sqrt{2}\right)^2}{\left(\mathrm{LSB}^2/12\right)} = 6.02\cdot n + 1.76$$

For example, an 8bit converter should have an SNR_{ideal}≈50dB.

Actually, within the ADC, other electronic noises superimpose on the analog input signal, and the result of the conversion will be affected by a ‘numerical’ noise far higher than the only quantization error. For example, as the frequency of the input signal increases, the internal circuitry commits more errors because of its limited bandwidth, settling times and so on, both of which are likely to markedly decrease the measured SNR, as shown in Fig. 8.22. This degradation is even more pronounced as the ADC is operated at high frequencies (high f_S and f_{in}) and has a high number of bits because of the difficulty in rejecting the broadband noise in the S&H at the input and in reducing the internal 1/f noise and all the disturbances internal to the ADC.

8.7.3 Effective Bits and ENOB

To quantify the total noise of a real ADC, the SNR_{real} actually measured at the output of the component itself could be provided. In place of this figure in dB, it is often more useful to speak in terms of effective bits or ENOB (Effective Number Of Bits) that the converter has, i.e. the number of bits that should have an ideal ADC (only affected by the quantization error) that converts a sine wave with a ‘quality similar’ to that obtained by the real ADC (affected also by real electronic noises). The ENOB is given by:

$$\mathrm{ENOB} = \frac{\mathrm{SNR}_{\mathrm{real}} - 1.76\mathrm{dB}}{6.02\mathrm{dB}}$$

So, a 10bit ADC (having a theoretical SNR_{max}=6.02·10+1.76=62dB) that converts an analog sine wave with the maximum amplitude, providing at the output a numerical sine wave with an SNR_{real} of only 56dB, would be similar to another ideal ADC with only (56-1.76)/6.02=9bit. That is, although the ADC has resolution of 10 bits (1024 possible codes at the output), it actually offers the same accuracy of a 9bit ADC (with only 512 codes): if the 10bit ADC provides the conversion 10'0000'0001=2.5049V, in reality, it would only mean that the input is about 2.5049V, but with a possible error of ±1LSB=±4.88mV (and not just of ±½LSB as it should be in an ideal 10bit ADC); this means that the correct code could be between 10'0000'0000 and 10'0000'0010.

The fact that the SNR_{real} is less than SNR_{ideal} may depend on two facts. The output of an ADC can lose resolution because the input sine wave does not exploit the whole dynamics allowed. For example, using a 10bit ADC with FSR=5V, to convert a sine wave with amplitude of $1.25V_{peak\text{-}to\text{-}peak}$, the digital code would exploit only 1.25V/5V=1/4 of the available levels, i.e. 64 levels instead of all 256. One could speak, in this case, of a theoretical SNR (i.e. due to the actual amplitude, but still considering the ADC ideal, in other words, noise-free):

$$SNR\big|_{theoretical} = \frac{\text{power of real signal}}{\text{only quantization power}}\bigg|_{dB} = 10\cdot\log\frac{\left(Vin_{peak}/2\sqrt{2}\right)^2}{\left(LSB^2/12\right)} = SNR\big|_{ideal} - 20\cdot\log\frac{Vin_{peak}}{FSR}$$

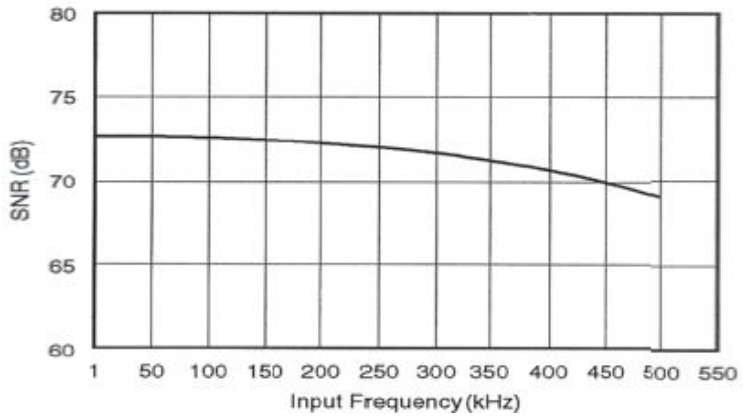


Fig. 8.22: Trend of the SNR as f_{in} increases.

In particular, it is clear that if the input signal of the ADC has a range of amplitude equal to half the entire FSR, the effective number of bits will be

reduced by one. This is why, in presence of an input signal with peak-peak amplitude much lower than the FSR of the ADC, it is convenient to properly pre-amplify the input signal before proceeding with the conversion. The other cause that reduces the SNR_{real} is the actual presence of noise in the real ADC, in respect of the ideal one.

8.7.4 Spectral Performance

The spectral quality of an ADC is evaluated by sending a very pure sine wave at the input and examining the spectrum at the output of the ADC. Since the latter is represented by series of numerical data, to be able to obtain its spectrum, we process them through the Fast Fourier Transform (FFT). This algorithm requires N numerical samples in the time domain (the data at the output of the ADC) to provide other N samples in the frequency domain, representing the spectral composition of the signal analyzed. The algorithm assumes that the sequence of N samples is periodic. Since the time samples are sampled with a frequency $f_S=1/T_S$, the frequency spectrum will repeat identically at multiples of f_S. Then, the N samples of the spectrum are not more than the values of the N frequency histograms, between f=0 and $f=f_S$, which are symmetric in the neighborhood of $f_S/2$. The width of each histogram is called '*bin-width*' and is equal to f_S/N. To have a well-defined frequency spectrum, it is necessary to have a very narrow bin-width and, therefore, acquire many samples N to be processed by the FFT algorithm.

Consider now an ideal ADC. If you gave at the input a pure analog sine wave with a frequency f_{in}, the samples at the output of the ADC would represent the same sine wave, but in the numerical form (we always have to meet the Nyquist sampling criterion, i.e. we choose to sample with $f_S>2f_{in}$). Because of the quantization, the sine wave is not perfectly identical to the input one, but it would have superimposed the quantization error. As said several times, if the ADC has a high number of bits (>6 about), this noise can be considered as white and distributed over the whole output spectrum, i.e. $0 \div f_S$.

Taking N samples of the numerical sine wave and calculating the FFT, we will find the N histograms in frequency. There will be one placed at f_{in}, with amplitude equal to the power of the output sine wave while all the others will represent the quantization noise (error) uniformly distributed over them. If

there were only one histogram, it would have amplitude equal to the whole power of the quantization noise, i.e.: $\sigma_q^2 = \frac{LSB^2}{12}$. For example, a 14bit ADC, with FSR=5V, LSB=305μV, the quantization noise power would be equal to $(88\mu V_{rms})^2$ while that of the fundamental would be $(2.5Vp/\sqrt{2})^2=(1.77V_{rms})^2$. From the definition of the theoretical signal to noise ratio, the amplitude of this single noise histogram would be lower than that of the sine wave exactly by SNR dB. Indeed, $1.77V_{rms}/88\mu V_{rms}=20091=86dB$, just equal to the definition SNR=6.02·n+1.76=86dB. It would be convenient to normalize all the amplitude values of the FFT and to match the fundamental with the power of $0dB_C$, where the subscript C indicates that it is referred to the power of $[(FSR/2)/\sqrt{2}]^2$.

Actually, since between $0 \div f_S$, there is not one, but N histograms in frequency, each will have amplitude equal to σ^2_q/N. Thus, by normalizing all with respect to the fundamental of $0dB_C$, the noise histograms would lie on an average level N times lower than the SNR. This level is called the '*noise-floor*', and, in dB, it should be equal to:

Noisefloor= -(6.02·n+1.76+10·logN)=-(SNR+10·logN)

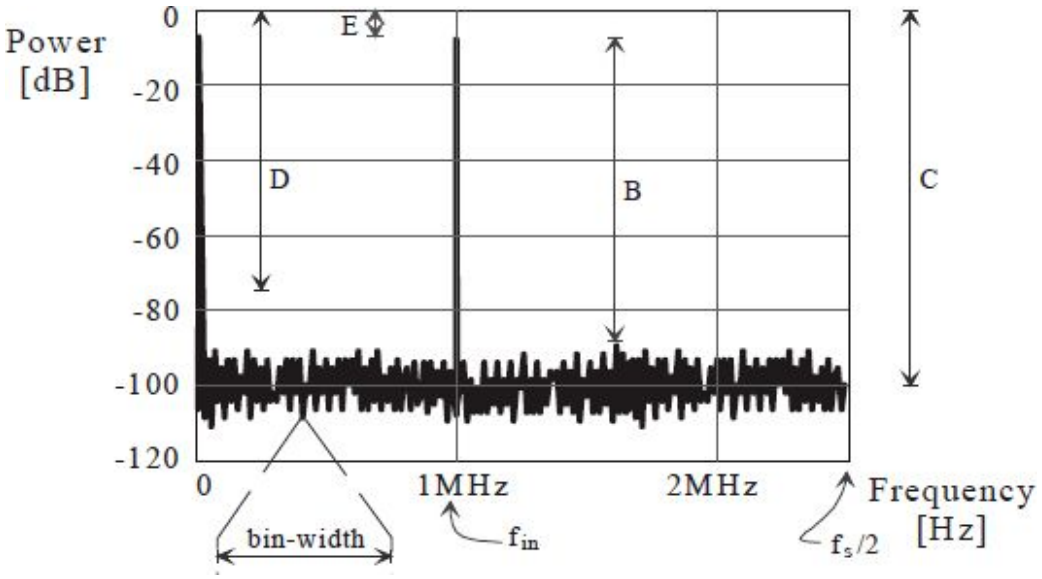


Fig. 8.23: Spectral power distribution at the output of a 14 bit ADC.

Indeed, since it is normal practice to plot only half the samples, i.e. only up to $f_S/2$, thanks to the symmetry of the samples between $0 \div f_S/2$ with those

between $f_S/2 \div f_S$, it is convenient that each histogram "accommodates" within itself the power of its symmetrical in order not to forget it in the calculations. For this reason, the noise-floor will be distributed over N/2 histograms and not over N; therefore, it will have a height twice that written above, i.e.:

$$\textbf{Noisefloor} = -(6.02 \cdot n + 1.76 + 10 \cdot \log N/2) = -(SNR + 10 \cdot \log N/2)$$

Note that, to have a lower Noisefloor, it is necessary to use a greater number of samples in order to calculate a more accurate FFT. Note also that it makes no sense to compare two spectra with different nose-floors if we first do not make sure that both had been calculated with the same number N of samples; in fact, a lower noise-floor could not indicate a less noisy ADC, but that the FFT has been calculated with a greater number of samples and, therefore, with a lower bin-width.

Usually, to avoid the risk of making the ADC saturate, at the input of the component, a sine wave not with the maximum amplitude, but with an amplitude equal to only 90%·FSR is sent; this margin of about 0.5dB is called the 'headroom' and avoids that, following undesired offsets, there are clipping phenomena of the digital sine wave at the output. This margin is shown with E in Fig. 8.23, which shows an example of output spectrum of a 14bit ADC614 with a sampling frequency f_S=5.12MHz (one sample each T_S=195ns) and input with f_{in}=2.35MHz and normalized amplitude of -0.5dB$_C$ (head-room); the plot has been obtained with N=8192 samples, collected for a total time of $T_{acq}=8192 \cdot T_S$=16μs.

8.7.5 SFDR and SiNAD

In Fig. 8.23, the 0.5dB is the headroom (E), and (A) is the amplitude of the fundamental equal to about 90% of the FSR. The section (B) is defined as the *Spurious-Free Dynamic Range* (SFDR) and represents the worst ratio between the fundamental and the highest unwanted disturbance or noise. The noisy background of graph (C) is just the noise-floor.

There is a further confirmation that we cannot see the theoretical SNR_{max} from the graph, but only the noise-floor, located in this case at –100dB$_C$. However, knowing that the number of histograms in the plot (between $0 \div f_S$) is just N, it is possible to calculate the total noise, N times greater than –100dB$_C$, i.e. equal to –100dB+10logN=-100dB$_C$+39dB=-61dB$_C$. This value (D) is

much higher than the theoretical $-86dB_C$ (equal to the SNR_{max}); then, in the ADC, there are other noises, much greater than the only quantization error.

In addition to the real SNR, which is measured, from the spectral diagram, it is possible to find also the *Signal to (Noise And Distortion) ratio* (SiNAD)

$$SiNAD = \left.\frac{\text{power of useful signal}}{\text{total power of real noise and harmonics}}\right|_{dB} = 10 \cdot \log \frac{\left(Vin_{peak}/\sqrt{2}\right)^2}{\sum \text{powers of all bin}}$$

As the SNR, also this cannot be represented on the graph.

8.7.6 THD and IMD

To quantify the harmonic distortion introduced by the non-linearities of the ADC, it is useful to relate the power of all fundamental harmonics (without considering other disturbances or the white noise) to the power of the fundamental itself. This ratio is entitled the *Total Harmonic Distortion* (THD), defined as:

$$THD = \left.\frac{\text{power of all harmonics}}{\text{power of the useful signal}}\right|_{dB} = 10 \cdot \log \frac{\sum_{k=2}^{9} P_{f=k \cdot fin}}{P_{f=fin}}$$

To further highlight the effects of non-linearities of the ADC, the component is stimulated not with one, but with two sine waves with slightly different frequencies. In this way, all the intermodulation products will arise, which are at all the frequencies that are a linear combination of the two at the input, as shown in Fig. 8.24. From these graphs, it is possible to find the *InterModulation Distortion* (IMD), equal to the ratio:

$$IMD = \left.\frac{\text{power of all intermodulation products}}{\text{power of the 2 useful signals at 2 different frequencies}}\right|_{dB}.$$

8.8. Requirements for Driving the ADC

In order to properly exploit the accuracy of an ADC, it is important to correctly provide it with the analog data to convert, avoiding introducing

offsets, drifts, and voltage drops along the path from the source to the ADC. Consider in this respect the circuit in Fig. 8.25. First, we must limit the output impedance of the source because it is the cause of undesirable effects that have an impact on the accuracy achieved by the converter.

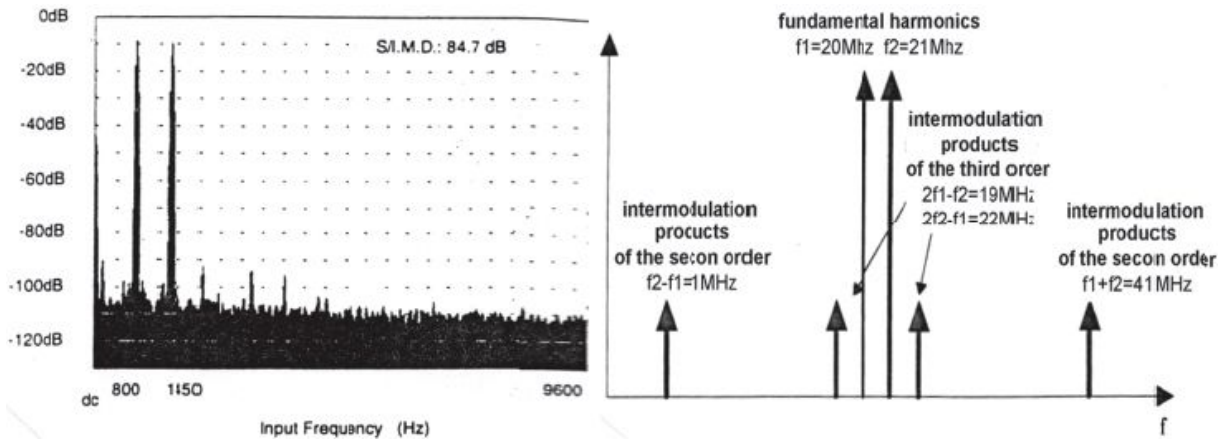


Fig. 8.24: Intermodulation products at the output of a non-linear circuit and intermodulation distortion at f_1=800Hz and f_2=1450Hz with f_s=19200Hz.

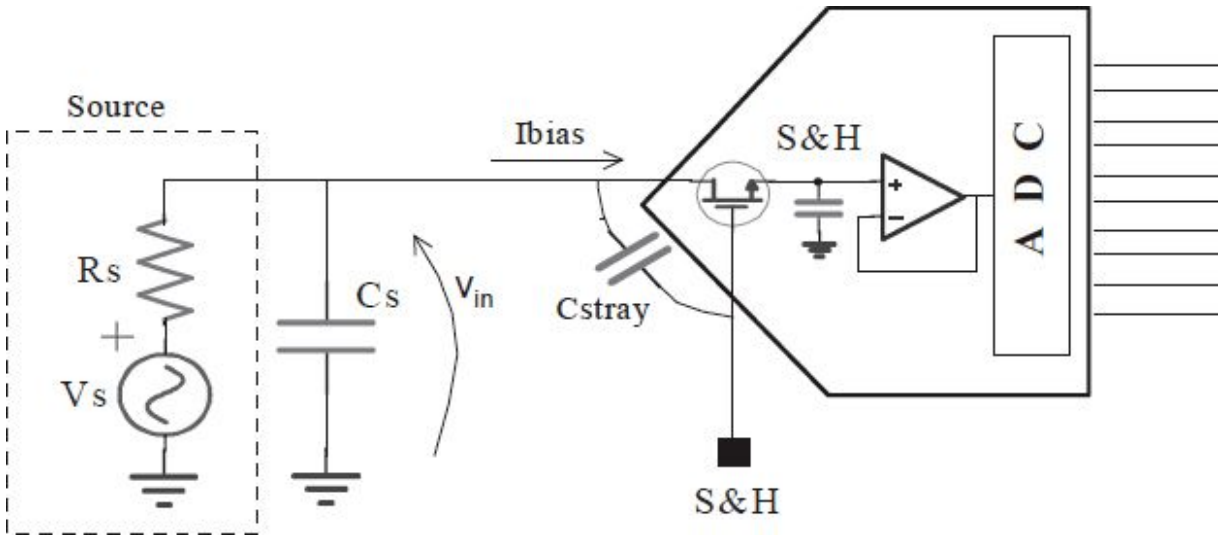


Fig. 8.25: Effect of the output impedance of the source.

There can be a static voltage drop across the output resistance of the source due to the bias current of the ADC. To keep the error <½LSB, we impose:

$$\Delta v_{in} = R_s \cdot I_{bias} \leq \frac{1}{2} LSB \Rightarrow R_s \leq \frac{LSB}{2 \cdot I_{bias}}$$

Therefore, in a 14bit ADC with bias currents of 50nA and FSR=5V, we must limit the output impedance of the source below Rs<3kΩ so as not to lose accuracy. Since it is a stringent constraint, it is often necessary to change the source or to introduce a buffer before the ADC. An additional problem may arise from the fact that R_s can vary as f_{in} changes: this would cause a voltage drop as a function of the frequency and, ultimately, a non-linearity in the acquisition.

If there is an S&H at the input of the ADC, we will have all the problems of *charge injection* and *feed-through* typical of S&Hs (see the corresponding chapter). Particularly if the source is capacitive (C_s), the injection of charge through C_{stray} of the MOS switch (Fig. 8.25) will determine an alteration in the signal, which is equal to:

$$\Delta v_{in} \cong \Delta V_{S\&H} \frac{C_{Stray}}{C_{Stray} + C_s}$$

For example, if the swing of the gate control voltage is 10V, with C_{Stray}=0.1pF and C_S=10pF, there will be a change in the input of the ADC of no less than 100mV, corresponding to 1LSB in a 5.6bit device. Since the time constant of discharge is equal to $\tau=R_sC_s$=3kΩ·10pF=30μs, this effect is felt only when working at frequencies higher than 8MHz, as can be seen from Fig. 8.26.

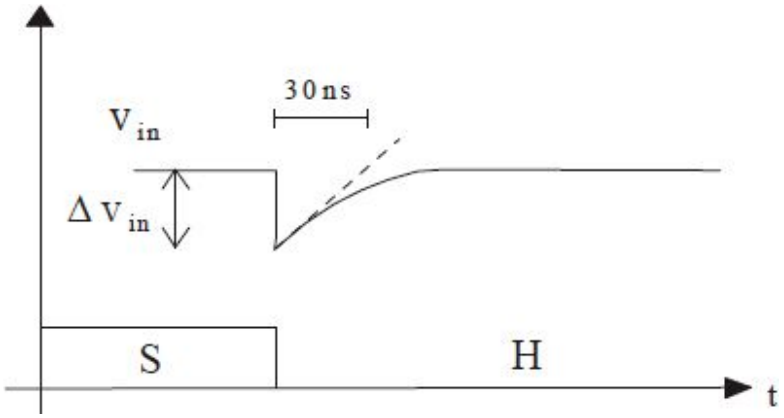


Fig. 8.26: Signal feed-through due to charge injection.

In case of v_s<<FSR, it would not exploit the full resolution of the ADC, as already discussed. This is why it would be convenient to amplify v_s in order that the maximum and minimum values exploit the full dynamic range. This operation is called ‘signal conditioning’, and it is done on the front-end, close to the sensor. Fig. 8.27 shows a simple example of a stage that translates the signal in the centre of the dynamics allowed (FSR/2) and amplifies it by a factor of G≈15kΩ/10kΩ=1.5; the coupling is AC, not suitable for continuous or very low frequency signals.

In order not to degrade the accuracy of the conversion, the OpAmp has to be chosen by taking care of the most important electrical characteristics. It should, for example, be able to supply the current peaks required to drive the converter (from 250μA to 20mA) and introduce a noise contribution lower than that which is theoretical of the ADC.

Whenever we have an AC coupling, such as that in Fig. 8.27 due to R_1 and C, and every time we risk having a non-flat frequency response of the OpAmp, it is fundamental to verify the transfer characteristics of the network. In fact, we can run the risk of underestimating capacitive attenuation and introduce a voltage drop before getting to the ADC, degrading again the accuracy of the conversion. For example, consider the high frequency response of the circuit. It is necessary that the loop gain be high enough to avoid the risk of being subjected to attenuation on high frequency signals. Although the high frequency pole is equal to GBWP/G (GainBandwidthProduct/Gain), as evident from Fig. 8.28, on the left, the maximum frequency of the input has to be much lower.

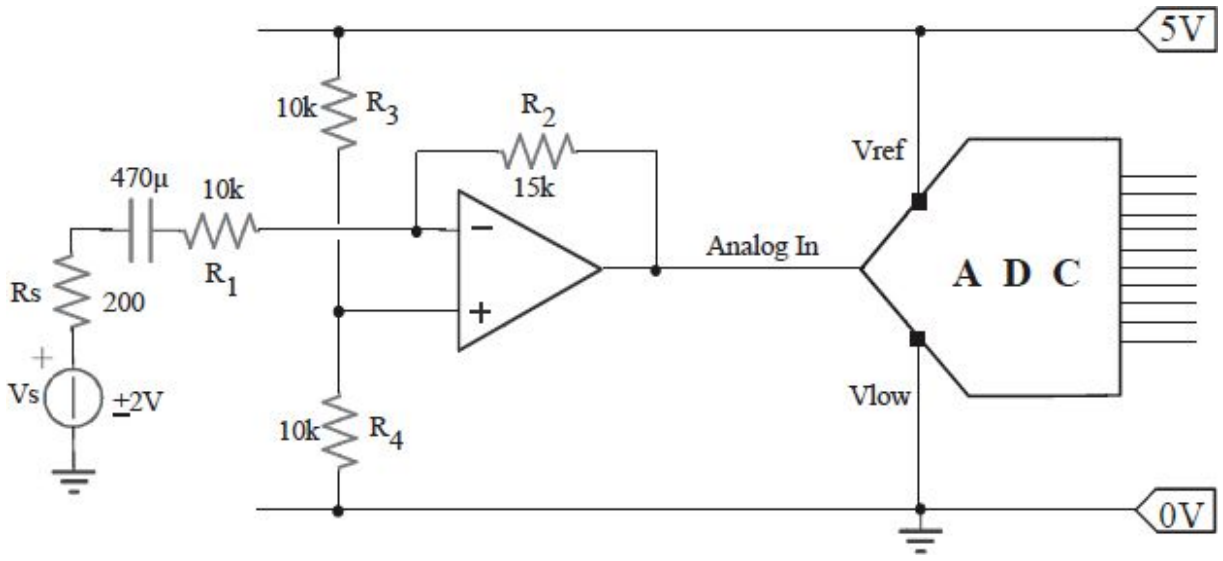


Fig. 8.27: Impedance matching and pre-amplification.

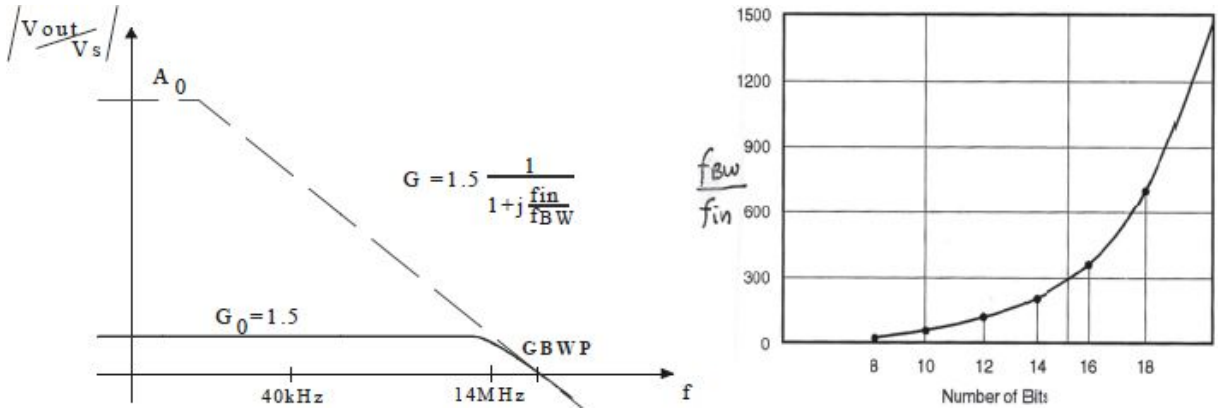


Fig. 8.28: The driver bandwidth (on the left) must be well above the maximum frequency of the input signal. Accuracy dependence on front-end bandwidth (on the right).

In fact, if we had a sine wave of $2V_{picco}$ just at the pole frequency, the attenuation of –3dB would lead to an input signal of the ADC (after the gain of 1.5) of 1.76V, instead of the expected 2.5V, with a subsequent loss of almost 1LSB in the case we were to use a simple 4bit ADC! Instead, desiring to ensure an accuracy of 1/4LSB throughout the input band f_{in}, it is necessary that the frequency of the dominant pole of the front-end (f_{BW}) is much higher than that of interest, according to the relationship:

$$\frac{f_{BW}}{f_{in}} \cong \sqrt{2 \cdot 2^n \frac{V_{inpp}}{FSR}} = (\text{if } V_{inpp} = FSR) = \sqrt{2 \cdot 2^n}$$

In Fig. 8.28, on the right, is plotted this relationship as the number of bits changes. For example, to ensure a 16-bit accuracy, one must have a ratio f_{BW}/f_{in} equal to 400. It will mean that, in this case, having fixed f_{BW}=GBWP/G=14MHz, the maximum frequency input to be applied would be only fin≤35kHz. This is why, even when we want to deal with 'medium frequency' signals (e.g. hundreds of kHz), it is still recommended to use 'wide-band' OpAmps, which have the advantage of maintaining the output impedance low, up to very high frequencies (limiting voltage drops).

Exactly the same principle must be applied for the lower end of the bandwidth. Desiring to make signals pass above 100Hz, it is not enough to size the pole $R_1 \cdot C$ of Fig. 8.27 choosing the usual factor 10 (a decade below), but we have to chose the factor 400 (always to use after a 16bit ADC), i.e. 0.25Hz, i.e. choose C≥63µF.

Slowly varying signals may not require an input S&H (or an ADC with an internal S&H), and the ADC can be chosen among many candidates because the signal is almost a DC. For signals at low frequencies, however, it is necessary to evaluate also the dynamic performance of the ADC (with neglecting the absolute errors of scale, offset, gain, etc. being possible). If the signal has sharp discontinuities, it would be convenient to use single-shot architecture of the type SAR or flash because integration (ramp) ADCs tend to integrate discontinuities providing smooth outputs, or even incorrect.

8.9. Commercial ADCs

The market offers ADCs with very different and new features at low prices, with very low-power features, differential input, 8pin package. Sometimes, this is to the detriment of ease of use (often lack internal blocks, such as reference, clock, and S&H). CMOS ADCs often lack the internal reference because, in this technological process, it is not possible to make precise Zeners or band-gap references since they require the implementation of bipolar transistors. The only possibility is to exploit the parasitic components of the CMOS process as the buried Zener (although the latter has a breakdown

voltage V_z=6V, which is incompatible with the power supply of the logic V_{cc}=5V). It is thus easier to use external band-gap references, individually available in very compact packages, such as the SOT23. In some ADCs, there is also an internal clock generator while, in others, there is not because it would consume power and chip area.

To reduce the power dissipation, the “power-management” is exploited, which deactivates the circuits that are not active, putting them in sleep-mode (also called standby-mode, shut-down, or power-down). Even if we get good results in terms of dissipation, it is necessary to pay attention to the reactivation of the reference, which may take a few milliseconds (depending on the value of the filter capacitor at the pin V_{ref}). Linear Technology, for example, offers two power-down modes: sleep and nap (1μs); some models also have an automatic power down.

Regarding the choice of the model to be used in a project, it is good to see how sometimes it is convenient to use a 12bit ADC ‘high speed’ even for 8bit applications with slow signals because their cost is comparable, and they offer benefits sometimes completely absent in the lower class of converters. For example, a 12bit ‘high speed’ ADC will have greater dissipation than a slower model, but since it can be put in sleep mode all the time of inactivity, it could however ensure lower consumption.

In the choice, we must carefully analyze the dynamic specifications: they are often reported only for sinusoidal signals at very low frequencies, far from the Nyquist. Instead, to access the quality of an ADC, data and graphs should cover f_{in} at least up to 90% of f_{Ny}. Other manufacturers give data at frequencies higher than f_s and then subsample. The reason is that they often lack the instrumentation in the laboratory or the FFT or simply they want to lower the costs of the tests by which they are made.

There are also ADCs with internal analog multiplexer; thus, they provide more analog inputs (8÷12) selectable, but usually they are slow and, therefore, are only used for DC or slowly varying signals. Many of them are SAR and have the function of power-management (they have no internal V_{ref}). Often, to save pins, they have serial output.

8.9.1 Examples of ADCs

12bit ADCs are very popular. The Harris HI5812 at 750kHz costs 4.60$ (in 1000 units) while the National Semiconductors ADC12041 costs 5.50$ (in

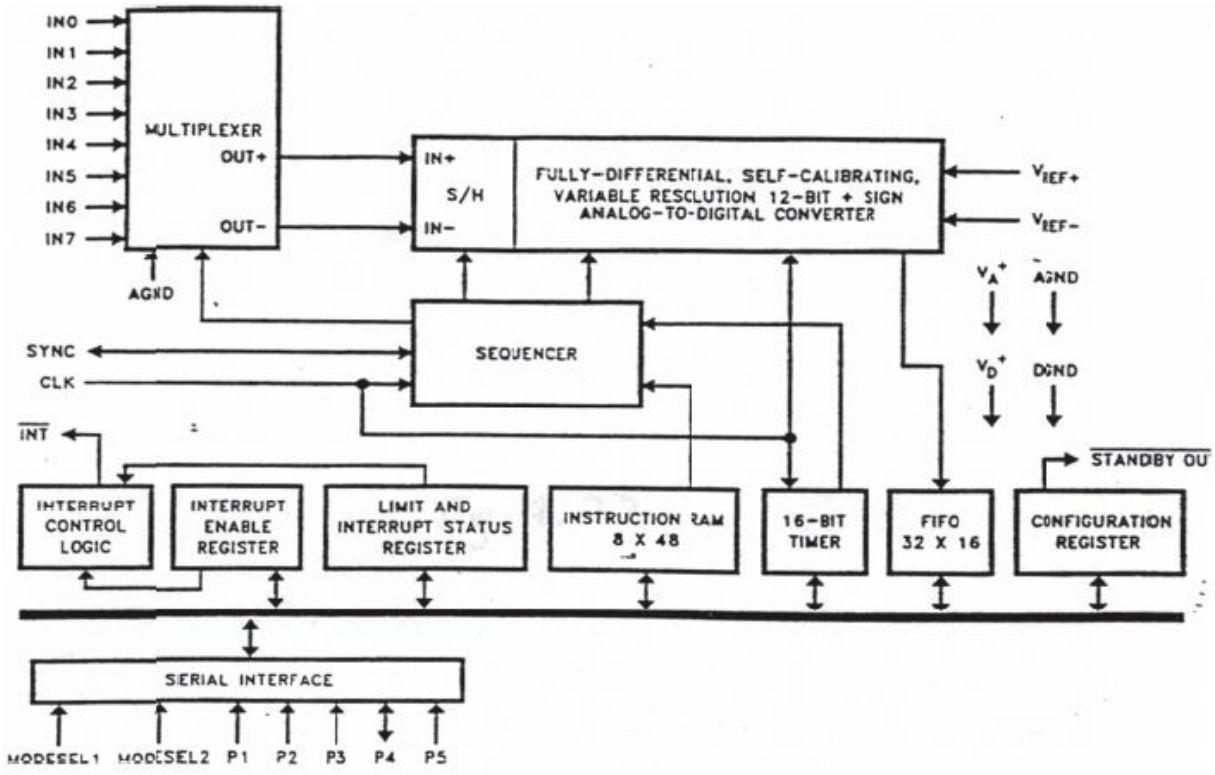
1000 units), but offers 12bit+sign, V_{inMax}=±5V (differential inputs), and a couple of reference inputs; auto-calibration and auto-zeroing; parallel; package type 28 SSOP and 28 PLCC. In addition, for this component, a TUE ("Total Unadjusted Error") of ±1LSB is ensured, which includes DNL, INL, Offset, and Gain and Common-mode Error. MAX1241, however, costs 5.80$ and has 0.5LSB of INL with V_{alim}=2.7V and f_s=70kHz.

LTC1286 – It is a compact 8pin 12bit ADC from Linear Technology and transfers data serially to the microcontroller using SPI ('Serial Peripheral Interface' of the µC). In addition to data, on the serial line travels also the signals that, if the converter allows, can change its gain, the number of bits, the digital format and so on, pointing to the miniaturization.

LTC2400 – Identical to the LTC1286, but with 24 bits, ensures resolution of $5V/2^{24}$, i.e. about 16M levels, but with little accuracy.

LM12438 – There are also more complete and powerful ADCs, such as the Analog Devices LM12438 that contains a 12bit+sign A/D with auto-calibration, an 8 channel multiplexer, an FIFO to record up to 32 results, and an RAM with 8 words of 48 bit each that can be programmed for a series of conversions and comparisons between some user-defined channels and has a full-differential ($6V_{in}$). Thanks to the serial interface of the µC, it is able to adapt to the outside world. The scheme of Fig. 8.29 shows how the LM12438 is a complete acquisition system; it has 3 operation modes: 12 bit+sign, 8 bit+sign, and watch-dog mode, in which no conversion is done, but samples and compares a selected input until it exceeds a predefined value stored in RAM by the user, generating then an interrupt to the µC. The dissipation is very low, maximum 45mW, and goes down to 25µW when in standby.

Concerning compact ADCs in SO-8 package, Fig. 8.30 lists the characteristics of a 24bit ADC, which costs only 6.00$.



INTERFACE	MODESEL1	MODESEL2	P1	P2	P3	P4	P5
Standard	0	1	$\overline{R}$/F	$\overline{CS}$	DI	DO	SCLK
8051	0	0	1*	1*	$\overline{CS}$	RXD	TXD
I2C	1	0	SAD0	SAD1	SAD2	SDA	SCL
TMS320	1	1	FSR	FSX	DX	DR	SCLK

Fig. 8.29: Scheme of the LM12438.

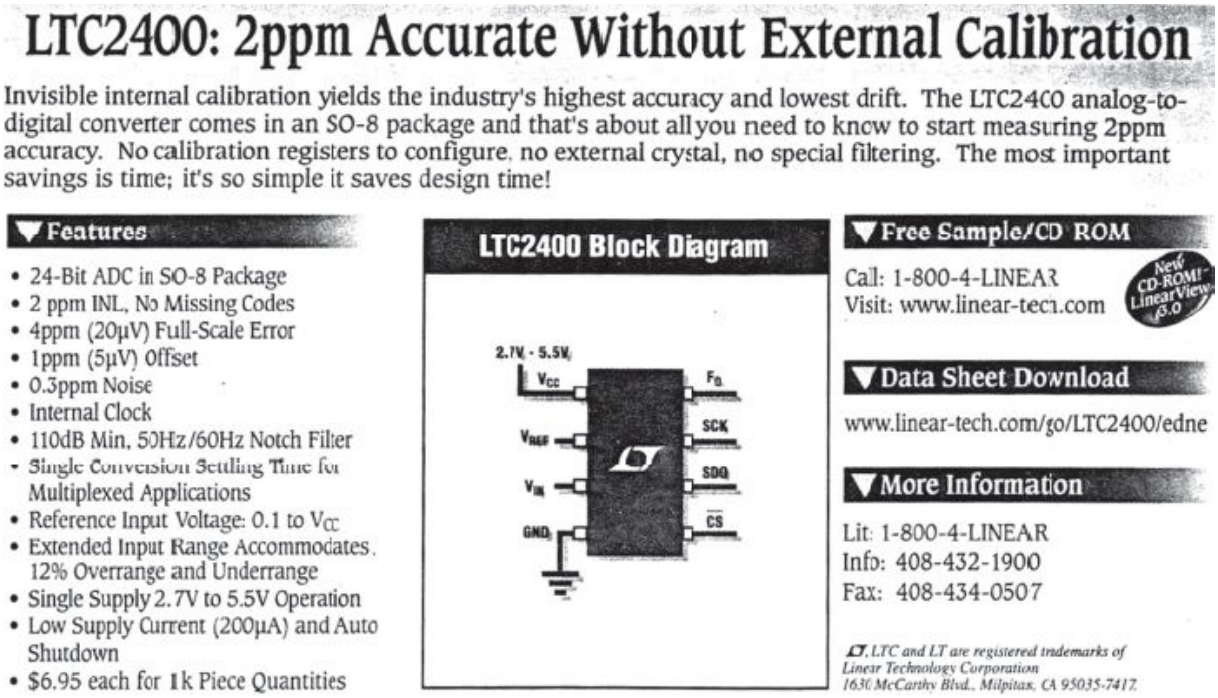


Fig. 8.30: Performance of a compact 24bit ADC.

8.10. DITHERING

The ‘dithering’ is a technique used to reduce non-linearity effects to which ADCs are subjected. Non-linearities are often measured with the parameters DNL and INL that, however, are only useful as long as the ADC is used to punctually acquire memory of the data, e.g. in measuring apparatus or waveform digitizers. For dynamic applications, with audio or video signals, it is important to control the amount of spurious tones and secondary harmonics in the output spectrum. The correlation between DNL and SFDR is not immediate.

To reduce the DNL of an ADC, we can deliberately vary the input signal through multiple adjacent codes in order to average the error superimposed on the signal. The “dither” is just a non-correlated signal, usually pseudo-random, injected into the analog input of the ADC together with the signal to be converted and then subtracted at the output. In Fig. 8.31 is shown the subtractive Dithering technique. We can use a broadband noise or simply use a pseudo-random values number generator. This technique of subtractive Dithering is valid for large dither signals.

Another method is to generate some “analog noise” that is still outside the band of interest. Usually is used a low frequency noise or a noise with bandwidth centered on the Nyquist to ensure a broad-band of available frequencies for the signal; Fig. 8.32 summarizes the out-of-band Dither technique. In both cases, the aim is to relocate and make the DNL of different codes of the converter random. Quantifying the dither added in terms of effective codes (steps), it is possible to analyze the effect of dithering on the resulting DNL. Fig. 8.33 shows the resulting DNL when to the ADC is added a dither of intensity (rms codes) variable and increasing (from top left to bottom right). For example, the last with 21.3 rms codes wants to give a dither 128 peak-peak codes with respect to the 12bit ADC (4096 codes). It makes no sense to dramatically increase the intensity of the dither since it would give very little relative improvements.

In the case of AD9042, the optimal dither is between 16 and 21.3 rms codes, corresponding to a noise power of –35dBm and –32.5dBm. Without the dither, this ADC has an SFDR=82dB, as shown by the spectrum in Fig. 8.34. The FFT is over 128K samples

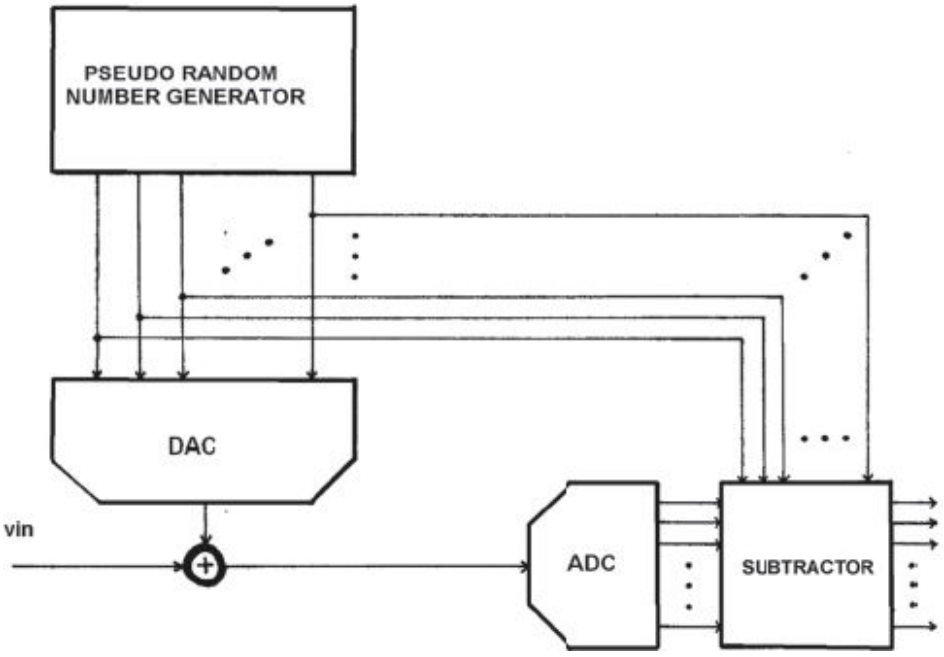


Fig. 8.31: Subtractive Dithering technique; some numerical noise in analog form is injected at the input, and then the corresponding numerical value is subtracted at the output.

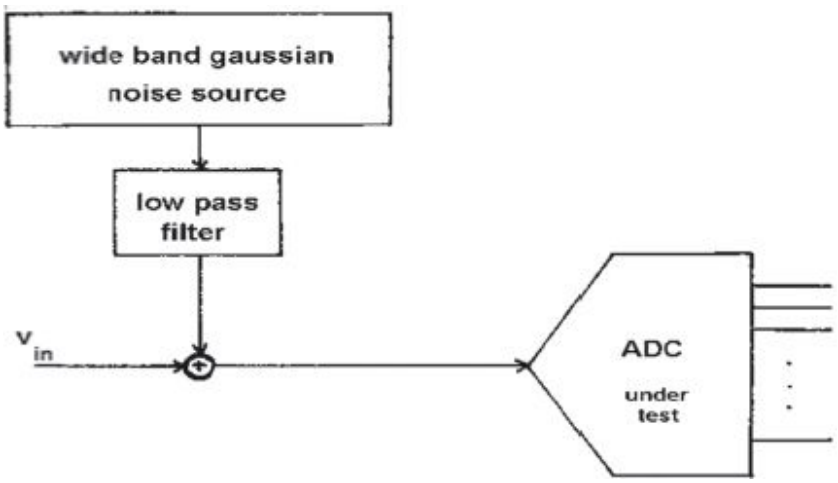


Fig. 8.32: Dither generated by means of an analog noise with spectrum outside the useful bandwidth of the signal; at the ADC output is made proper filtering (average) of the samples to eliminate the out of band dithering.

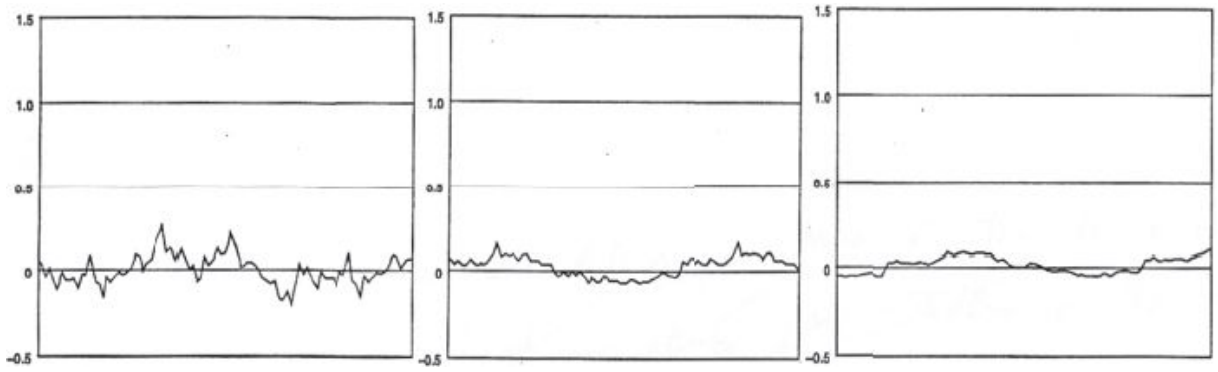


Fig. 8.33: Effect on the DNL of dithering, by applying a fluctuation in effective value, corresponding to 5.3, 16, and 21.3 codes RMS.

when we expect a Noise-Floor=6.02·12-1.76-10·log(128·1024/2), i.e. -120dB theoretical. Actually, other electronic noise determines an elevation of about 10dB, as well as the appearance of many spurious harmonics. The improvement obtained with the dither is clearly visible from the two spectra shown in Fig. 8.34. By applying the dither, the spurious harmonics pass from –82dB$_c$ to even –103dB$_c$. We can clearly see the dither noise outside the useful bandwidth, assumed around 1kHz.

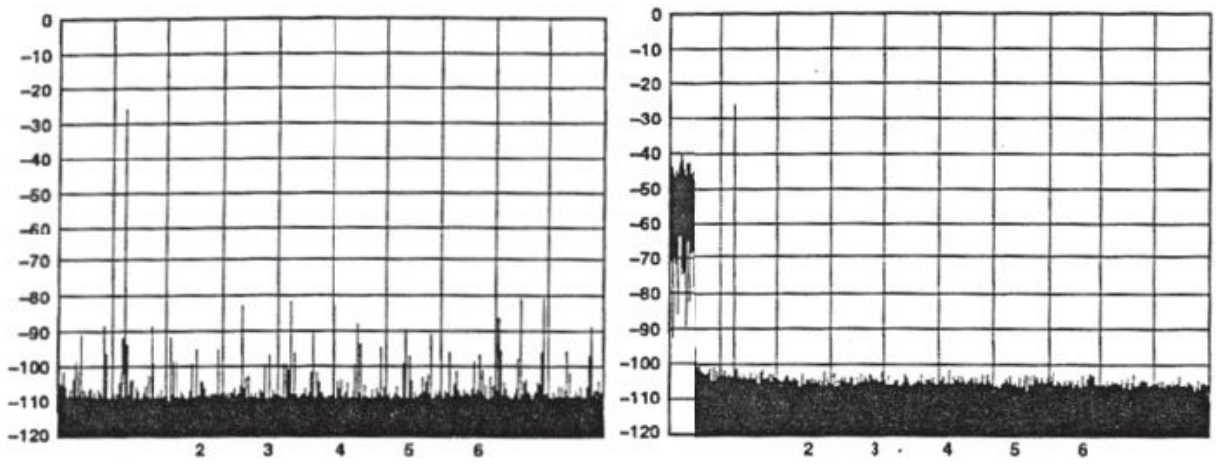


Fig. 8.34: Spectra without (on the left) and with (on the right) the use of dither (FFT over 128k samples).

8.11. EXAMPLES

In this section, we will analyze some criteria for the selection of an ADC, based on project requirements. Suppose that we need to measure, with precision of 0.1°C, a temperature in the range –100°C÷+300°C. We use a resistance thermometer PT100, which is equal to 100Ω at 20°C and has a temperature gradient of +0.385Ω each 1°C; in case of maximum thermal excursion, it will assume the limit values 61.5Ω÷215.5Ω. To achieve the desired accuracy, it will be necessary to divide the available range in $4096=2^{12}$ parts, corresponding to resolution of 1LSB, i.e. 38.5mΩ or just 0.1°C.

8.11.1 Choice A

The simplest choice seems to be adopting a 12bit ADC, shown in Fig. 8.35. In this first solution, we can think of injecting a high current, I=23mA, into the resistance thermometer in an attempt to exploit the whole dynamics of the converter. In this way, in fact, at 300°C, the voltage at the ADC input will be just equal to the power supply +5V. Unfortunately, at –100°C, the voltage will not reach 0V, but will go down only to +1.43V. Therefore, we can exploit only the 72% of the FSR, resulting in the maximum resolution of only 11.5bit, instead of the available 12 bits (V_{ref}=3.57V instead of 5V).

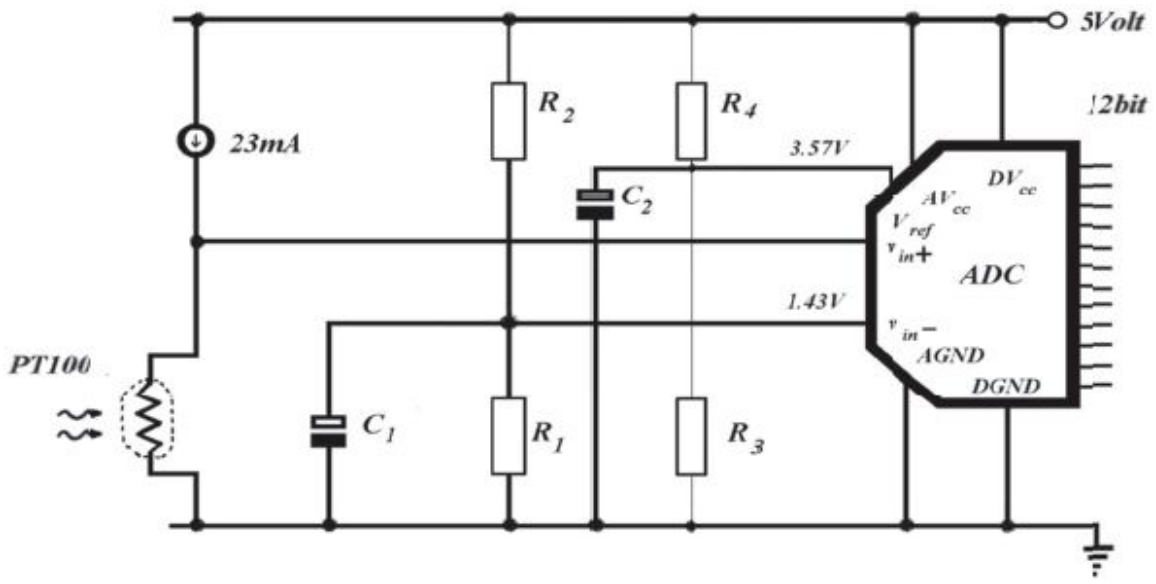


Fig. 8.35: Measure with a 12bit ADC; implementation of the choice A.

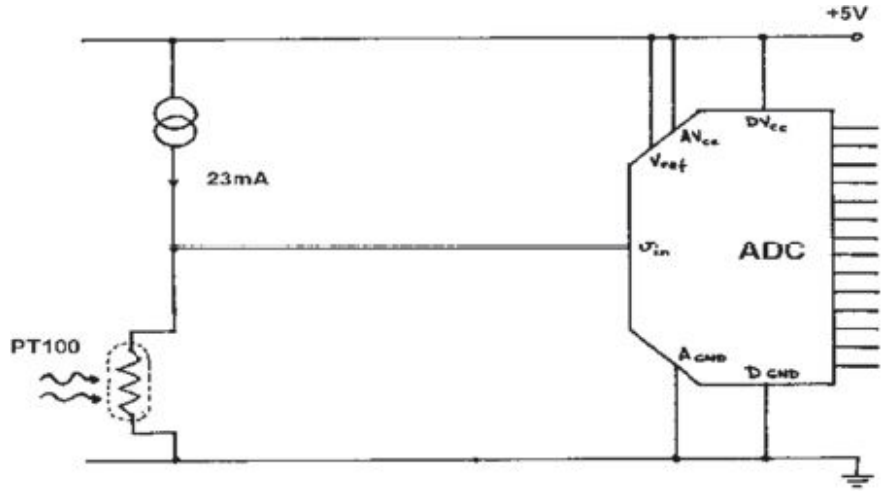


Fig. 8.36: Measure with a 14bit ADC; implementation of the choice B.

To overcome this difficulty, we can take a converter with differential input, introducing a common-mode offset of 1.43V and using a reference voltage of 3.57V. This type of ADC is more expensive and requires further stabilization of V_{ref} through appropriate filtering. We should also consider the thermal noises and resistance tolerance degrees, which can lead to inaccuracies in reference levels. In fact, a 1% tolerance is sufficient to cause the same change in the nominal value of the input (equal to 1.43V), according to the equation:

$$V_{in-} = +5V\frac{R_1(1\pm toll)}{R_1(1\pm toll)+R_2(1\mp toll)} = +5V\frac{R_1(1\pm toll)}{R_1+R_2} \cong +5V\frac{R_1}{R_1+R_2}(1\pm toll)$$

With the degree of precision required in this project, the variation of (±1.43V)·1%=14.3mV corresponds to a variation of ±16 codes (being 1LSB=3.57V/4096=872μV). The accuracy is therefore reduced to just 3.57/14.3mV=8 effective bits.

8.11.2 Choice B

A second solution is to use an ADC with higher resolution, which requires less external components (Fig. 8.36). With the demand for precision of 0.1ºC, corresponding to 1LSB=23mA·0.1·0.385Ω=1mV on 5V of the FSR, 5V/1mV=13bits would be enough. Adopting a 14 or even a 16bit ADC, despite not exploiting the full dynamic range, we get resolution higher than the base 12 bits.

Unfortunately, both solutions suffer from the problem of self-heating since the dissipation $R\cdot I^2$=100·(23mA)^2=78mW can cause a change in the temperature of the PT100 itself. This causes an additional error in evaluating the temperature read by the sensor. To avoid this, it is necessary to bias the PT100 with currents lower than mA: a current of only 500μA would produce a signal of only 77mV between 30.75mV and 107.75mV. It is possible to use this solution only by introducing a gain block because the full range would no longer be covered with the same precision of the 4096 levels.

8.11.3 Choice C

This solution, very expensive, suggests the use of an INA (compensates for the reduction in the current source) that has to be able to operate with a dual polarity power supply, not only with +5V (Fig. 8.37). By selecting the gain with an external resistor R_g, in order to have an amplification of 64.8V/V, we can use the full FSR=5V available.

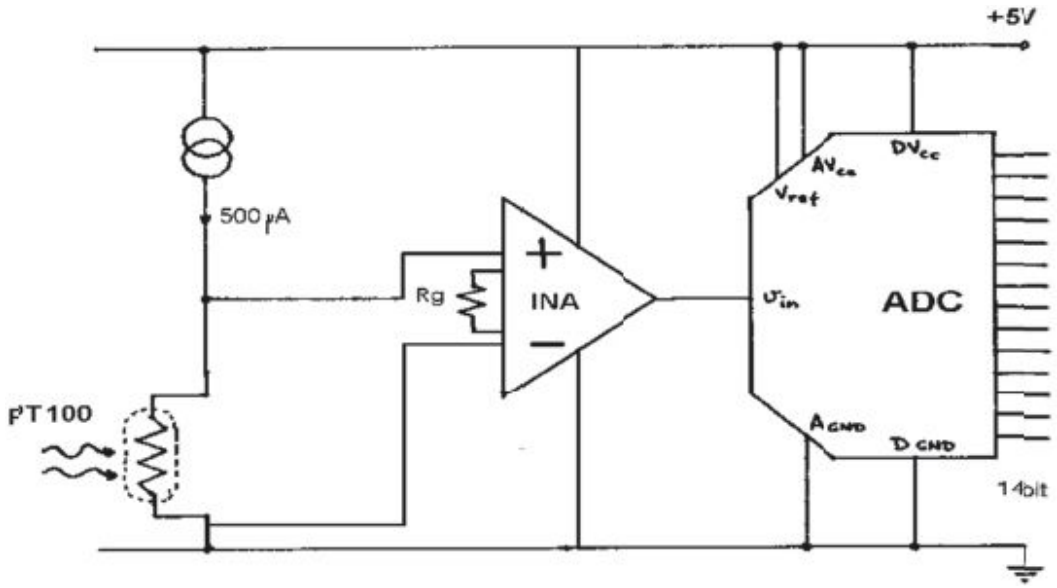


Fig. 8.37: Measure with a 14bit ADC and an INA; implementation of the choice C.

If, however, the INA were not able to guarantee the rail-to-rail output-swing (most likely working with only 5V of power supply), we would be obliged to reduce the gain and the output range of the stage. Again, we must use a 14 or 16bit ADC to ensure the accuracy required after the conversion.

8.11.4 Choice D

It proposes again the choice B using only an 18bit ADC; thanks to the development and cost reduction of ADCs, this could be the most economical solution. Since, with a current I=500μA, we want to resolve 1LSB=19.25μV with FSR=77mV using $n=\log_2(5V/1LSB)=18bit$, we can choose the ADC that can provide such precision even without a gain stage. The choice of an 18bit ADC (high resolution), for reasons of precision and accuracy, is likely to lead to the use of an oversampled ∑Δ modulator.

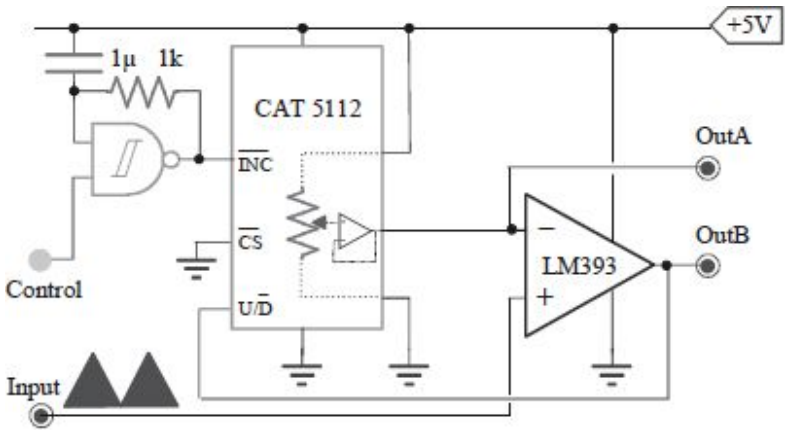
8.12. EXERCISES

Ex. 1

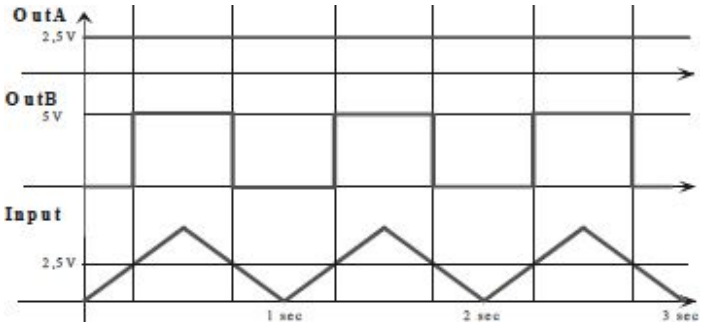
The digital potentiometer has *64taps* (**INCrement** moves the cursor up/down depending on the level of the pin **Up/Down**; at the turn on of the circuit, the cursor is in the middle of the dynamics).

a) Draw **OutA** and **OutB** when **Control**=*low* and with a triangular wave *0÷5V* with a period of *1s*.

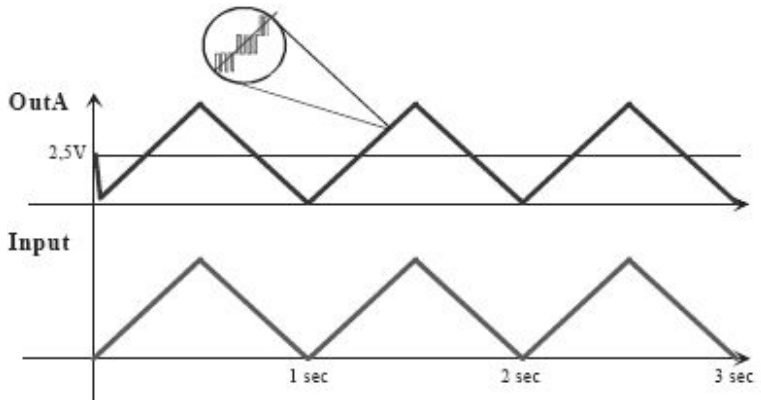
b) With the same input, explain the operation of the circuit when **Control**=*high*.



a) When Control is LOW, the CLK at the input of /INC is off; then, if Control has not been asserted before, the digital potentiometer is located halfway, with OutA fixed at 2,5V. The OpAmp works as a comparator between the output signal of the potentiometer (OutA) and Input. For Input>OutA, the output (OutB) is saturated at 5V. For Input<OutA, OutB=0. Graphically, we have the following operation.



b) With Control High, the potentiometer works “following” the signal Input. As can be seen, the control over the direction of the count is done by the comparator: when the code expresses a voltage higher than the input, at every clock pulse, the potentiometer decreases its output voltage and, when it is lower, increases it. When the circuit is switched on, the potentiometer immediately hooks the signal (in less than (64/2)·1μF·1kΩ=32ms) and provides an output that is a quantized version of the signal in Input. OutB indicates instant by instant if OutA<OutB. Graphically, this is the operation.

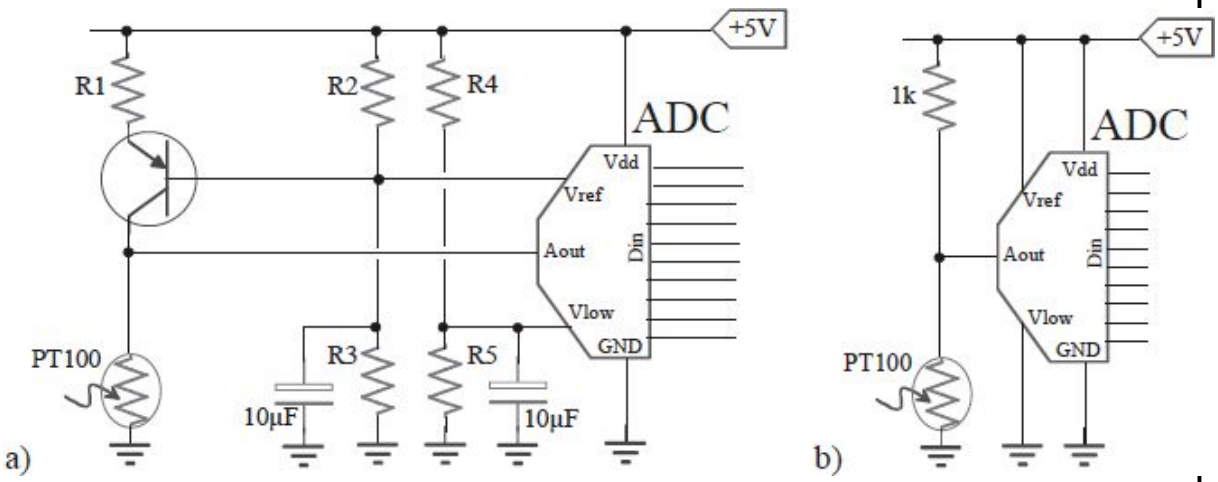


Ex. 2

We want to measure a temperature between *0÷500°C* with *10bits* with a PT100, which is *100Ω* at *0°C* and increases by *0.4Ω/°C*.

a) Using a *10bit* ADC, size the components of the circuit (a) to fully exploit the allowed dynamics.

b) Considerably simplify the circuit, as shown in (b): analyze the non-linearity introduced in this way and choose the proper ADC to ensure the same precision.



Ex. 3

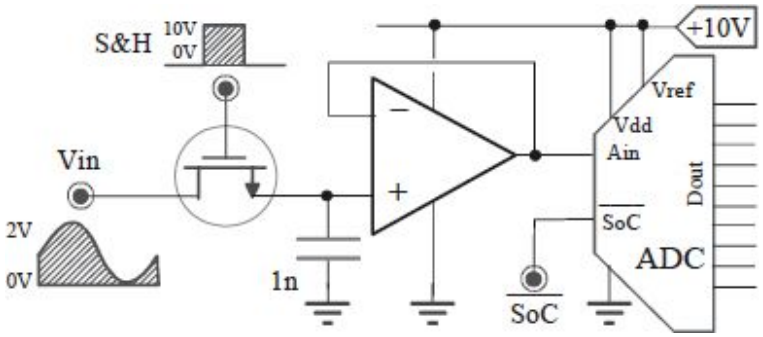
OpAmp with $A_0=100dB$ and $I_b \leq 100nA$; MOS with $V_t=1.5V$, $C_{GD}=C_{GS}=10pF$, and $R_{DSon} \leq 50\Omega$; input between $0 \div 2V$; *12bit* ADC.

a) Calculate *SNRideal* and *SNRreal* at the output of the ADC, which are caused by the only quantization error.

b) Calculate the R_S limit of the source to ensure an acquisition time of less than $1\mu s$

c) Calculate the static errors accumulated passing in Hold.

d) Modify the circuit to exploit the whole FSR of the ADC.



a) The SNR_{ideal} of a converter is the ratio between the power of the signal with full dynamics and the power (variance) of the only quantization noise introduced by the conversion. Remember that the power of the white quantization noise is equal to:

$$\sigma^2 = \int_{-\frac{LSB}{2}}^{\frac{LSB}{2}} \varepsilon_q^2 \frac{1}{LSB} d\varepsilon_q = \frac{1}{LSB}\left[\frac{\varepsilon_q^3}{3}\right]_{-\frac{LSB}{2}}^{\frac{LSB}{2}} = \frac{1}{LSB}\frac{2}{3}\frac{LSB^3}{8} = \frac{LSB^2}{12}$$

So $SNR_{IDEAL}|_{dB}$ is: 10Log(Power of the full dynamic signal /σ^2), namely:

$$10Log\frac{\left(\frac{FSR/2}{\sqrt{2}}\right)^2}{\frac{LSB^2}{12}} = 10Log\left(\frac{12}{8}2^{2n}\right) = 6.02n + 1.76\ [\text{dB}]$$ where n is the number of bits of the ADC.

In this case, therefore, 6.02·12+1.76=74dB.

Even if the quantization error were the only noise, the conversion is often not performed on a full dynamic signal. This affects the quality of the conversion itself, reducing the $SNR_{theoretical|dB}$=10·Log(Power of the input signal /σ^2), namely:

$$10Log\frac{\left(\frac{Range/2}{\sqrt{2}}\right)^2}{\frac{LSB^2}{12}} = 10Log\frac{\left(\frac{2V/2}{\sqrt{2}}\right)^2}{\frac{\left(\frac{10V}{2^{12}}\right)^2}{12}} = 60\ [\text{dB}]$$

If another noise is also added, then we come to $SNR_{real|dB}$=10·log(Power of the input signal /total variance of all the noises present).

b) R_S is to be considered in series with the resistance $R_{DS,ON}$ of the conducting MOS. From the equation of charge of an RC network, we find that the charging time should be $t_{ACQ} \geq \tau \cdot \ln(\Delta/\varepsilon)$, where τ=RC is the time constant of the network, Δ is the excursion in voltage required, and ε is the error allowed. In our case, sizing to meet the specification $\varepsilon \leq LSB/2$ in the worst case (Δ=2V), we get: t_{ACQ}=1μs≥ $\tau \ln\left(\frac{2V}{2.44mV/2}\right)$, from which $\tau \leq 135 \cdot 10^{-9}$ → $R \leq 135 \cdot 10^{-9}/1nF = 135\Omega$. Since $R=R_S+R_{DS,ON}$, we have $R_S \leq 135\Omega - 50\Omega = 85\Omega$.

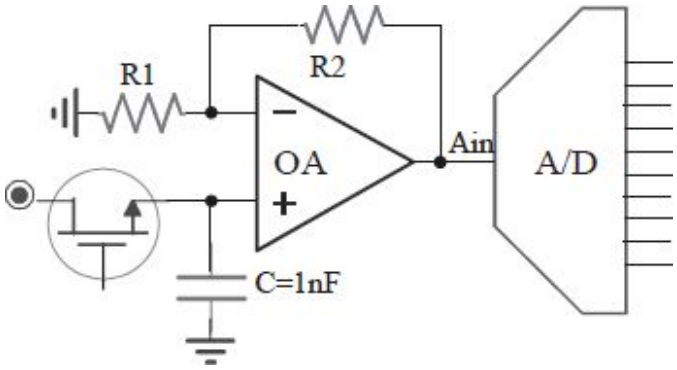
c) The main static error in Hold is the Charge Injection due to the presence of the parasitic capacitance C_{GS} that causes a shift in the voltage stored on C_H during the transition of the voltage applied to the GATE from 3.5V ($V_{IN,MAX}+V_T$=2+1.5) to 0V. At the opening of the switch, a charge injection of unwanted charge occurs on C_H, which is equal to:

$$\Delta V = 3.5V \frac{C_{GS}}{C_{GS} + C_H} = 3.5V \frac{10pF}{10pF + 1nF} = 35mV$$

In addition to this, the droop effect of bias currents on the Hold capacitance:

$$\frac{dV}{dt} = \frac{I_B}{C_H} = \frac{100nA}{1nF} = 100\frac{V}{s} = 0.1\frac{mV}{\mu s}$$

d) Desiring to exploit the whole dynamics (0÷10V) with an input 0÷2V, it is sufficient to amplify the signal by a factor of 5. This can be trivially obtained by exploiting this OpAmp. As in the figure:



As is known, the gain V_{OUT}/V_{IN} is equal to $1+R_2/R_1$; we have to use then two resistors such that R2=4·R1, for example R2=40kΩ and R1=10kΩ.

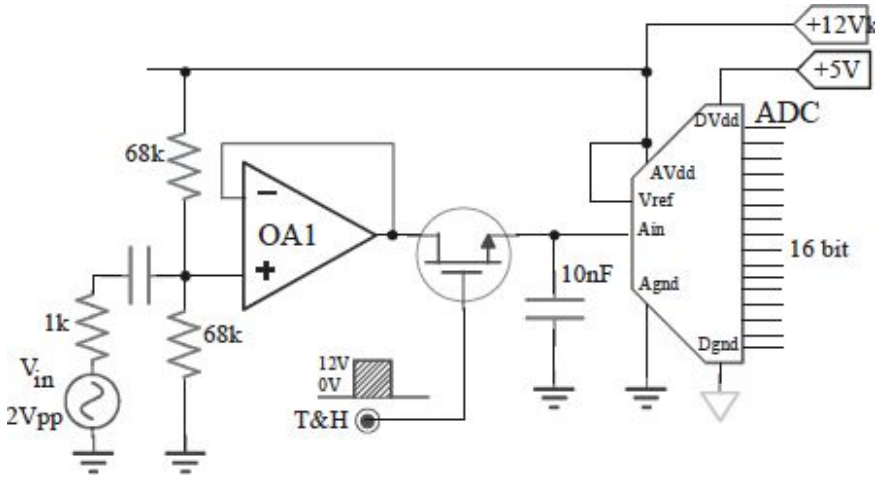
Ex. 4

The OpAmp has *IB=+300nA, VOS=3mV,* and *A0=100dB.* Downstream, there is a 16bit ADC with *Vref=12V*.

a) Calculate the maximum static errors at the output of the OpAmp, in *LSB*, due to *IB, VOS*, and A_0.

b) Calculate the ENOB with a sinusoidal input of $1V_p$.

c) Modify the front-end stage to exploit the whole dynamics of the ADC input.

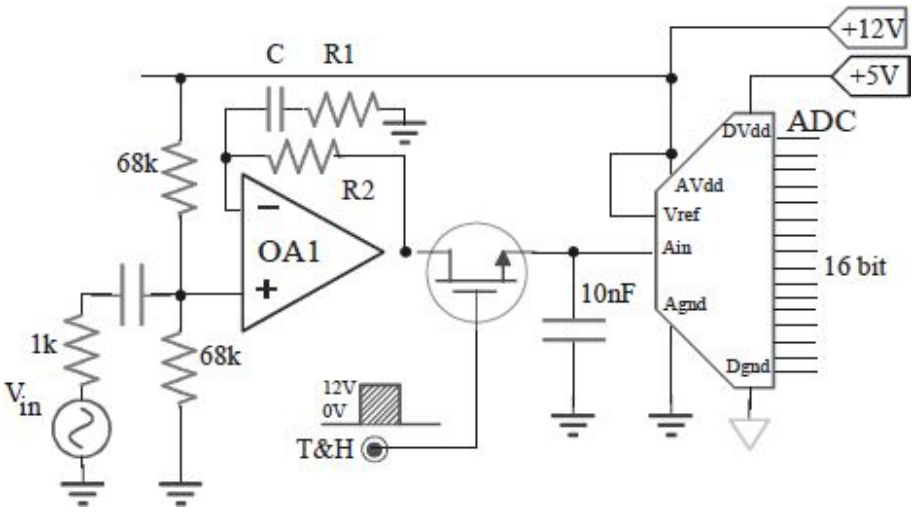


a) The 16bit ADC has V_{REF}=12V, then LSB=$LSB=\frac{12V}{2^{16}}=183\mu V$. Since the OpAmp is in buffer configuration, V_{OS} is transferred to the output with unity gain. In terms of LSBs, it is equal to 3mV/183μV=16.4 LSB. The bias current coming out from the positive terminal flows half into the branch to the ground and the other half into the branch to the power supply (in DC, the input capacitor acts as an open circuit). Considering the branch to the ground, the current produces a voltage drop at the non-inverting terminal, which is equal to: $\frac{300nA}{2}\cdot 68k\Omega=10{,}2mV=56LSB$. As can be seen, even in this case, the voltage at the non-inverting terminal is transferred to the output with unity gain by the buffer. The bias current coming out from the inverting terminal,

instead, does not cause changes in the output voltage. Since the gain A_0 is limited, the output voltage is $V_O=(V_{IN}-V_O)\cdot A_0$ (V_O is brought to the input by the feedback branch of the buffer). The error ε, i.e. the non-zero difference ($V_{IN}-V_O$) due to a finite gain is hence equal to: $(V_{IN}-V_O)=\varepsilon=V_O/A_O$
In the worst case, i.e. with the input voltage equal to 7V (A_0=100dB=100000), ε is 7V/100000=70μV=½LSB.

b) Since there is only 1V of peak of input signal, the entire conversion dynamics of the ADC is not used, causing a loss in terms of performance.
The input signal, peak to peak, has a range of 2V, i.e. only one-sixth of the dynamics (12V). This means that, of the 2^{16}=65536 available levels, only 65536/6=10923 are used, i.e. only those of an equivalent ADC with 13 bits (ENOB) are exploited, operated at full dynamics.

c) To exploit the whole dynamics of the signal, the amplifier stage has to introduce a gain equal to 12V/2V=6. In non-inverting configuration, this means to introduce two resistors in the feedback branch, as shown in the figure.



Since the gain in this configuration is equal to 1+R2/R1, we choose R2 and R1 in 5:1 ratio. A capacitor C is inserted in the branch of R1 in order to isolate R1 in DC. To cancel the contribution of the input bias currents, we make the OpAmp see the same resistance on both terminals. Then, R2 is chosen 68kΩ//68kΩ=34kΩ. R1, therefore, is 34kΩ/5=6.8kΩ. Obviously, introducing

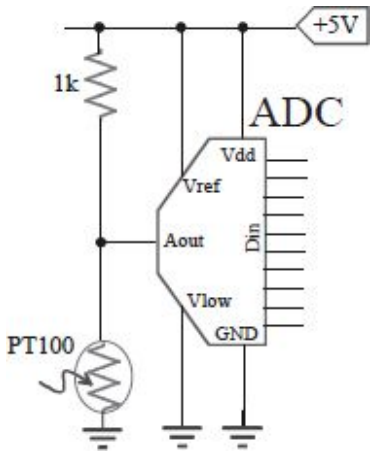
the capacitance C, we insert a pole in the transfer function, which, in order not to alter the function of the circuit, must be placed outside the useful bandwidth.

Ex. 5

Measure the temperature between *0÷500°C* with resolution of *0.1°C*. The PT100 is *100Ω* at *0°C* and increases by *0.4Ω/°C*.

a) Determine the range of variability of the voltage across the PT100 and determine the bits needed for the ADC.

b) Due to the presence of a ripple on the power supply of *100mV* at *100Hz*, size the filter capacitor to be placed at the input of the ADC to ensure an error of *½LSB*.



a) The temperature sensor consists of a resistance thermometer with the temperature coefficient equal to 0.4Ω/°C. The voltage that develops across the resistance thermometer varies depending on the operating temperature; this tension is also the input of an ADC whose reference voltages V_{ref} and V_{low} are respectively put to 5V and 0V. If we desire to measure a temperature between 0÷500°C, the voltage at the input of the ADC will be:

$$Vin = 5V \cdot \frac{PT100}{PT100 + 1k\Omega} = \begin{cases} 5V \cdot \dfrac{100\Omega}{100\Omega + 1k\Omega} = 0.45V & @0°C \\ 5V \cdot \dfrac{300\Omega}{300\Omega + 1k\Omega} = 1.15V & @500°C \end{cases}$$

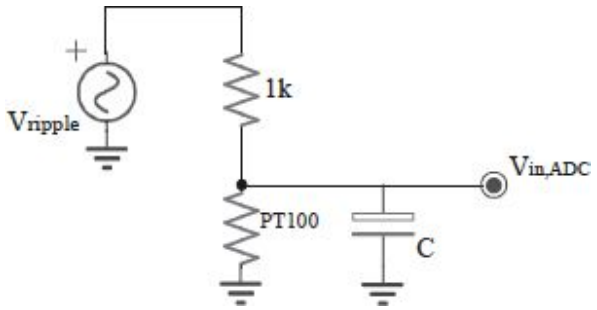
Since the required resolution is 0.1°C, in the range of tensions now calculated, there should be 500°C/0.1°C=5000 digital levels. Then, the entire Full Scale Range (FSR) will have to contain 5000 *levels* : $(1.15V - 0.45V) = x : 5V \rightarrow x = 35714$ *levels*. Therefore, the number of bits needed by the ADC is equal to $n \geq {}_{log}235714 = 15.1bit \rightarrow n = 16bit$.

Unfortunately, the circuit does not have a perfectly linear behavior; in fact, the current flowing in the PT100 is not constant, but varies depending on the value of the PT100 itself. In other words, the current flowing in the PT100 varies depending on the voltage that develops across it and, ultimately, on the temperature; this causes a non-linearity and implies that the value of the LSB is not constant over the entire voltage range, but is a function of the operating working point. A possible solution to this problem is to replace the resistance with an "ideal" current source, such as a pnp bipolar transistor emitter degenerated or a MOSFET source degenerated.

b) If the power supply is not adequately stabilized, some oscillations, just ripple, may occur, which cause errors in the stage. A ripple on the power supply, if not properly filtered, could create, in the worst case, a fluctuation in the input, which is equal to:

$$\Delta V_{in,senza\ filtraggio} = V_{ripple} \cdot \frac{PT100_{max}}{PT100_{max} + 1k\Omega} = 23mV = 302LSB!!$$

There are two solutions to be adopted: we can use a 5V voltage regulator with tighter specifications or introduce a filter capacitor in parallel to the PT100 at the input of the ADC. The equivalent circuit at the input of the converter is:



The error voltage at the input ε_{in} is now equal to:

$$\varepsilon_{in} = V_{ripple} \cdot \frac{PT100 \,//\, \frac{1}{sC}}{\left(PT100 \,//\, \frac{1}{sC}\right) + 1k\Omega} \cong V_{ripple} \cdot \frac{\frac{1}{sC}}{\frac{1}{sC} + 1k\Omega}$$

The approximation made, which may be justified in retrospect, comes from the fact that, at 100Hz, the impedance of the lower branch must be very small to make the error negligible, which is introduced by the ripple. In addition, 1/sC<<1kΩ. So, desiring an error of less than ½LSB, it has to be:

½LSB, it has to be:

$$\varepsilon_{in} = V_{ripple} \cdot \frac{\left|\frac{1}{sC}\right|}{1k\Omega} \le \frac{1}{2} LSB \quad \rightarrow \quad \frac{1}{\omega C} \le \frac{1}{2} LSB \cdot \frac{1k\Omega}{V_{ripple}} = 0.38\Omega$$

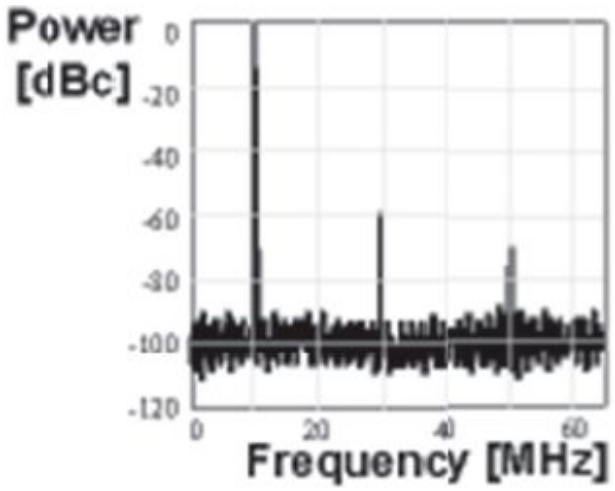
Occurs that 1/sC << PT100. Thus, we find:

$$C \ge \frac{1}{2\pi \cdot 100Hz \cdot 0.38\Omega} = 4.2mF!!$$

In doing so, we have introduced a filter that lowers also the signal components having frequencies above 100Hz. If this were a problem, it would be necessary to remove the capacitor and finally resort to a current source in place of the 1kΩ resistor.

Ex. 6

The S&H SHC605 works at f_s=*128MHz* and has a sinusoidal input of ±*1V*. The output spectrum shown on the right has a *bin-width=1kHz*.
a) Calculate *SiNAD* and *SFDR*.
b) Calculate the total harmonic distortion.



a) In the figure is shown the spectrum calculated at the output of a stage S&H and ADC when at its input is sent a very pure analog sine wave. The FFT algorithm uses N samples of the signal (in time domain) for total duration of $T_{acquisition}=N\cdot T_{sampling}$ and provides as many samples in the frequency domain. The frequency resolution, entitled *bin-width*, is equal to $BW=f_{sampling}/N$. In this case, the number N of samples used can be obtained as: $N=\frac{f_{sampling}}{BW}=\frac{128MHz}{1kHz}=128000$
It is useful to note that the sampling frequency $f_{sampling}$ is equal to twice the maximum frequency shown in the spectrum. This follows from the fact that the spectrum is calculated through the FFT, as well as being discrete (i.e. consisting of N samples), it is also symmetric with respect to $f_{sampling}/2$ (a property of all real signals). Accordingly, it is sufficient to graph only half of the spectrum, implying that the second half is the exact replica.

The SiNAD is the ratio between the power of the input sinusoidal signal and that obtained by adding the power of all the noise and all the unwanted harmonics generated by the distortion made by the stage. The signal, the harmonics, and the noise may be obtained directly from the spectrum. These are expressed in power and in dBc, where the subscript stands for “carrier”, i.e. they are related to the power of a sine wave having the maximum peak amplitude allowed, which is FSR/2:

$$P_{0dBc} = \left(\frac{FSR/2}{\sqrt{2}} \right)^2 = 0.5V^2 = (0.707V)^2$$

For what concerns the noise, we must consider that, having a flat spectrum (white noise), its power is divided into the N/2 histograms, which are shown in the spectrum; in practice, in each histogram falls only a fraction equal to 2/N of the whole noise of the stage. For this reason, the total noise is calculated by multiplying the average value of the background (called Noise-Floor) by the number N/2 of samples, making sure to use consistent sizes:

$$N_{reale} = -100dBc \cdot 128000/2 = !!! = -100dBc + 10\log(128000/2) = -52dBc$$

The exclamation marks highlight the relations of equalities mathematically incorrect, but are written in a useful way from a practical point of view to easily move to dB.
The harmonic distortion produces spurious spectral components at frequencies multiple of that of the fundamental. The total power of the harmonics is calculated as follows:

$$D_{real} = -60dBc - 70dBc = !!! = 0.5V^2 \cdot \left(10^{-\frac{60dBc}{10}} + 10^{-\frac{70dBc}{10}} \right) = (742\mu V)^2 = -59.6dBc$$

We find:

$$SiNAD = 0dBc - [(-52dBc) + (-59.6dBc)] = !!! = \frac{0dBc}{10\log\left(\frac{0.5V^2 \cdot \left(10^{\frac{52}{10}} + 10^{\frac{59.6}{10}} \right)}{0.5V^2} \right)} = 51.3dB$$

The SFDR (*Spurious Free Dynamic Range*) is defined as the ratio between the useful signal and the highest spectral contribution. From the graph, it is equal to -60dBc, from which:

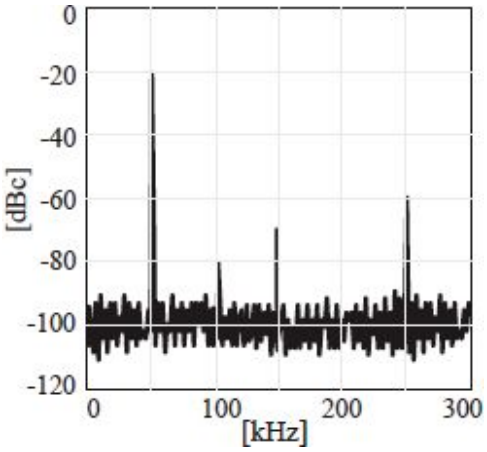
$$SFDR = 0dBc - (-60dBc) = 60dB$$

b) The THD (*Total Harmonic Distortion*) takes into account the non-linearities introduced by the stage and is calculated as the ratio between the power of the harmonics and the power of the fundamental. Since it is NOT possible to add powers in dBc, it is necessary to convert all of them into voltage values, then add them up, and eventually convert all of them into dBc, as shown in the following expression:

$$THD = \frac{10\log\left(\frac{0.5V^2 \cdot 10^{-60/10} + 0.5V^2 \cdot 10^{-70/10}}{0.5V^2}\right)}{0dBc} = \frac{D_{real}}{S_{real}} - 59.6dB$$

Ex. 7

The output spectrum of a *14bit* ADC with *FSR=5V* is shown. The FFT was calculated on a data stream with duration of *213ms*.
a) Calculate the theoretical and real *SNR* and the *ENOB.*
b) Calculate the *THD* and the *SFDR*.



a) The figure shows the spectrum calculated at the output of an ADC when at its input is sent an analog sine wave. In this case, the data stream used for the calculation of the spectrum is equal to 213ms, which yields the number N of used samples:

$$N = \frac{T_{acquisition}}{T_{sampling}} = T_{acquisition} \cdot f_{sampling} = 213ms \cdot 600kHz = 127800$$

The theoretical signal to noise ratio is the upper limit reachable from the ADC; it can be calculated by considering a sinusoidal input with the maximum amplitude allowed (FSR/2) and the only quantization error that is inevitably present in each ADC, whose total power is σ^2_q=LSB^2/12. We find that:

$$SNR_{theoretical} = 20\log\left(\frac{\frac{FSR/2}{\sqrt{2}}}{\sqrt{\frac{LSB^2}{12}}}\right) = 6.02 \cdot n + 1.76 = 86dB$$

Actually, the real SNR measured at the output of the ADC will never be equal to that just calculated, and this is due to both the presence of more noise than that of the only quantization and a signal level lower than the maximum limit allowed. Regarding the signal, in fact, it should be noted that often its amplitude does not reach the value FSR/2, for example because we desire to avoid running into the saturation phenomena of the device (clipping of the sinusoidal input). Consequently, the dynamics of the ADC is not fully exploited, resulting in a deterioration of the SNR.
For what concerns the noise, instead, in addition to the quantization noise, there is the noise of the internal electronics of the ADC, which, adding to the other, gives an increase in the overall noise and, finally, a further reduction in SNR.
All these effects can be quantified by calculating both the real SNR and the ENOB (*Effective Number of Bits*) which expresses the number of bits that an ideal ADC (only affected by the quantization noise) should have in order to produce at the output a sine wave (numerical) of a quality equal to that of the real we are examining.
In the first case, the signal and the noise may be obtained directly from the spectrum. These are expressed in power and in dBc, where the subscript stands for “carrier”, i.e. related to the power of a sine wave with the maximum amplitude allowed:

$$P_{0dBc} = \left(\frac{FSR/2}{\sqrt{2}}\right)^2 = 3.125V^2 = (1.77V)^2$$

For what concerns the noise, it must be considered that, having a flat spectrum (white noise), its power is divided into N/2 histograms, which are shown in the spectrum. In practice, in each histogram falls only a fraction equal to 2/N of the whole noise of the ADC. For this reason, the total noise is calculated by multiplying the average level of the background (said Noise Floor) by half the total number of samples, making sure to use consistent sizes:

$N_{real} = -100dBc \cdot 127800/2 = !!! = -100dBc + 10log(127800/2) = -52dBc$ We find $SNR_{real} = -20dBc - (-52dBc) = 32dB$. It is now possible to calculate the ENOB:

$$ENOB = \frac{SNR_{real} - 1.76}{6.02} = 5bit$$

b) The SFDR (*Spurious Free Dynamic Range*) is defined as the ratio between the power of the useful signal and the highest spectral contribution. From the graph, it is equal to -60dBc, from which $SFDR = -20dBc - (-60dBc) = 40dB$is found.
The THD (*Total Harmonic Distortion*), instead, gives account for the non-linearities introduced by the ADC and is calculated as the ratio between the power of the harmonics and that of the fundamental, as shown in the following expression:

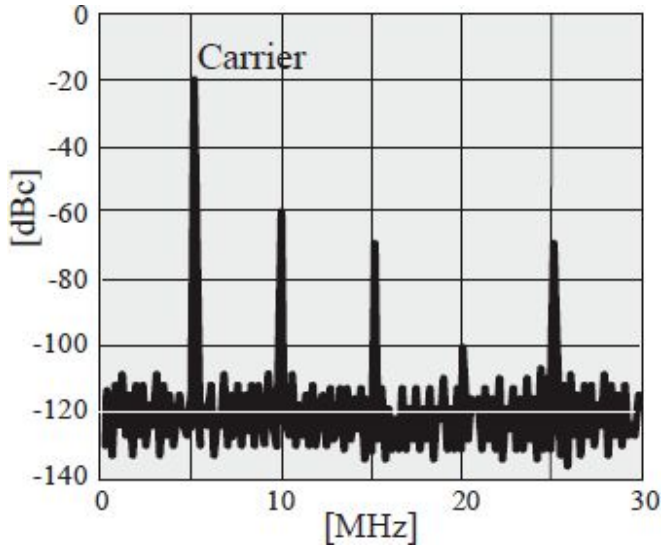
$$THD = \frac{10\log\left(\frac{3.125V^2 \cdot 10^{-80/10} + 3.125V^2 \cdot 10^{-70/10} + 3.125V^2 \cdot 10^{-60/10}}{3.125V^2}\right)}{-20dBc} = -40dB$$

Ex. 8

Unilateral spectrum of a *16bit* ADC, with *FSR=5V* and *bin-width* of *3kHz*.

a) Calculate SNR_{ideal} (maximum amplitude and only quantization noise), theoretical *noise-floor*, and effective values of the fundamental (f_0*=5MHz*) and of the second harmonic (at $2 \cdot f_0$).

b) Calculate $SNR_{theoretical}$ (actual amplitude, but only quantization) and $SNR_{measured}$ (actual amplitude and noise) and the resulting *ENOB*.



a) SNR_{ideal}=98dB, NF=-138dBc, fundamental power=31.3mV^2, second harmonic power=3.13μV^2.

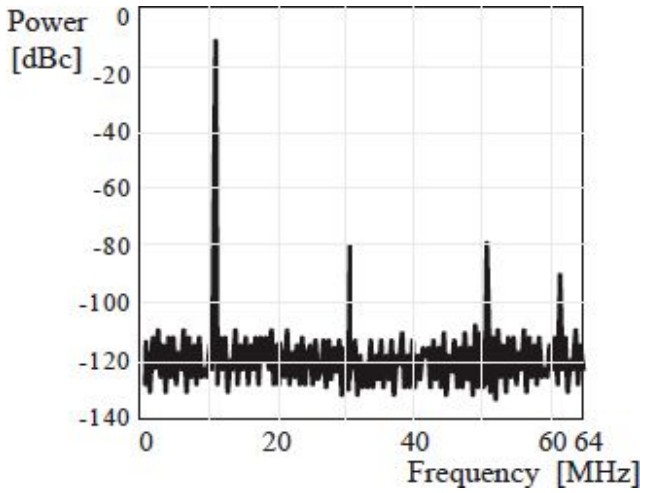
b) $SNR_{theoretical}$=78dB, $SNR_{measured}$=60dB, ENOB=10bit.

Ex. 9

A two-quadrant ADC has an FSR of *5V* (from *–2.5V* to *+2.5V*). The output spectrum shown on the right has a *bin-width=2kHz*.

a) Calculate the real *SNR* and estimate the number of bits that the ADC should have.

b) Calculate *SiNAD* and *SFDR*.



a) SNR = 62dB, ENOB = 10bit

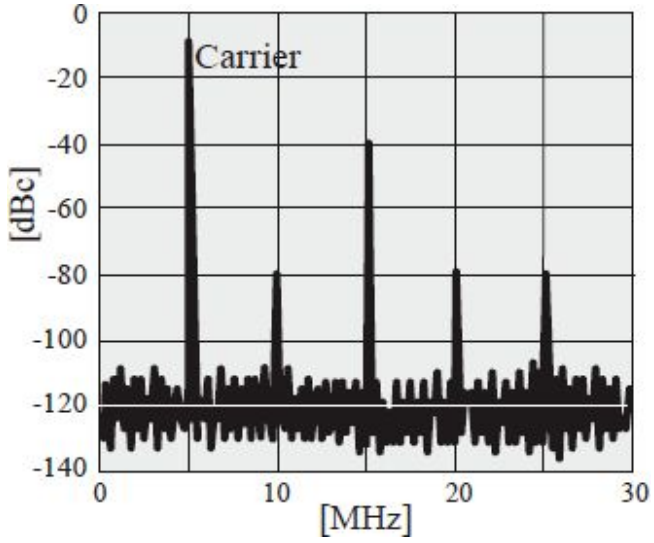
b) SiNAD = 61dB, SFDR=70dB

Ex. 10

Spectrum of an ADC with *16bit*, *FSR=5V*, and *bin-width=3kHz*.

a) Calculate SNR_{ideal} (maximum amplitude and only quantization noise), theoretical *noise-floor*, and effective values of the fundamental (f_0=*5MHz*) and of the third harmonic (at $3 \cdot f_0$).

b) Calculate $SNR_{theoretical}$ (actual amplitude with only quantization) and $SNR_{measured}$ (actual amplitude and noise) and the resulting *ENOB*.



a) SNR_{ideal}=98dB, NF=-128dBc, fundamental=313mV^2, third harmonic =31.3µV^2.

b) $SNR_{theoretical}$=88dB, $SNR_{measured}$=78dB, ENOB=12.7bit.

Ex. 11

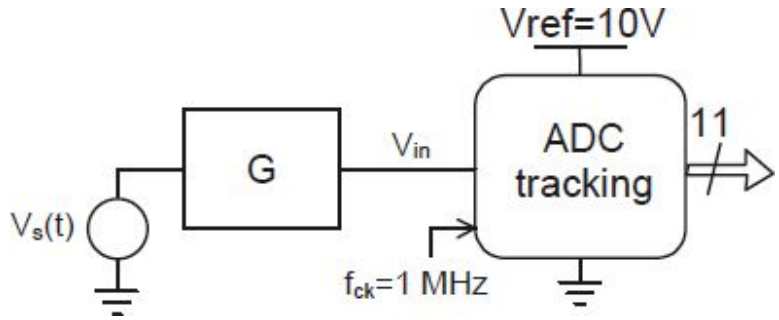
Consider the following circuit for the digital conversion of the input signal $V_s(t)$:

a) Assuming that the input signal $V_s(t)$ is contained in the range 0-1.2 V, find the minimum value of the gain G to ensure resolution of the input signal better than 0.1%.

b) Draw the block scheme of the A/D tracking converter and describe its functioning.

c) If at the output of the block G there was a square wave with amplitude of 10 V and a frequency of 1 kHz, would the ADC be able to hook it? Justify your answer.

d) If the output of the block G were $V_{in}(t)=5V \sin(2\pi f\, t) + 5V$, what would the maximum frequency of the sine wave that the ADC is able to hook be?



a) The resolution of the ADC, referred to the input of the ADC itself, is LSB=10V/(2^{11})=4.88mV. If the resolution is instead referred to the input where $V_S(t)$ is applied, it is LSB/G. According with the specification of this exercise, the resolution at the input is better than 0.1% of the signal, i.e. LSB/G_{min}<0.1%·1.2V=1.2V/1000=1.2mV, from which G_{min}=4.069 is found.

c) The staircase of the ADC has the following trend. Considering the height and length of the steps, the maximum average slope (i.e. the slope of an ideal line which connects the beginning and the ending of the stairs) is equal to

$P=V_{ref}/(2^n)\cdot f_{CK}=4.88\cdot10^{-3}V/\mu s$. The square wave has a half period of ½·1/1kHz=500μs. In 500μs the ADC, with a staircase with a slope P, P·500μs=2.44V is reached, not 10V. Then, the signal is not hooked.

d) The maximum frequency of the sine wave that the ADC is able to hook is that with a maximum slope lower than the maximum slope P of the staircase that the ADC can produce. Given a sine wave $V_{in}(t)=5V\cdot\sin(2\pi\cdot f\cdot t)+5V$, the maximum slope is $5V\cdot2\pi\cdot f$. Therefore, we must impose $5V\cdot2\pi\cdot f_{max}<4.88\cdot10^{-3}V/\mu s$, from which $f_{max}=155Hz$ is found.

9

Advanced ADCs

“Gazing through the window at the world outside
Wondering will mother earth survive
Hoping that mankind will stop abusing her sometime
After all there’s only just the two of us
And here we are still fighting for our lives
Watching all of history repeat itself
Time after time
I’m just a dreamer
I dream my life away
I’m just a dreamer
Who dreams of better days
I watch the sun go down like everyone of us
I’m hoping that the dawn will bring a sign
A better place for those Who will come after us ...
This time
Your higher power may be God or Jesus Christ
It doesn’t really matter much to me
Without each others help there ain’t no hope for us
I’m living in a dream of fantasy
Oh yeah, yeah, yeah

“Dreamer”, Ozzy Osbourne

9.1. INTERPOLATING FLASH

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.

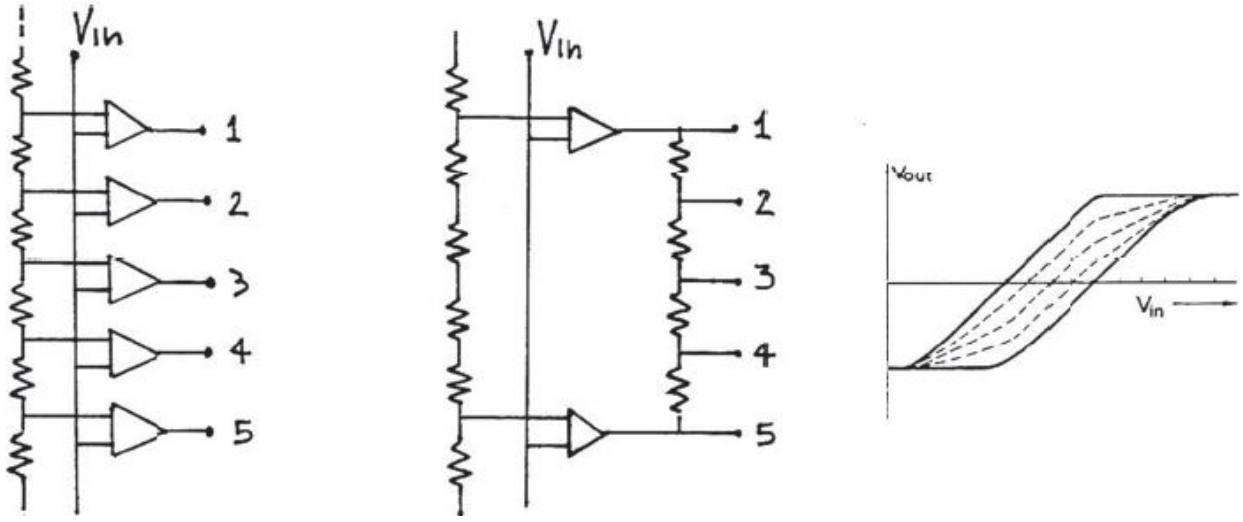


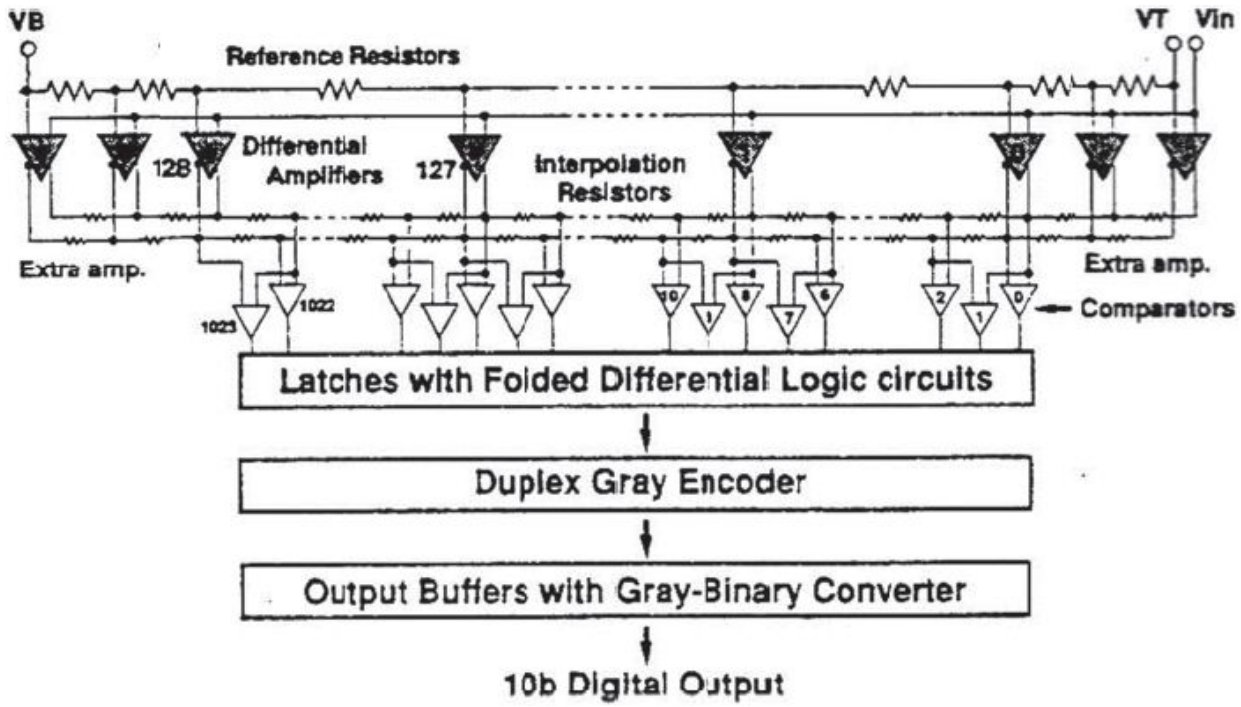
Fig. 9.1: Principle scheme for a classical flash ADC (on the left) and interpolation ADC (in the center). The output resistive partition effectively replaces the three comparators characteristics (on the right).

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.


VB
VT
Vin
Reference Resistors
128
Differential Amplifiers
127
Interpolation Resistors
Extra amp.
Extra amp.
1023
1022
10
9
8
7
6
2
1
0
Comparators
Latches with Folded Differential Logic circuits
Duplex Gray Encoder
Output Buffers with Gray-Binary Converter
10b Digital Output

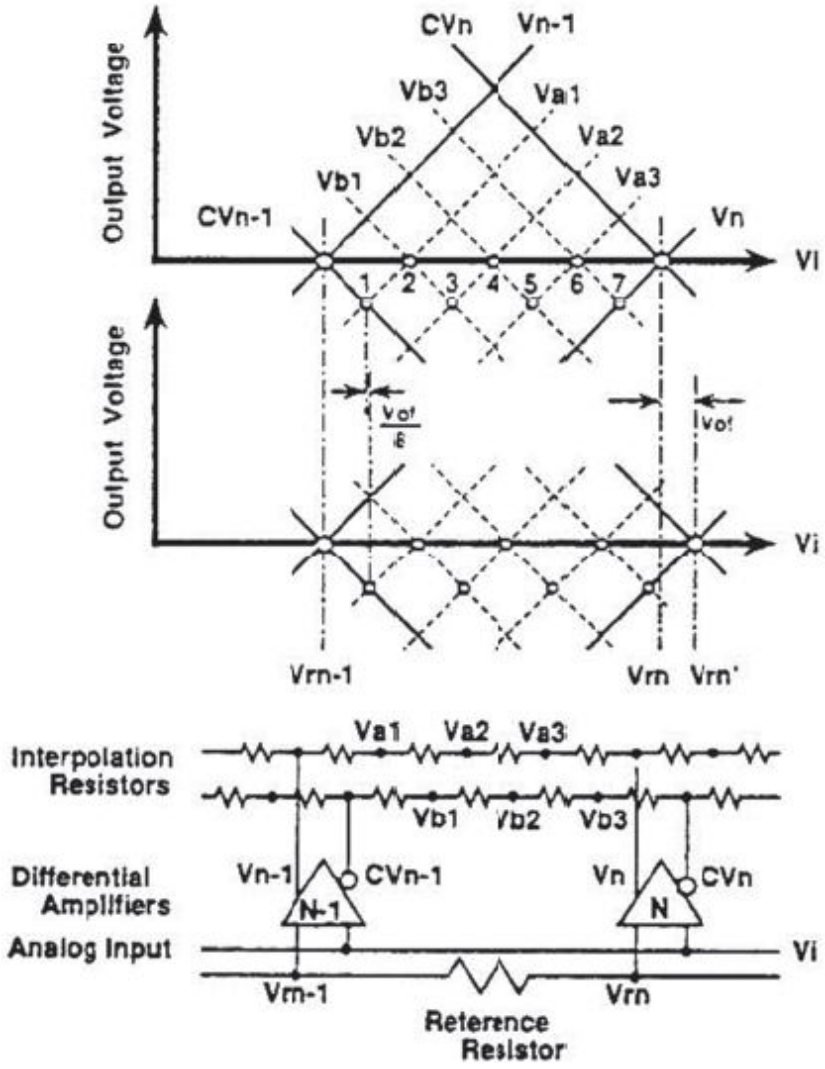


Fig. 9.2: Architecture of an interpolated-parallel flash ADC.

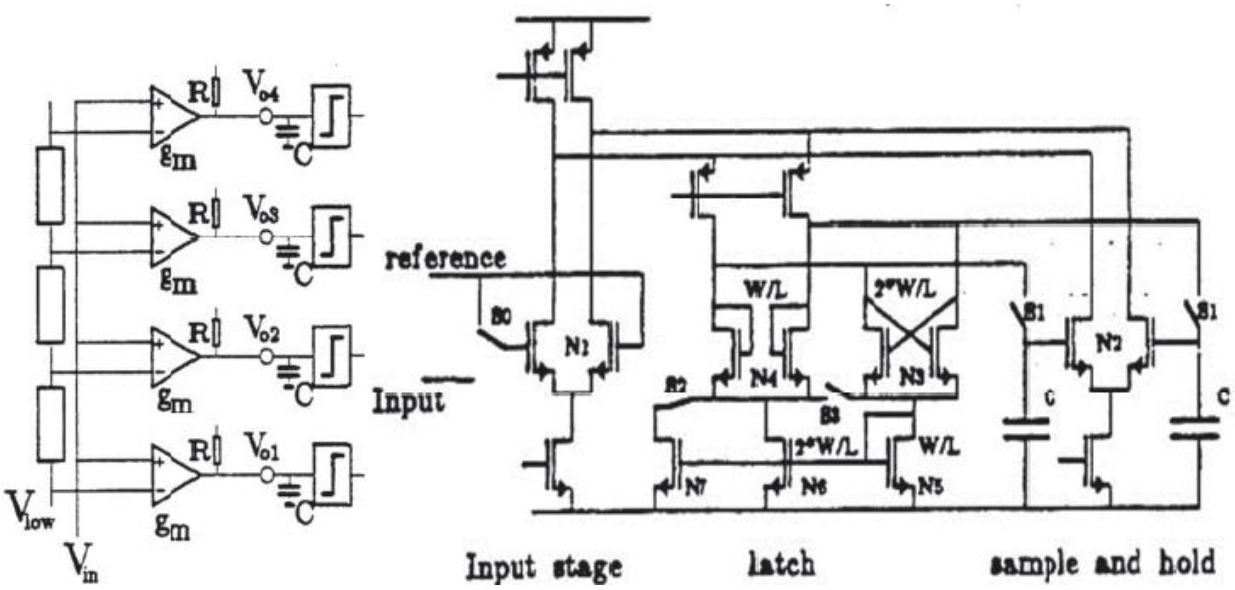


Fig. 9.3: Comparator offset reduction, thanks to the presence of the OpAmps upstream (on the left) and by means of an appropriate alternation in the phase within the comparator (on the right).

In this respect, consider Fig. 9.3. The comparators are considered ideal, except for the presence of capacitance C and a random offset source V_{oi} with an r.m.s. value σ_{ov}, which depend on the transistor geometry. The input OpAmp (trans-conductance in this case) allows reducing the input offset in accordance with the following expression:

$$\sigma=\sigma_{ov}\cdot(1+j\omega RC)/(g_m\cdot R)$$

For example, with a 1μm CMOS technology with $\sigma_{ov}\approx 20mV$ and $g_m/(2\pi C)\approx 250MHz$, we obtain $\sigma=0.8mV$ at 10Msps.

Fig. 9.3 shows a possible scheme for the comparator which allows reducing the offset. It is an input differential stage N_1, through which the signal reaches the S&H stage constituted by N_2, S_1, and C. The working principle is divided into three phases. During the sampling phase, S_1 is closed, and the voltage $V_{in}-V_{ref}+V_{ov}$ is stored on both the capacitors. During the amplifying phase (shown in the figure), the switch S_0 connects the input transistors to the reference voltage: S_1 is open, and the S&H stage generates a current proportional to the voltage across the capacitors while the input stage produces a current proportional to V_{ov}. In this way, the difference

between the two currents has no offset. In the last phase, S_3 conduces, activating the latch, whose "decision" depends only on the difference V_{in}–V_{ref}. The actual offset is 0.4mV.

The scheme just seen does not effectively act on the input capacitance C (shown in Fig. 9.3): greater capacitance requires, in fact, the use of an equivalent resistor and then the introduction of a buffer stage with high performance at the input, which can greatly penalize the power consumption.

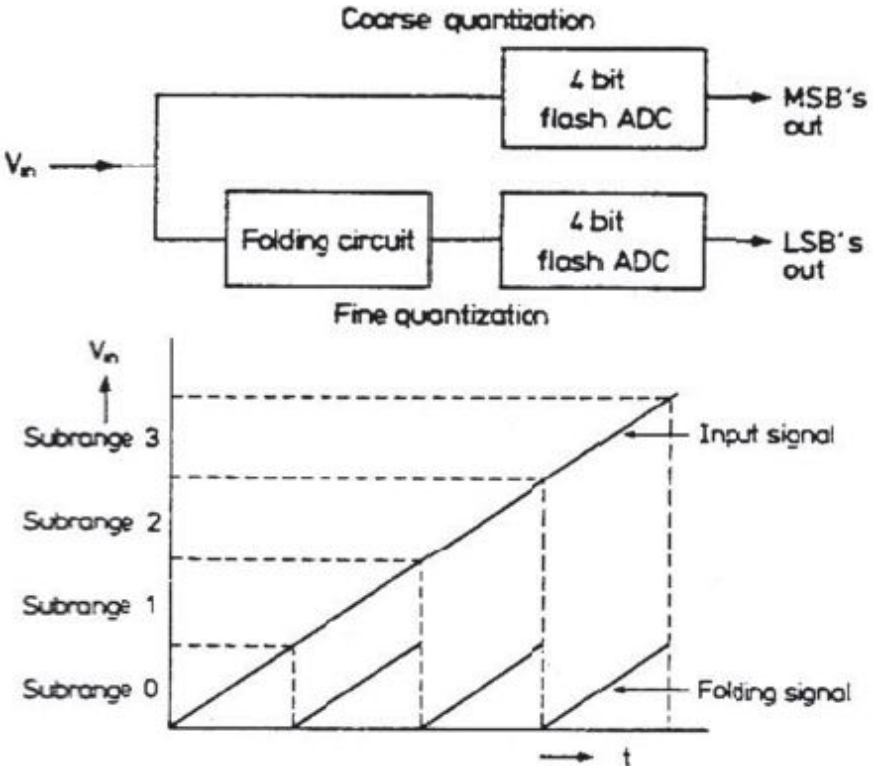


Fig. 9.4: Working principle for a folding ADC.

9.2. FLASH FOLDING

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.

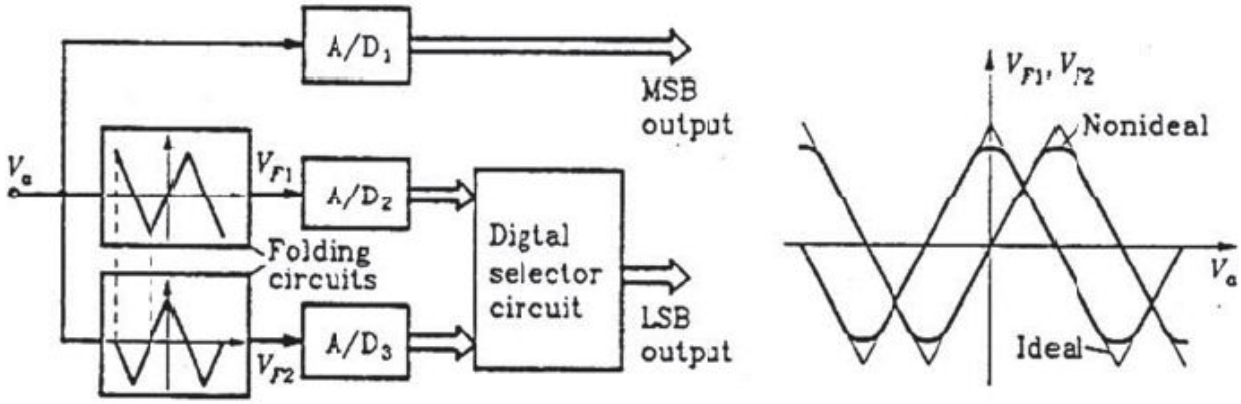


Fig. 9.5: Implementation of the folding by means of a triangular signal, instead of the saw tooth.

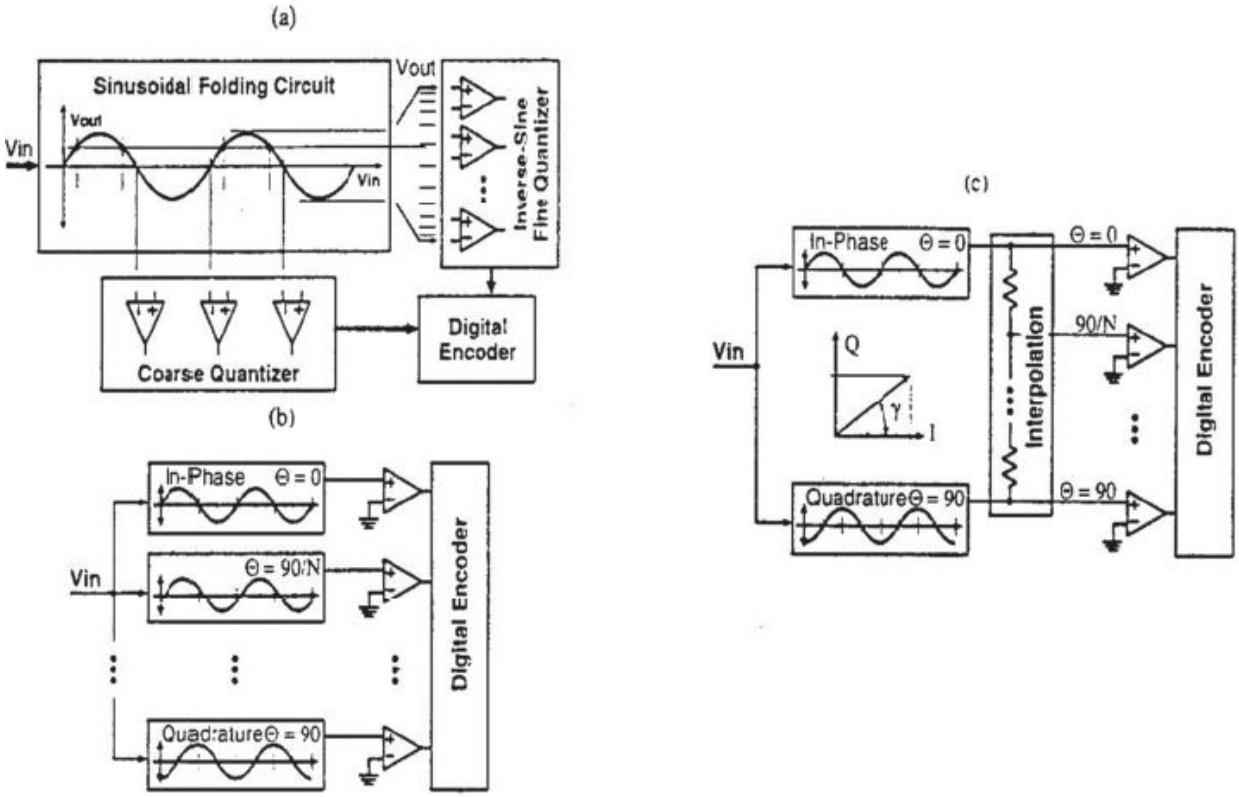


Fig. 9.6: ADC with sinusoidal folding (a), with a folding circuit for every level (b), and interpolating folding (c).

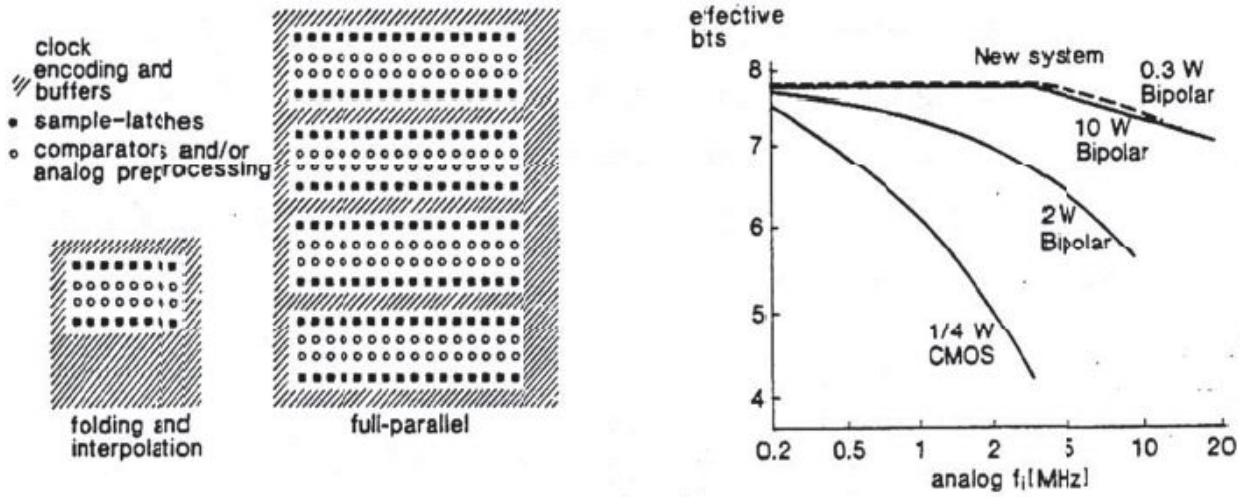


Fig. 9.7: Advantages in terms of occupied area (on the left) and power dissipation (on the right) for the flash ADC with interpolation and folding of Fig. 9.6c.

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 $90^o/N$. 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 with 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.

9.3. Half-Flash Converter

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.

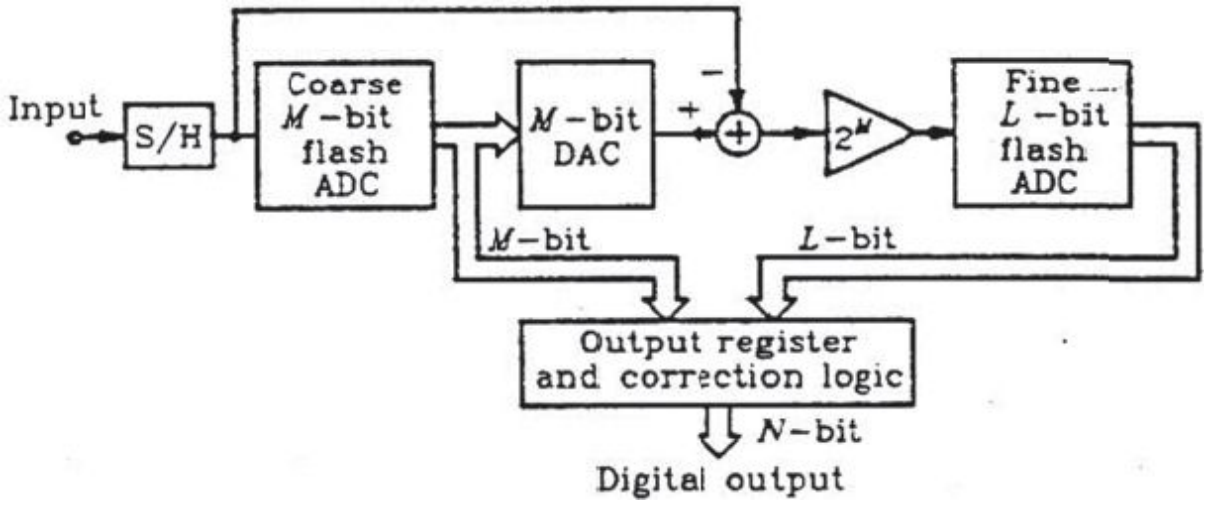


Fig. 9.8: Half-flash ADC structure.

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.

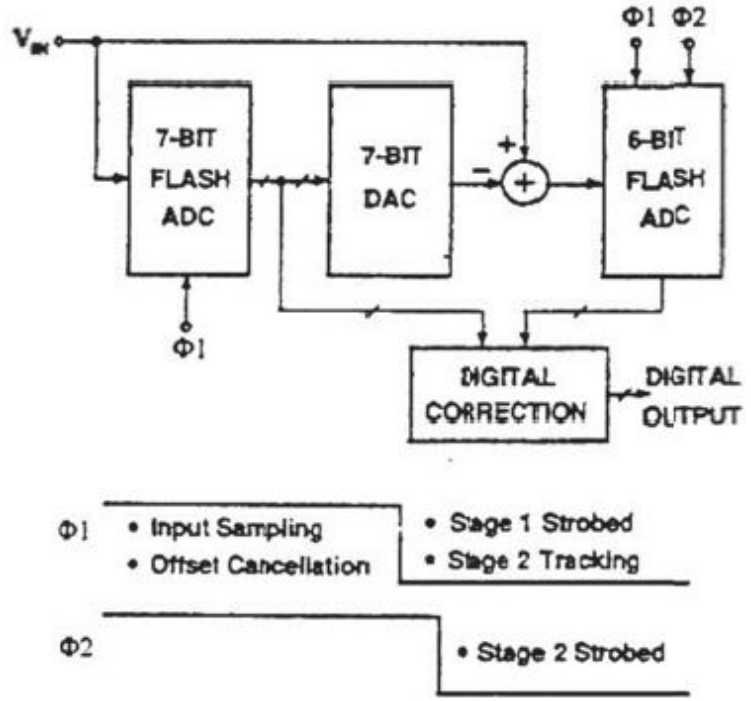


Fig. 9.9: 12 bits half-flash ADC with internal temporizations.

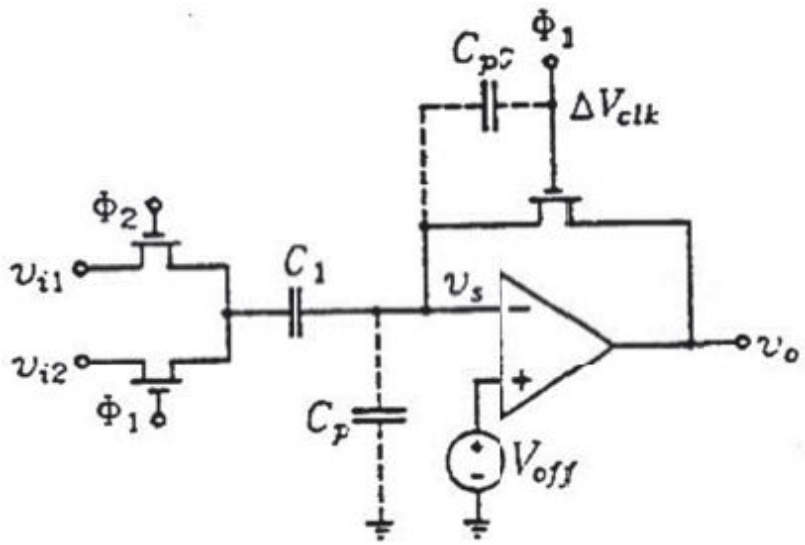


Fig. 9.10: Basic structure of a switched capacitance CMOS comparator.

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}.

Fig. 9.11 shows the comparators used in the first stage of the ADC, in which the offset reduction method (named auto-zeroing) is the same as that depicted in Fig. 9.10. In this case, the comparator structure is fully-differential, and, for this reason, at the input, we compare not only $+V_{in}$ with $+V_{ref}$, but also $-V_{in}$ with $-V_{ref}$; and the outputs are two because there is also the complementary logic value. This architecture, more complex than the single-ended one, allows minimizing some problems of the MOS technology, first of all, the charge injection effect due to the parasitic capacitance of the switches. The comparator obtained with this scheme dissipates 0.55mW and presents an offset of 5mV.

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.

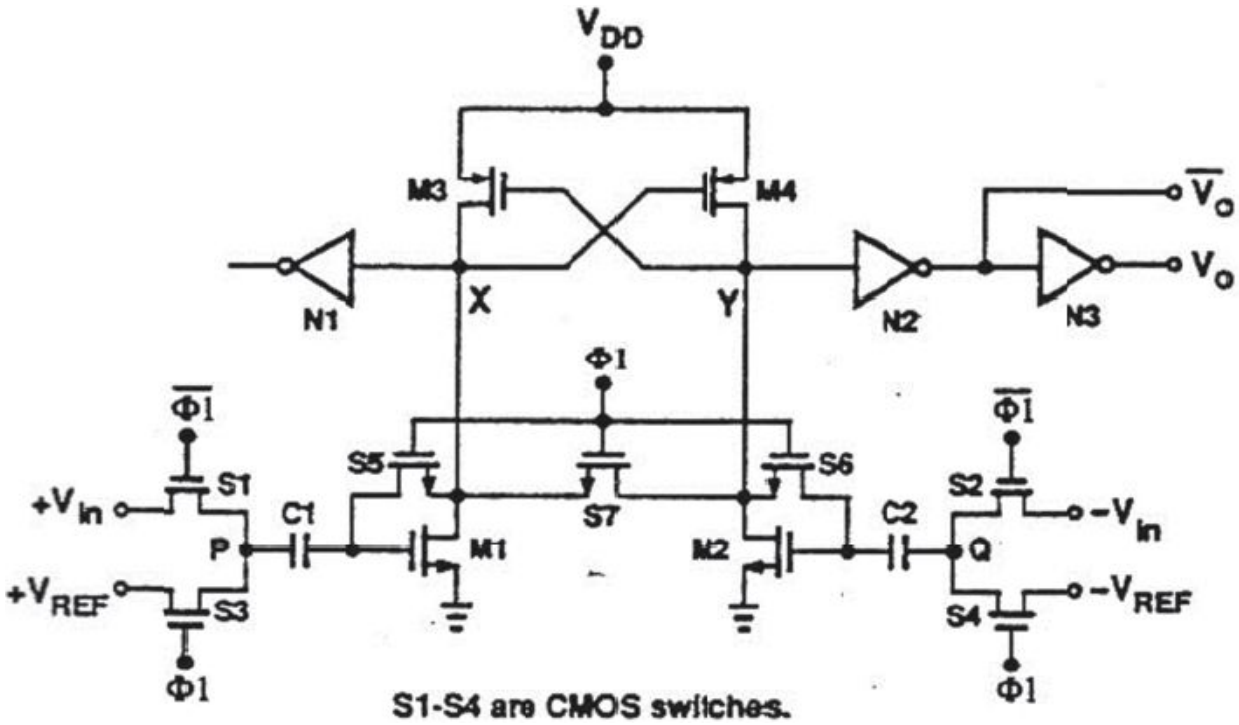


Fig. 9.11: Another implementation of a switching capacitance comparator.

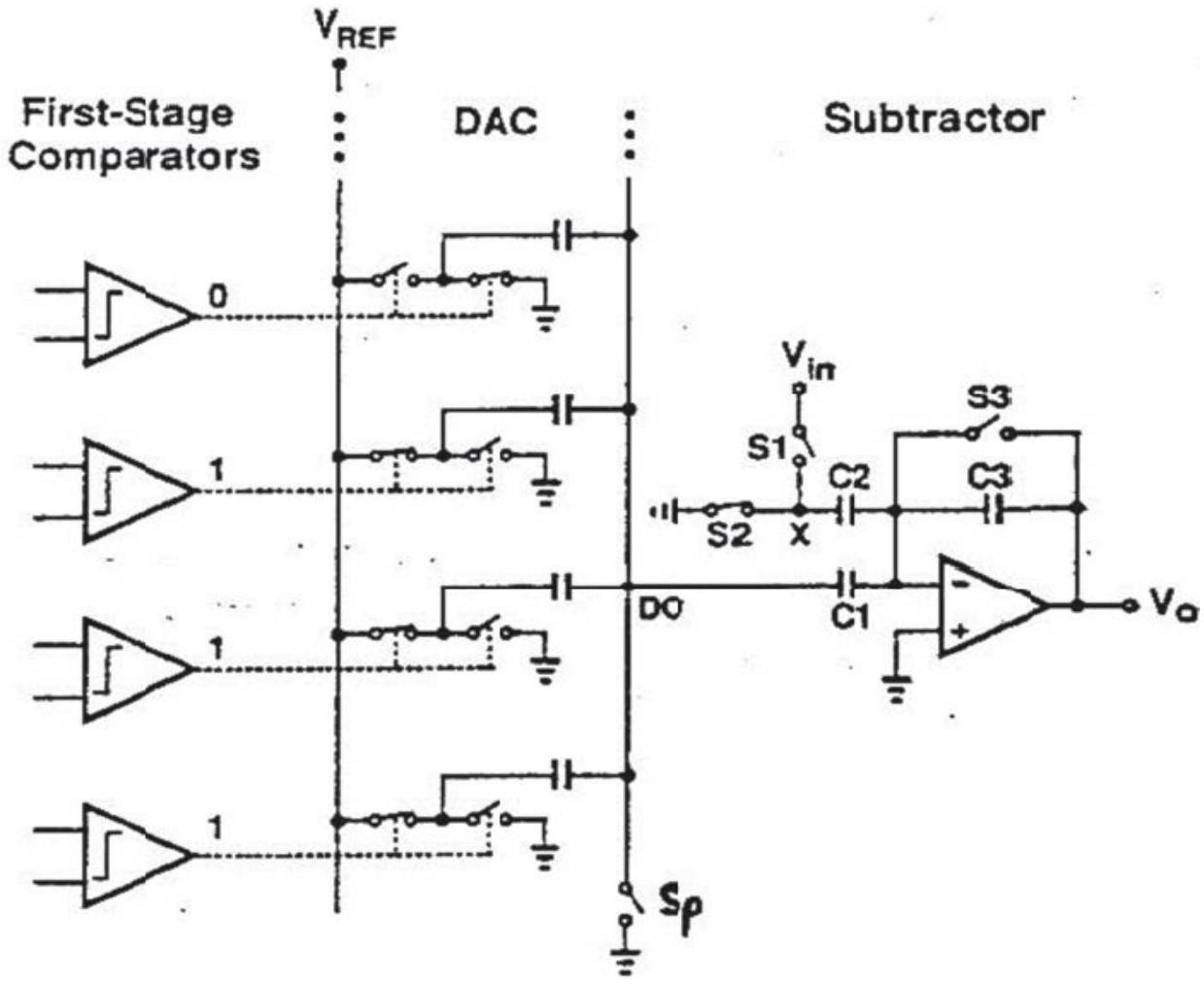


Fig. 9.12: Implementation of a part of a half-flash ADC.

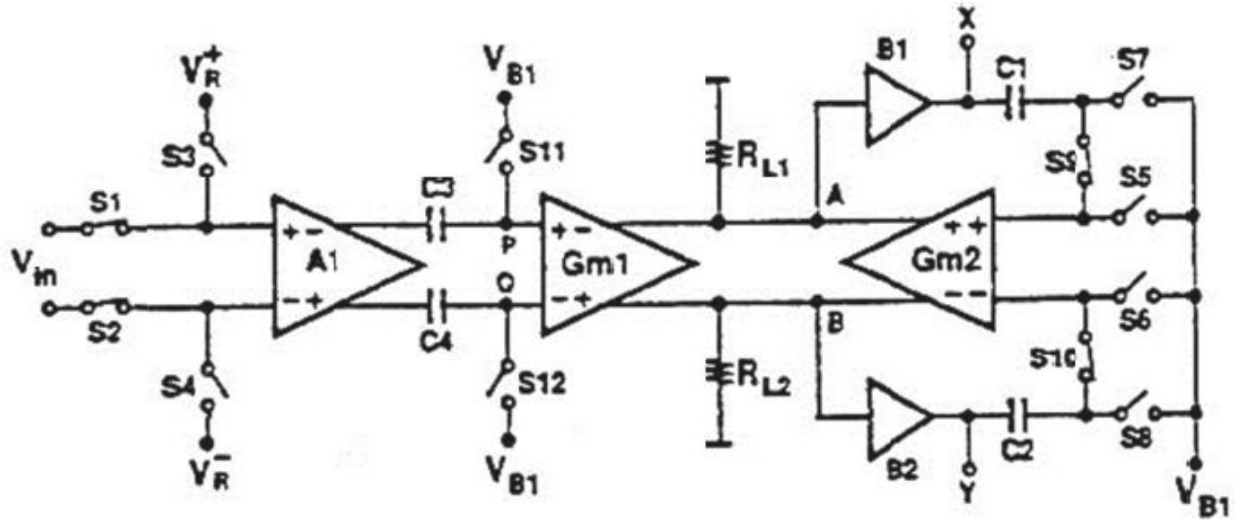


Fig. 9.13: Internal structure of the second stage comparators of the half-flash ADC.

About the second stage, the comparators should have a very low offset because the residual is not amplified by the adder. The scheme of these comparators, shown in Fig. 9.13, is more complex than that of the first stage comparators. The offset obtained is 0.3mV with consumption of 1.7mW.

We talked about the error correction: this technique allows the relaxation of the precision required for the first stage comparators. In fact, the first stage is studied to reach a very high conversion frequency, which means that it cannot have high precision. In this example, n overlaps of 1bit between two stages is used in order to ensure that the first conversion residual (i.e. the difference between the input signal and the analog reconstruction of the first stage coded bits) is within the second stage input range. In other words, the second flash range is greater than the ideal one. If the first flash were ideal, the residual would be limited between 0 and 32 LSB, which means that it should be coded by a second stage with 5bit. As shown in Fig. 9.14, if the j-th first stage comparator has an offset of $\Delta V_j(<0)$, and the (j+1)-th has an offset of $\Delta V_{j+1}(>0)$, then the residual varies between ΔV_j and (32 LSB)$\cdot\Delta V_{j+1}$. Therefore, a range of 32 LSBs is not enough to codify the residual.

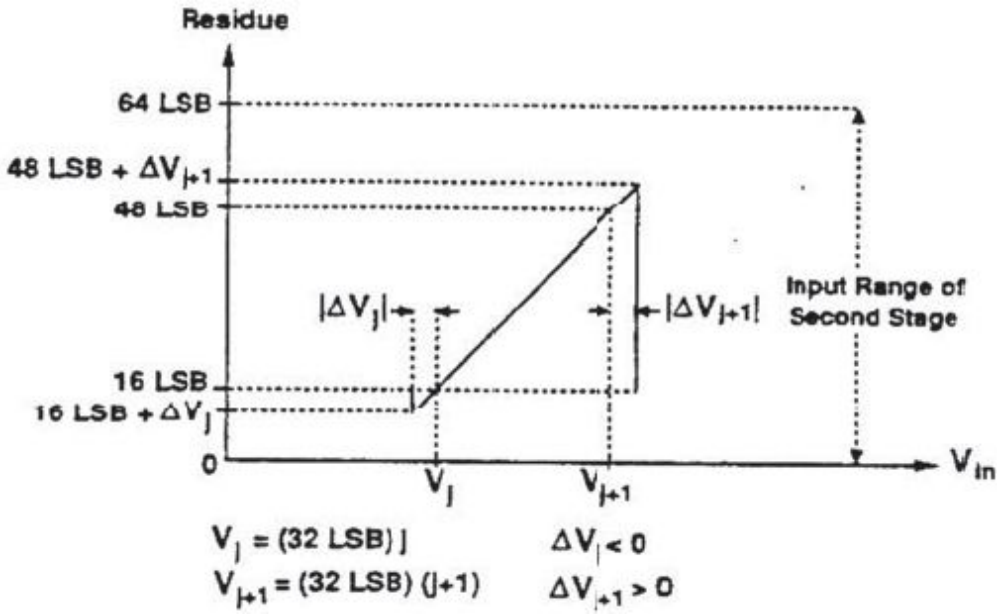


Fig. 9.14: Range composition for the half-flash conversion.

Instead, using a 6bits second stage, the residual can vary between 0 and 64 LSBs, which means that the comparator offset can be equal to ±16 LSB (in this case around 40mV). In order to have a residual within the second stage range, it must be shifted by 16 LSBs (Fig. 9.14); this can be obtained by shifting down the logic signal reconstructed by the DAC.

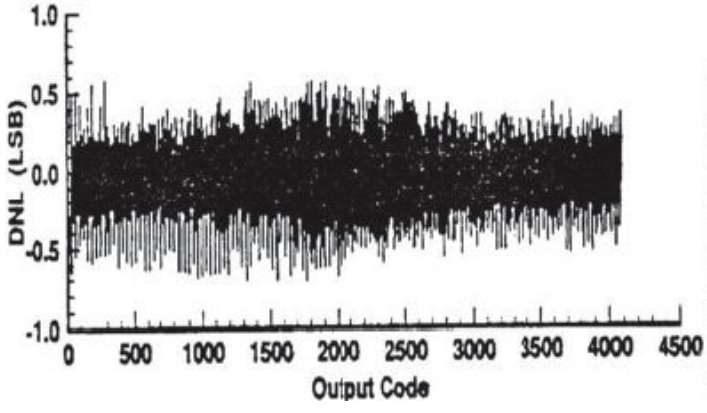
At this point, the 12bits output code is obtained by merging the 7 MSBs with the 6 LSBs. If X1 is the analog value associated with the MSBs, represented by the code $a_{11}…a_5$, the value normalized to the LSB is: $X1=a_{11}\cdot 2^{11}+…+a_5\cdot 2^5$. Similarly, the 6 LSBs with $b_5…b_0$ have the analog value: $X2=b_5\cdot 2^5+…+b_0\cdot 2^0$.

Because the DAC does not produce the value X1, but X1–16 LSB, X1 must be shifted by the same quantity. The obtained number is added to X2. The final value is:

$$Y=(a_{11}\cdot 2^{11}+…+a_5\cdot 2^5)–16+(b_5\cdot 2^5+…+b_0\cdot 2^0)$$

Therefore, the digital correction algorithm includes the following steps: shift $a_{11}…a_5$ towards left by 5bit; subtract the binary number 10000; add $b_5…b_0$.

Fig. 9.15 (on the left) shows the DNL of the ADC obtained by the static test. A sinusoid with amplitude equal to the FSR of the ADC and a frequency equal to 5kHz is sampled at 5Msps, setting the number of times during which the codes are present. The obtained histogram is normalized taking into account the “U-shape” of the difference in the potential of the sinusoid. From the normalized histogram, we obtained the DNL. Fig. 9.15 (on the right) shows a summary of the converter performance.



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

Fig. 9.15: DNL measurement (on the left) of the ADC, performed through the static test and its performance (on the right).

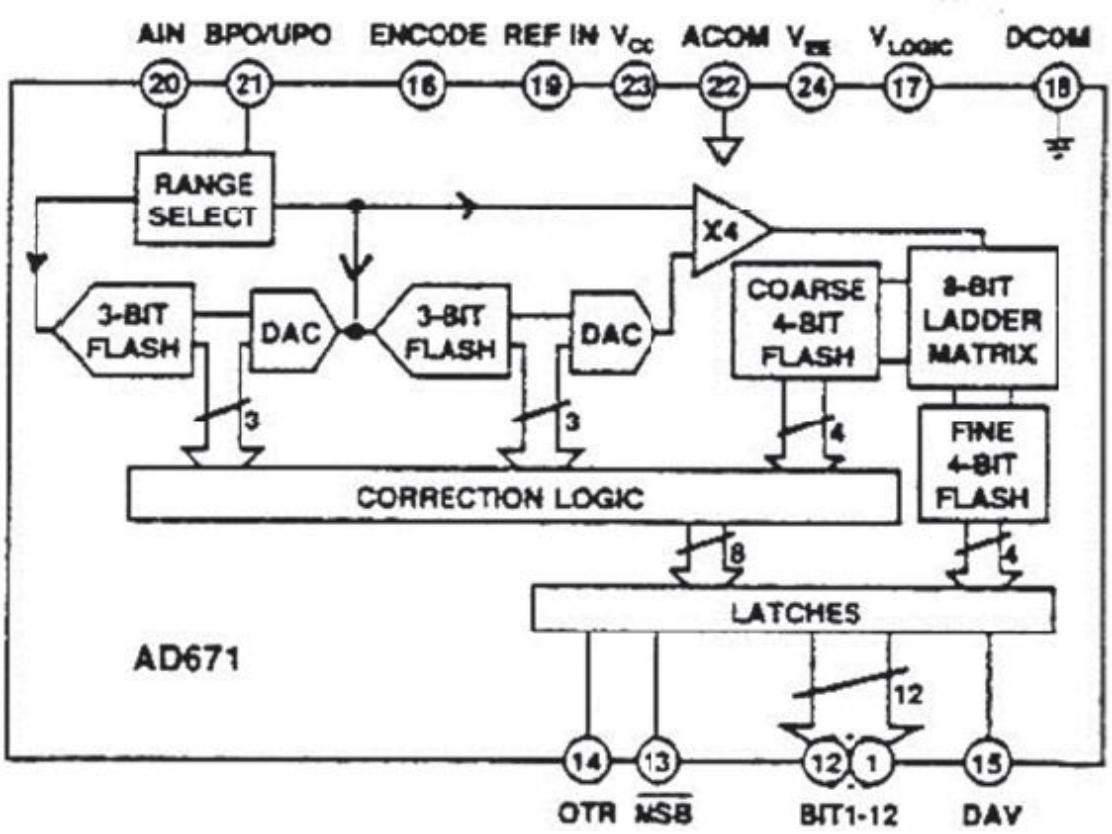


Fig. 9.16: Multistep flash ADC (AD671).

9.1 Multistep flash

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.

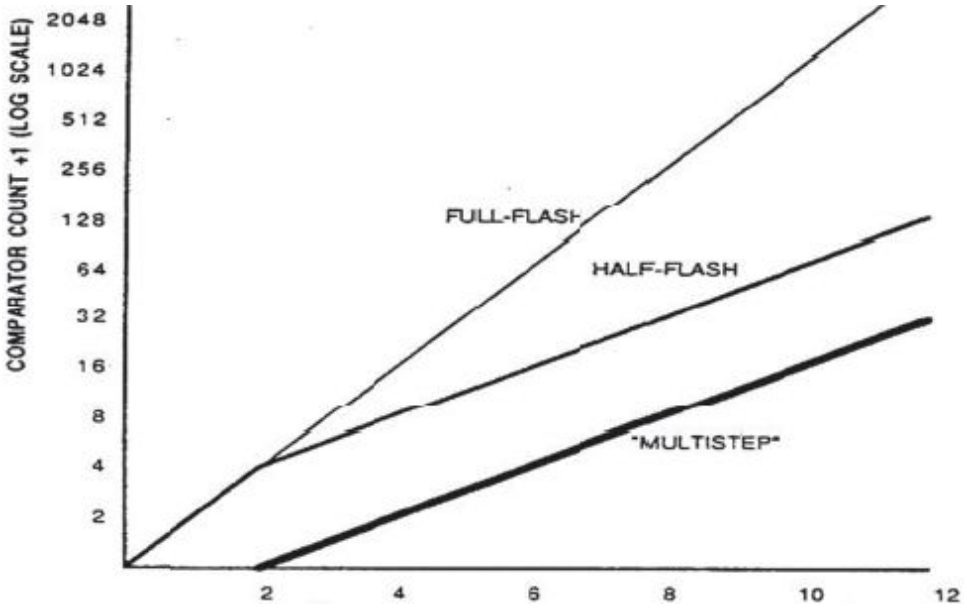


Fig. 9.17: Comparison of performance between some flash ADCs.

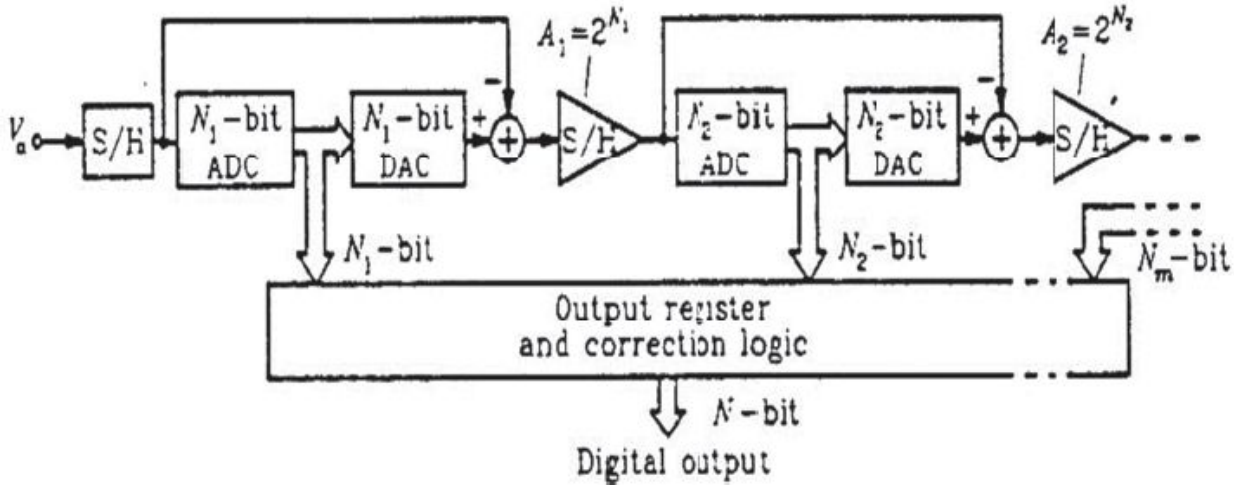


Fig. 9.18: Pipeline ADC converter architecture.

9.4. PIPELINE CONVERTERS

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.

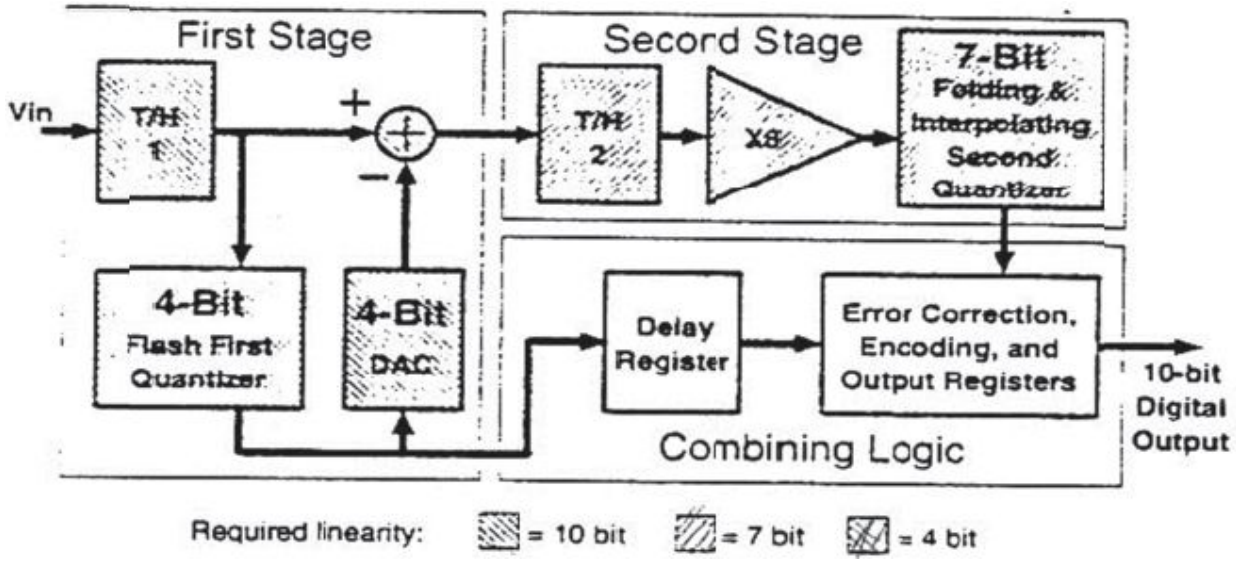


Fig. 9.19: 10bits two stages pipelined ADC example.

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.

Fig. 9.19 shows an example of two stages pipelined ADC with 10bits, obtained by means of two stages with 4 and 7bits, respectively. Such a choice is due to the fact that the second stage uses folding and interpolation techniques, which make the circuitry complexity of the second stage similar to the first one, thanks to the consequent layout saving. If we use, instead, two traditional flash ADCs, a choice can be 6bits for the first and 5bits for the second, or vice-versa. From the figure, we can note that the T&H1 and the DAC must have a linearity of at least 10bits while the first flash requires a linearity of only 4bits, thanks to the error correction.

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.

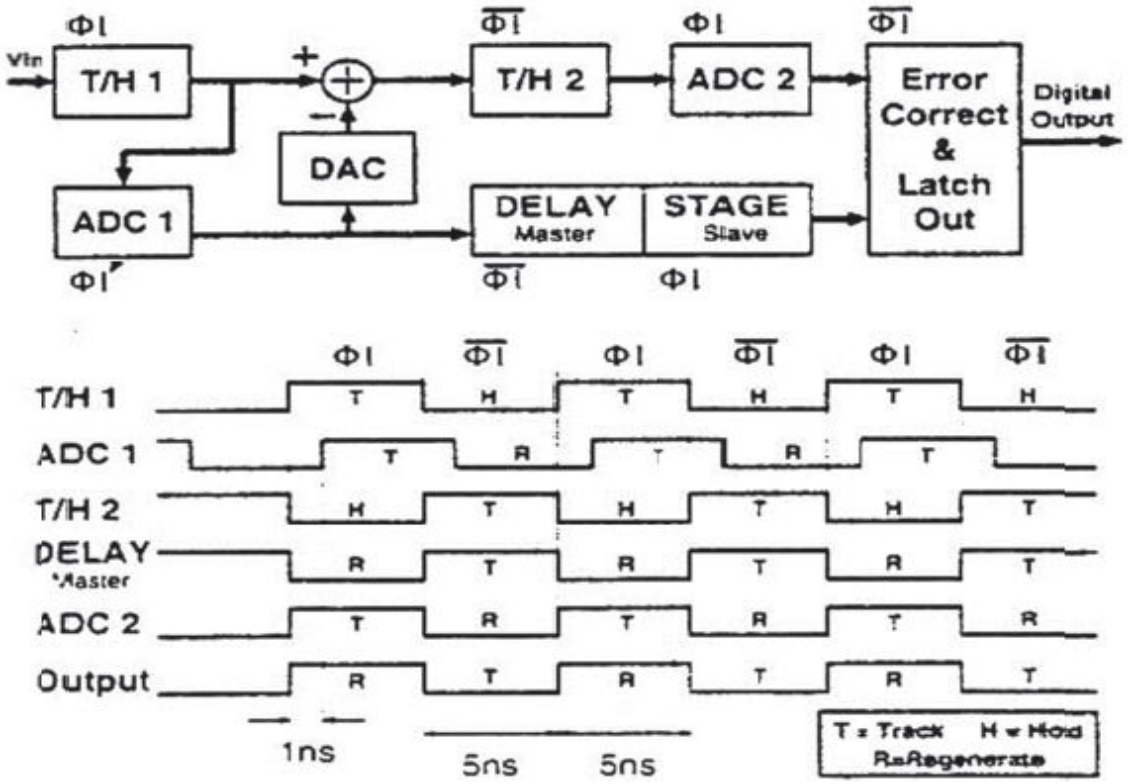


Fig. 9.20: Pipelined ADC timing graph.

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:

$$V(i+1)=2\cdot[V(i)-b_i\cdot V_r] \text{ for } i=1, 2, \ldots N$$

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.

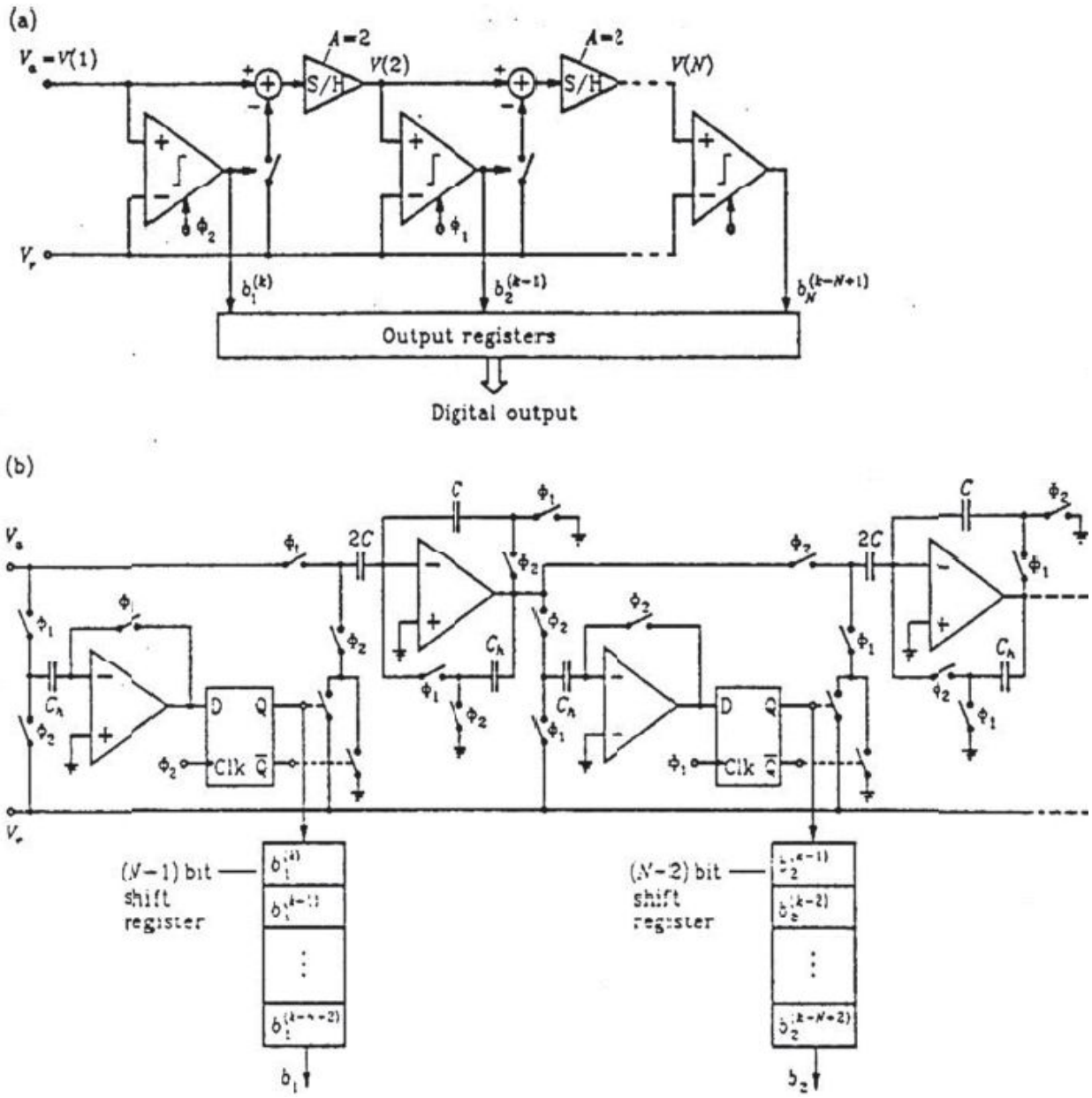


Fig. 9.21: Pipelined ADC with N bits made with N 1bit stages (a) and its implementation in CMOS technology with switched capacitors (b).

Error sources		Problems	Solutions
S/H	Offset Nonlinearity Gain error	DC Offset Nonlinearity & missing code	Input-referred offset Negligible Calibration
ADC	Offset Nonlinearity Gain error	Nonlinearity & missing code	Digital error correction
DAC	Offset Nonlinearity Gain error	DC offset Nonlinearity & missing code	Input-referred offset Calibration
Amp	Offset Nonlinearity Gain error	DC offset Nonlinearity & missing code	Offset cancellation Negligible Calibration

Fig. 9.22: Main causes of errors in a pipelined ADC.

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.

9.5. Successive Approximations Converters

The Successive Approximations (SAR) ADCs are largely used because they allow obtaining good performance levels with low costs. They form a great part of the ADC market, and the application field is, typically, for resolutions of 10÷16bits and a conversion time in the range of μs.

The structure for a SAR ADC is formed by the following blocks: an S&H, a comparator, a DAC, and digital control logic named the *successive approximations registry* (SAR). These blocks are connected with a feedback

configuration, as shown in Fig. 9.23. The S&H block maintains the analog signal V_a during the conversion process, which proceeds in the following way: after the reset, the SAR brings the DAC output to the value $V_r/2$, where V_r is $(V_a)_{max}$. This means that the first step consists of imposing the MSB to the value 1 and the others bits to 0. If V_a is greater than $V_r/2$, then b_1 remains equal to the value 1; otherwise, it is brought to 0. In the following step, the SAR imposes the DAC output to the value $MSB\cdot(V_r/2)+V_r/4$, where MSB is 1 or 0. This voltage is newly compared with V_a, and b_2 is established on the basis of the converter response. The process continues until all the N bits are determined with successive approximations.

This technique requires N clock cycles to convert N bits and is intrinsically slower than that described in the preceding paragraphs. However, the extreme simplicity of the structure and, particularly, the use of a unique comparator allow obtaining good resolution with moderate power consumption and area occupation.

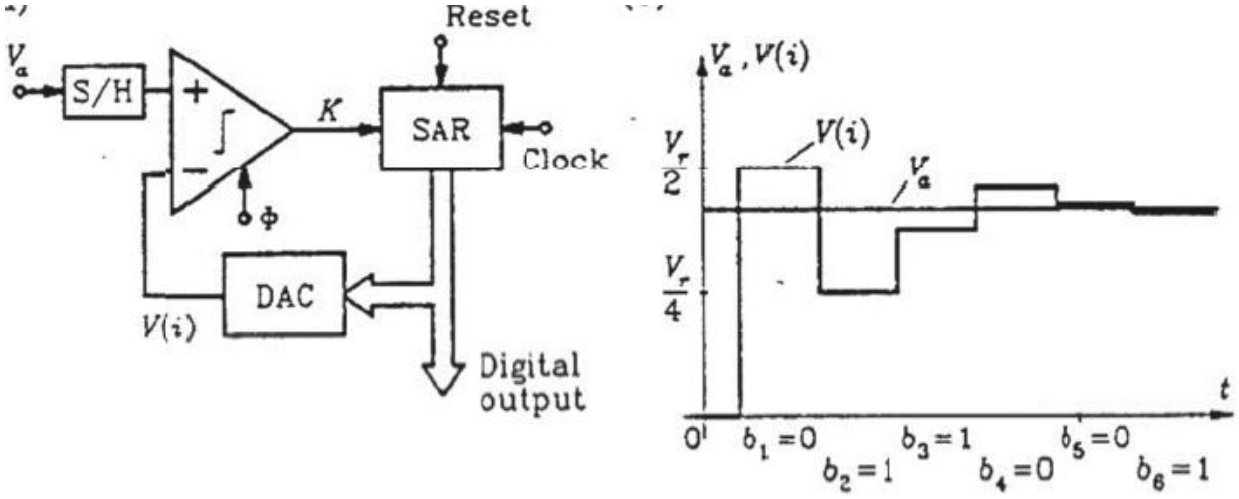


Fig. 9.23: Block scheme for a SAR ADC.

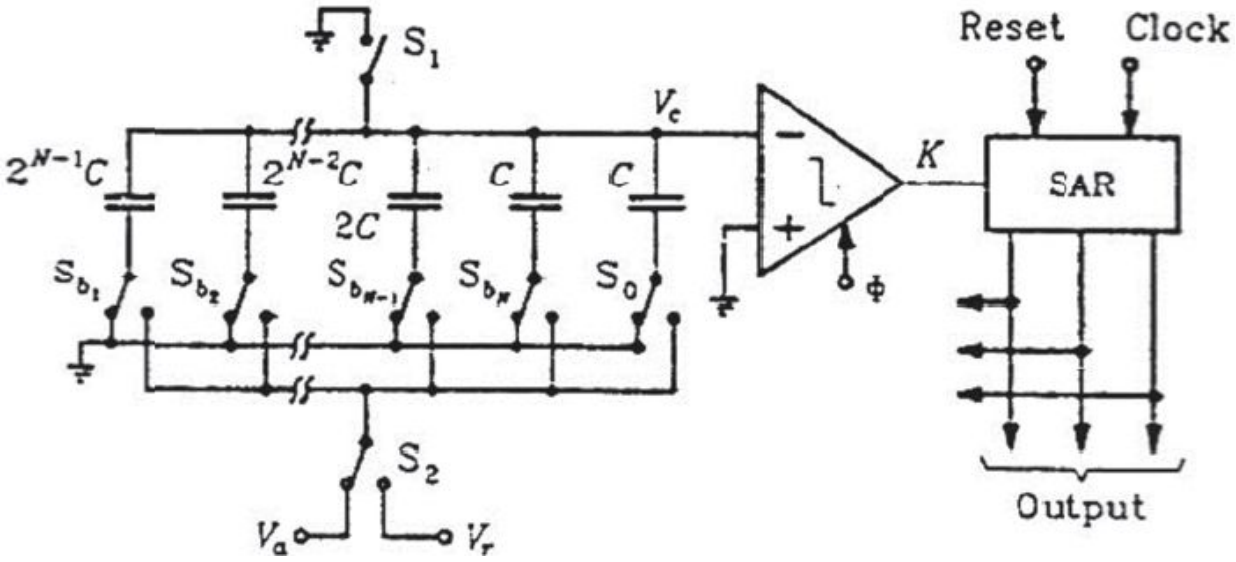


Fig. 9.24: Simplified structure of a charge redistribution ADC.

9.1 Charge redistribution ADC

A variant of the SAR philosophy is the *charge redistribution* technique, which can be easily made with MOS technology. A simplified scheme for a charge redistribution ADC is shown in Fig. 9.24 and consists of an array of capacitors with growing capacitance, a comparator, and an SAR. The capacitor array is a charge redistribution DAC. The conversion process is made of three phases: sampling, hold, and redistribution/comparison. During the sampling phase, the upper plates of the capacitors are connected to the ground towards the switch S_1 while the lower are connected to the input V_a by means of the switch S_2 (the other switches are connected to the bus connected to S_2). The total charge stored in the upper plates is: $Q_a=-2^N \cdot C \cdot V_a$

During the hold phase, S_1 is open while the switches S_{b1}, S_{b2}, … , S_{bN} are connected to the ground. In this way, the comparator's input voltage becomes $V_c=-V_a$. Finally, in the redistribution/comparison phase, which requires N steps, we can obtain the bits (one for each step). In the first step, the upper plate of the greater capacitor ($2^{N-1} \cdot C$) is connected to the reference voltage V_r towards the switches S_2 and S_{b1} while those of the other capacitors is connected to the ground. In this way, in respect of the reference voltage, the capacitors network works as a voltage divider 2:1. Therefore, the comparator's input voltage becomes $V_c=-V_a+V_r/2$. The comparator compares this voltage with zero and determines the *most significant bit*, b_1. In fact, if $V_a>V_r/2$, $V_c<0$ and b1=1 hold while, if $V_a<V_r/2$, $V_c>0$ and b1=0 hold. If b1=0, the switch S_{b1} remains connected to V_r, otherwise is grounded.

During the second step, the capacitor with capacitance $2^{N-2} \cdot C$ is connected to V_r, and the comparator's input voltage becomes $V_c=-V_a+(b1 \cdot V_r/2)+V_r/4$. Comparing this voltage with zero, we can determine the bit b_2. The process continues until all the bits are obtained. With this principle, we can obtain converters with resolution in the range 10÷16bit, conversion times in the order of μs, and power consumption limited to 10÷20mW.

A great advantage of this technique is that the S&H stage is not required anymore because the working principle just described involves also the sampling and the storage of the signal. A drawback is the requirement of a great range of various capacitance values. The maximum ratio between capacitance values exponentially grows, in fact, with the resolution. A partial solution to this problem can be the use of a DAC with two or three stages, obtained by means of a combined use of resistors and capacitors.

The accuracy obtainable by a charge redistribution converter depends mainly on the precision reached in the capacitance construction. This is closely linked to the conversion speed for which the ADC is designed. In fact, to obtain high speed, we need to use capacitors with smaller capacitance. The percentage error made in the realization of a capacitor grows with decreasing capacitance. To obtain high speed, we therefore need to adopt calibration techniques.

9.6. Time-Interleaved Converters

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).

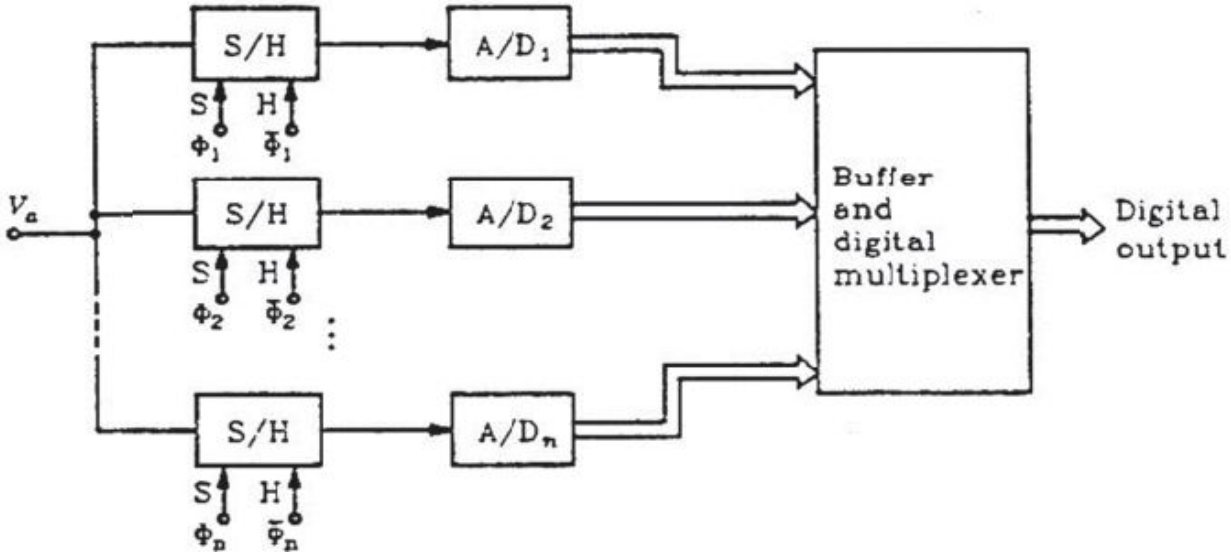


Fig. 9.25: Time-interleaved converter architecture constituted by many ADCs connected in parallel.

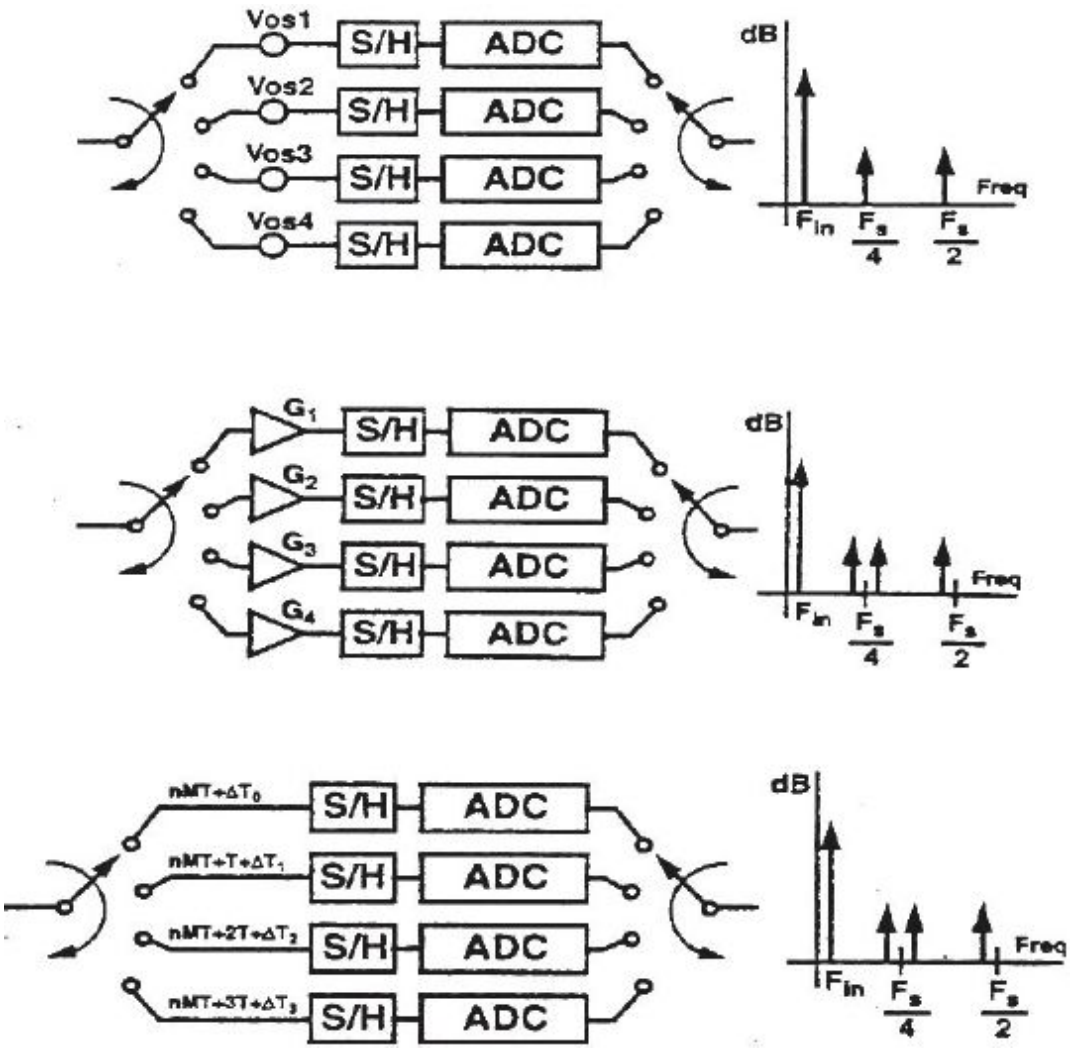


Fig. 9.26: Effect on the spectrum due to the offset errors (at the top), due to the gain (in the center), and due to the sampling interval (at the bottom).

The time interleaved technique allows obtaining converters with very high speed; in this case, the sampling interval regularity assumes a fundamental importance. To limit the effect produced by the timing difference of the S&H stages, adding an input S&H was proposed in literature so as to sample the signal at a frequency of 1Gsps (Fig. 9.27). In this way, the four S&Hs for each channel have a simpler task because they switch between the sampling phase and the hold phase when the input S&H is in the hold phase, i.e. maintains a constant input signal. In this way, the maximum error admissible to ensure bandwidth of 1GHz and resolution of 6bits pass from 2.5ps to 50ps (less stringent condition).

Fig. 9.27 shows the converter scheme and the timing graph for the different blocks. The input S&H must ensure a sampling frequency four times than those of the single channels; for this reason, we use the MESFET technology. The four ADCs are made with the folding technique, which allows significantly reducing the input capacitance, facilitating the S&H's task. A following realization allows reaching 8bits with a frequency of 4Gsps in a hybrid technology. Such high performance levels are obtained with power dissipation in the order of tens of Watt.

The preparation phase (named auto-zeroing) is needed to avoid the offset and charge injection (into the CMOS comparators) effects. In the Flash, half-flash, and pipeline techniques, the comparators' work is dominated by the preparation period, which is needed before each single comparison and which requires, generally, two-four clock cycles. In the SAR converters, instead, the comparator needs only a single preparation step, after which the N comparisons for the N bits are made.

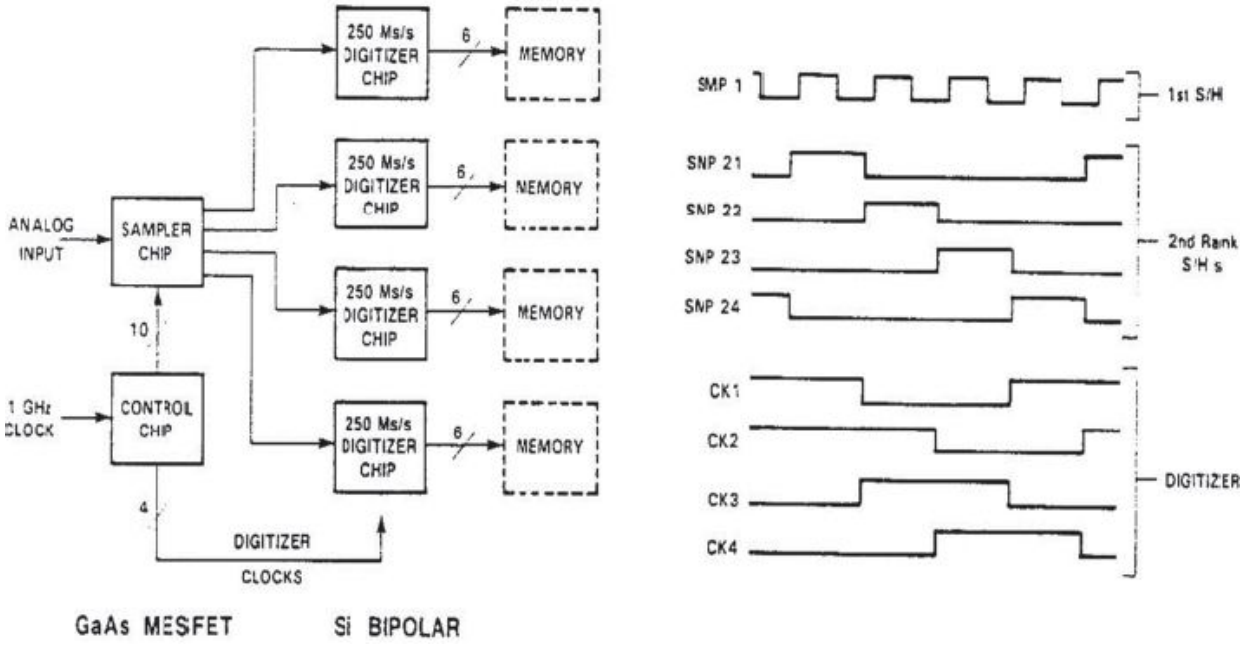


Fig. 9.27: Time-interleaved converter and its timing graph.

The SAR converters in CMOS technology are particularly useful in the time interleaved technique. The great advantage of the SAR ADC is the efficiency in the area occupation in respect of the Flash, half-flash, and

pipeline techniques. Imposing different SAR converters in parallel, we can obtain the time-interleaved parallel SAR architecture (PSA).

The converter efficiency can be expressed in different ways; between these, there is the “decision cost”, equal to the product of the number of comparators and the clock cycle number required to complete the conversion. The lower is the coefficient, the higher is the efficiency. In Fig. 9.28 is shown a comparison of different architecture: N is the number of bits while k is the number of clock cycles for the preparation period and the comparator auto-zeroing. As we can see in the last row of the table, the architecture PSA interleaved proportionally increases the converter speed, but maintains the efficiency unchanged.

Fig. 9.29 shows another time-interleaved configuration, in which every single SAR has resolution of 10bit and a sampling equal to 5Msps. There are 14 converters in parallel, therefore obtaining 70Msps. As a demonstration of the efficiency of the PSA technique, the power consumption was limited to only 267 mW in this realization.

Fig. 9.30 shows four pipelined ADCs in parallel, and each ADC has four stages. The working principle of each stage requires 4 clock cycles. In this case, 8bits and 85Msps are obtained, with consumption of 1.1W.

	comparators	clock periods	decision cost (k=4, N=10)
flash	2^N	k+1	5100
half flash	$2^{(1-N/2)}$	k+2	384
pipelined	n	k+1	50
SA	1	k+n	14
PSA	k+n	1	14

Fig. 9.28: Comparison of different converter architecture.

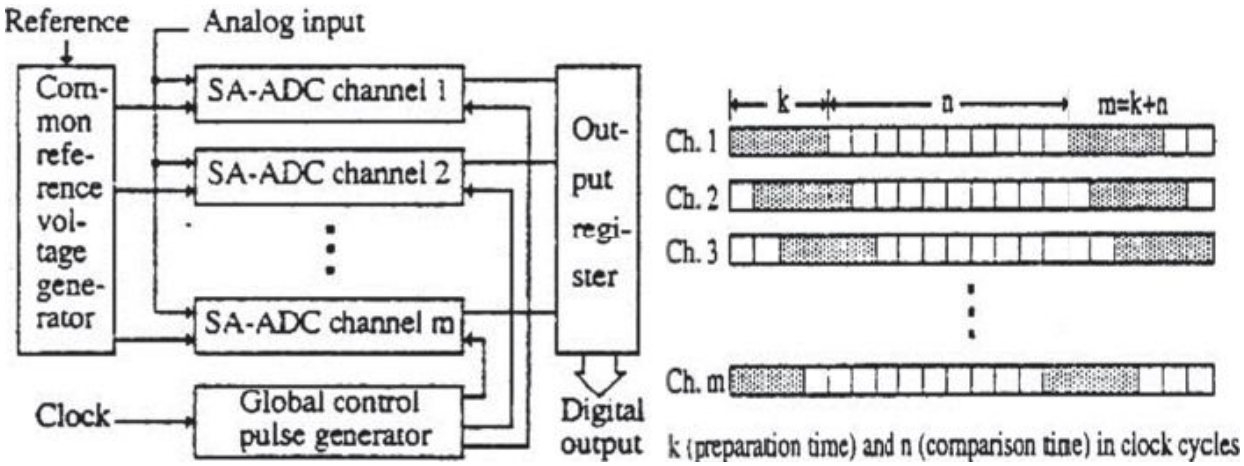


Fig. 9.29: Another time-interleaved converter and the relative temporizations.

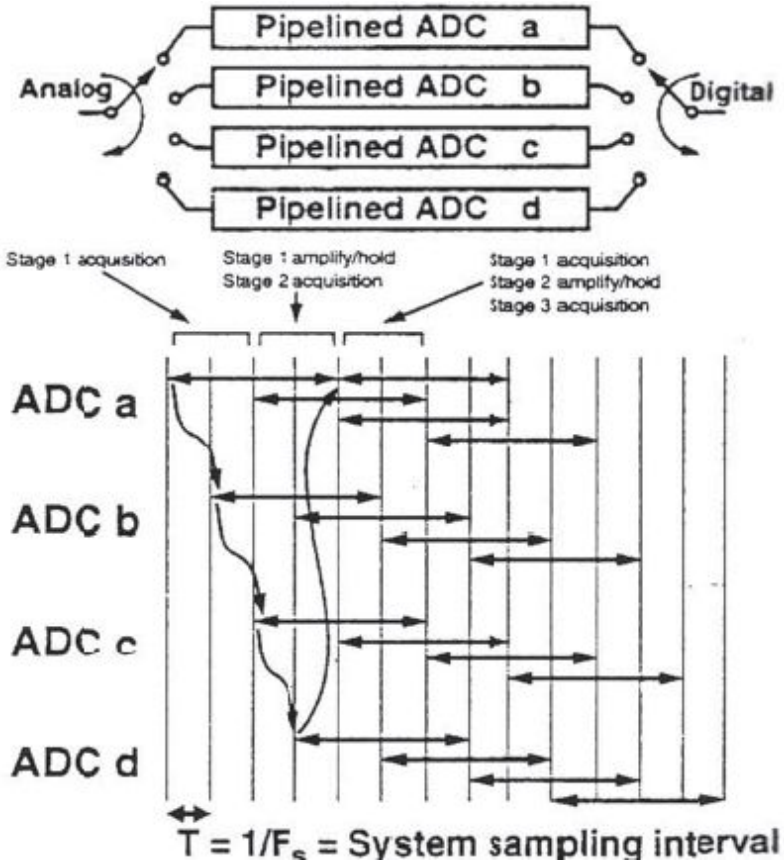


Fig. 9.30: Time-interleaved converter with four pipelined channels.

9.7. Integration Converters

The conversion techniques described so far require considerable technology efforts to obtain high resolution, over 12bits. The linearity of these converters is set by the precision which can be obtained in the components realization. For example, with a standard CMOS process, it is possible to obtain 10÷12bits. Higher resolution requires expensive trimming processes for resistors and capacitors or complex self-calibration algorithm.

The indirect conversion is based, instead, on the translation of the analog voltage signal into an intermediate quantity, time, or frequency, for which the digital value can be extracted, thanks to a counting algorithm. With this philosophy, we can reach very high resolution without the use of sophisticated technologies, at the expense of a long conversion time. However, in some applications, an ADC with very high resolution and accuracy are requested while it is not necessary that the conversion frequency is high because the signal varies slowly. It is the case, for example, of digital voltmeters or monitoring of quantities as, for example, the temperature.

In the integration converters, the input signal is converted into a frequency or a time period; a digital counter measures the frequency or the time period. The final value of the counter, at the end of the conversion period, determines the digital output. In the case of the voltage-frequency converters (VFC), the counter counts the number of pulses generated within a pre-determined period of time. The total number of pulses is proportional to the input signal value. In the case of the voltage-time conversion, the time interval is proportional to the input signal. Using a clock source and a digital counter, we can obtain a digital output proportional to the analog input signal.

The working principle for the integration converter implies a long conversion time, generally a function of signal amplitude and dependent on the clock frequency. Typical conversion times are in the range of 1ms÷500ms. The main advantages of the integration converters are the noise suppression (for example the 50Hz network noise), the monotonicity of the transfer characteristic, the extreme linearity, and the long term stability performance levels. The most popular integration ADCs are the *double ramp* converters and the *charge balancing* converters, which allow reaching resolution up to 22bit.

9.1 Double ramp ADC

The basic scheme for the dual-slope converter is shown in Fig. 9.31. The conversion takes place in two different integration periods: during the first interval whose duration is T_1, the analog signal V_a is integrated. Suddenly, the integrator input is connected to the reference voltage $-V_r$ to bring the integrator output to zero with a fixed slope. The time T_2 necessary to bring the integrator output to zero is proportional to the input signal amplitude V_a. The output code is obtained by the ratio between the numbers of clock periods counted within the two intervals. If N_0 is the number of bits, we impose $T_1=2^{N0}/f_c$, where f_c is the clock frequency.

The control logic, at the beginning of each conversion, resets the integrator, and the counter commutes the switch S_1 on V_a. At the end of the interval T_1, the counter overflows, and the control logic switches S_1 on $-V_r$. At the end of the interval T_2, the comparator changes its state and blocks the counter on the digital output value: the conversion has reached the end. From these considerations, we obtain:

$$\frac{1}{\tau_i}\cdot\int_0^{T1} V_a dt - \frac{1}{\tau_i}\cdot\int_0^{T2} V_r dt = 0$$

where τ_i is the integration time constant. We can write $V_{am}\cdot T_1 - V_r\cdot T_2=0$, in which V_{am} is the mean value of the signal V_a within T_1. Given a numerical value D, accumulated by the counter into the second interval T_2, and considering that $T_1=2^{N0}/f_c$ e $T_2=D/f_c$, we obtain:

$$D=2^{N0}\cdot V_{am}/V_r$$

The main characteristic for this result is that D does not depend on the time constant τ_i of the integrator and on the clock frequency f_c because these parameters are involved in the same way in both the first and the second intervals.

The accuracy is limited only by the accuracy of the reference voltage, the offset, and the comparator and OpAmp noises. Such effects can be greatly reduced by introducing a second cycle, named calibration. In the modified converter, the conversion of the input signal is followed by a further conversion made, connecting the input to the ground.

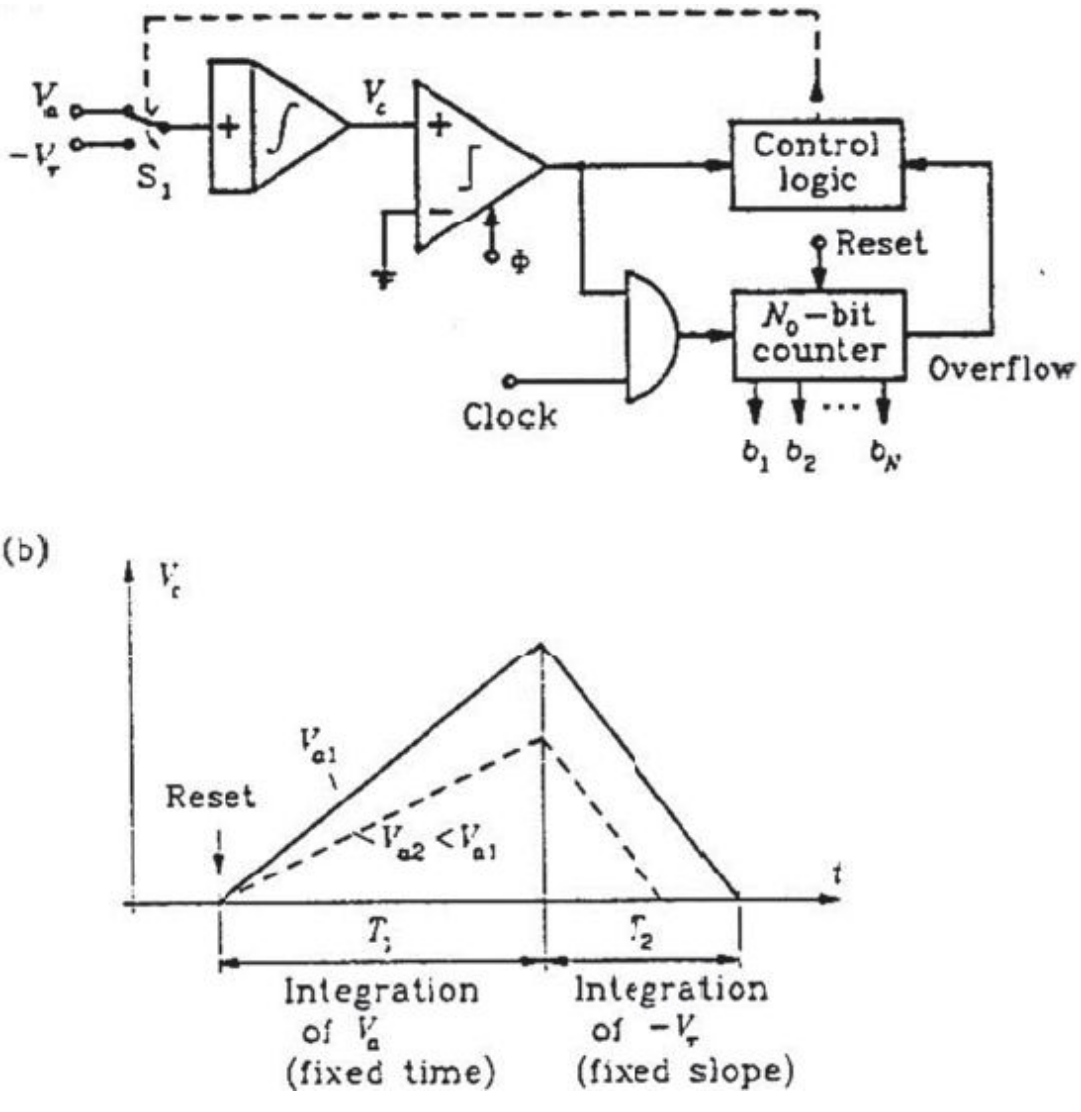


Fig. 9.31: Double ramp ADC architecture and its temporization.

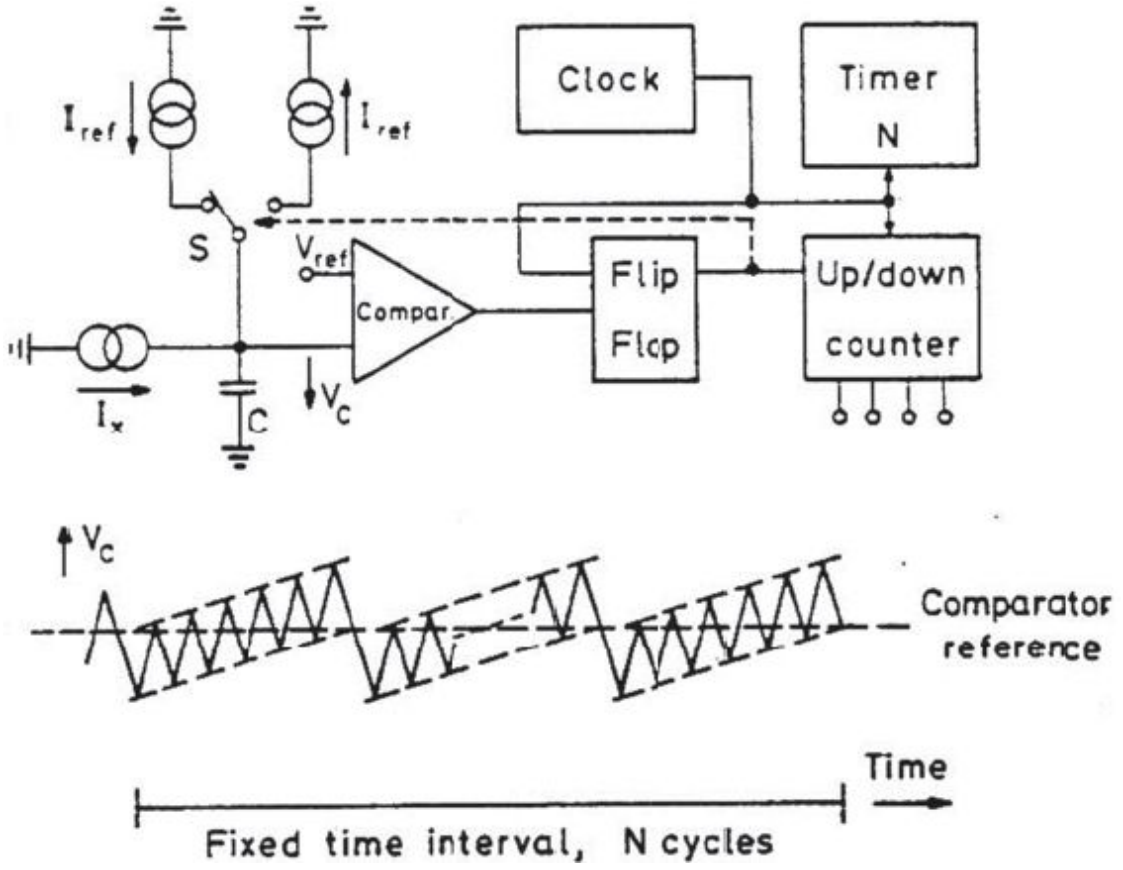


Fig. 9.32: Charge-balancing ADC structure.

By subtracting the second result from the first one, the errors and the comparators' imperfections are reduced. In the triple ramp or multi-ramp converters, we achieve greater accuracy, thanks to complex circuitry.

The advantage introduced by the double ramp converters is that the integration of the signal determines the high frequency noise rejection because it means that the signal is in the first interval. Each AC signal component whose frequency is a multiple of the integration period T_1 will have a null integral and will not disturb the conversion result. Often, this characteristic is exploited to eliminate the noise caused by the power supply, opportunely choosing T_1 equal to a multiple of the period of the supply voltage (1/50s or 1/60s). The use of the double ramp is in the field of instrumentation and the measurement when the conversion frequency required is not so high (for example into the digital voltmeters).

9.2 Charge balancing ADCs

The *charge balancing A/D converters* belong to the category of the voltage-frequency converters and have some similarities with the double ramp converters. The working principle is based on the simultaneous integration of the input signal, and the reference voltage alternatively takes the negative and the positive signs in order to maintain the charge on the integration capacitor balanced, generally around the zero level. An up-down converter measures the difference $N_{up}–N_{down}$, where N_{up} and N_{down} indicate the number of cycles during which the integrator integrates the reference voltage with positive and negative signs, respectively. In other words, the counter records the number of charge packets that the reference source must provide in a pre-determined time period to maintain the charge into the capacitor constant. The value $N_{up}–N_{down}$ is proportional to the amplitude of the input signal.

Because the voltage on the capacitor is maintained approximately constant, the dielectric absorption does not produce errors. Moreover, unlike the double ramp converters, in the charge balancing converters, the input range of the integrator can be more limited and does not depend on the signal amplitude. Furthermore, the conversion time is fixed.

Fig. 9.32 shows a possible charge balancing ADC implementation. In this case, the input and the reference signals are the currents I_x and I_{ref}. After every clock cycle, the flip-flop controls the switch S on the basis of the comparator output. This means that the voltage V_c on the capacitance is above or below the reference voltage V_{ref}. The same Fig. 9.32 shows a possible trend for the voltage V_c. In this case, there is a predominance of N_{down} in respect of N_{up}. In fact, we can see that the waveform is repeated for every 13 clock cycles, during which $N_{down}=7$ and $N_{up}=6$. If N is equal to the sum between N_{up} and N_{down}, the input signal will be proportional to $(N_{down}–N_{up})/N$; in this case, we should have $I_x=I_{ref}\cdot 1/7$.

The charge balancing converted architecture is similar to that used in the Σ-Δ ADC, with the difference of the up-down converter substituted with a digital filter, named the decimator.

10

Oversampling and ΣΔ

"Five houses are painted in different colors. In each house lives a person with a different nationality. The 5 owners each drink a certain type of beverage, smoke a certain brand of cigar, and keep a certain pet. No owners have the same pet, smoke the same brand of cigar, nor drink the same beverage. Some hints are:

- The Brit lives in the red house.
- The Swede keeps dogs as pets.
- The Dane drinks tea.
- The green house is on the left of the white house.
- The green house's owner drinks coffee.
- The person who smokes Pall Mall rears birds.
- The owner of the yellow house smokes Dunhill.
- The person living in the center house drinks milk.
- The Norwegian lives in the leftmost house.
- The person who smokes Blends lives next to the one who keeps cats.
- The person who keeps the horse lives next to the person who smokes Dunhill.
- The owner who smokes Bluemasters drinks beer.
- The German smokes Prince.
- The Norwegian lives next to the blue house.
- The person who smokes Blends has a neighbor who drinks water.

Discover the matches."

10.1. UNDER-SAMPLING

With the under-sampling technique, we sample a signal with a frequency lower than the Nyquist limit. Because of the violation of the limit imposed by the Shannon sampling theorem, aliasing phenomenon occurs, which makes impossible the signal reconstruction of the original signal with a simple filter at the sampler output. Nevertheless, under-sampling can be usefully employed to process very high frequency signals. In fact, by exploiting the periodical repetitions of the original spectrum, due to sampling, it is possible to fold the spectrum of those signals back to lower frequencies, thus performing a frequency conversion of the signal itself.

To highlight the under-sampling effects, we can simply consider the processing of a sinusoidal signal at a frequency f_0 and the following reconstruction phase. Fig. 10.1 shows the sampling spectrum, and the Fig. 10.2 shows the reconstructed spectra:

$$x(t)=\cos(2\pi f_0 t) \quad \rightarrow \quad X(f)=\frac{1}{2}[\delta(f-f_0)+\delta(f+f_0)]$$

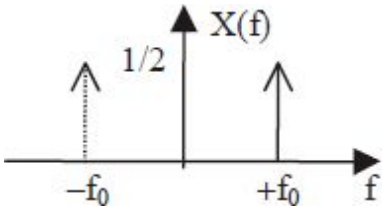


Fig. 10.1: Spectrum of a continuous-time signal.

The bandwidth of a reconstruction filter is maintained constant in the two examples with a cut-off frequency between f_0 and $f_S–f_0$. In the first case, varying the sampling frequency, we can observe how, when f_S suffices to determine the under-sampling (f_S $2f_0$), the reconstructed signal is the following:

$$x_R(t)=\cos(2\pi f_0 t)+\cos(2\pi(f_S-f_0)t)$$

which is the sum of two sinusoids with frequencies f_0 and f_S-f_0.

In the second case, instead, maintaining f_S constant and varying the input signal frequency, the reconstructed signal, with under-sampling conditions

(f_S $2f_0$), becomes:

$$x_R(t) = \cos(2\pi(f_S - f_0)t)$$

In both cases, the under-sampling involves, during the reconstruction phase, an input signal conversion towards a lower frequency: such an effect is exploited in the under-sampling techniques.

Particularly, for $f_S=f_0$, the reconstructed signal is constant, in fact:

$$f_S = f_0 \quad \rightarrow \quad x_R(t) = \cos(2\pi(f_S - f_0)t) = \cos(2\pi \cdot 0 \cdot t) = 1$$

To use the under-sampling technique, we need to understand the input signal characteristics precisely. We can, in fact, under-sample only the signals whose, despite their high frequency components, bandwidth is less than half of the chosen sampling frequency. We need, moreover, that the input signal components lie around integer multiples of $f_S/2$.

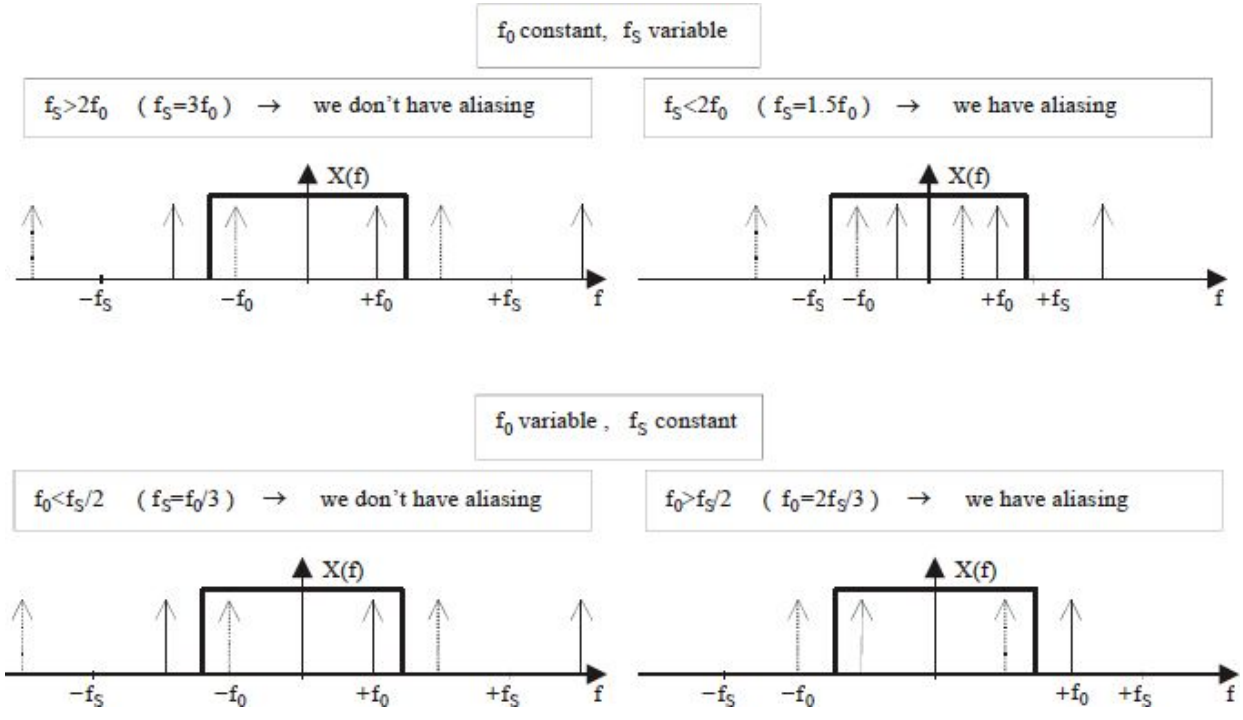


Fig. 10.2: Sampling and reconstruction of a signal with frequency f_0.

To clarify the characteristics of an under-sampling device, consider now a numerical example. We can under-sample a pure 9.1MHz sine-wave with a sampling frequency of only f_s=10MHz. We choose the sampling frequency in order to have the input signal spectrum close to a multiple of $f_S/2$ after the sampling operation. As a result of the sampling operation, we have infinite replicas of the input signal spectrum around the frequencies $|f_{in}–n\cdot f_S|$. Particularly, in the case of the considered sine-wave, we can find a spectrum at only 900kHz, as highlighted in Fig. 10.3. In doing so, we obtain the double result to have an amplitude demodulation of the signal and to employ converters with a working frequency lower than that imposed by the Shannon Theorem. In this example, we should employ an ADC with a sampling frequency higher than 18MSps ($f_s/2f_{Ny}$/(18MSps).

A slightly more complex case, but in which we can effectively implement the under-sampling technique, is the following: assume that the original signal's spectral components are far away from each other, but that the total bandwidth is less than $f_s/2$ (Fig. 10.3). The mapping between the original input frequency 'f_{in}' and the repetition frequency of the spectra '$f_{Sampled}$', as discussed before, is provided by the relationship:

$$f_{Sampled} = \left| f_{in} - n \cdot f_s \right|$$

choosing 'n' to have $f_{sampled}<f_s/2$. The frequency interval $f_{sampled}$ between 0-$f_s/2$ is cyclically repeated up to infinity, around the integer multiples of f_s. The mirror has images which re-'mirrored' on the side of every multiple because all the information of $f_{sampled}$ is available in the intervals 0-$f_s/2$ (and multiples) and $f_s/2$-f_s (multiples). We have also a sign '+' or '-' for the translation of the baseband, depending on the position of the origin of the input signal, as shown in Fig. 10.4.

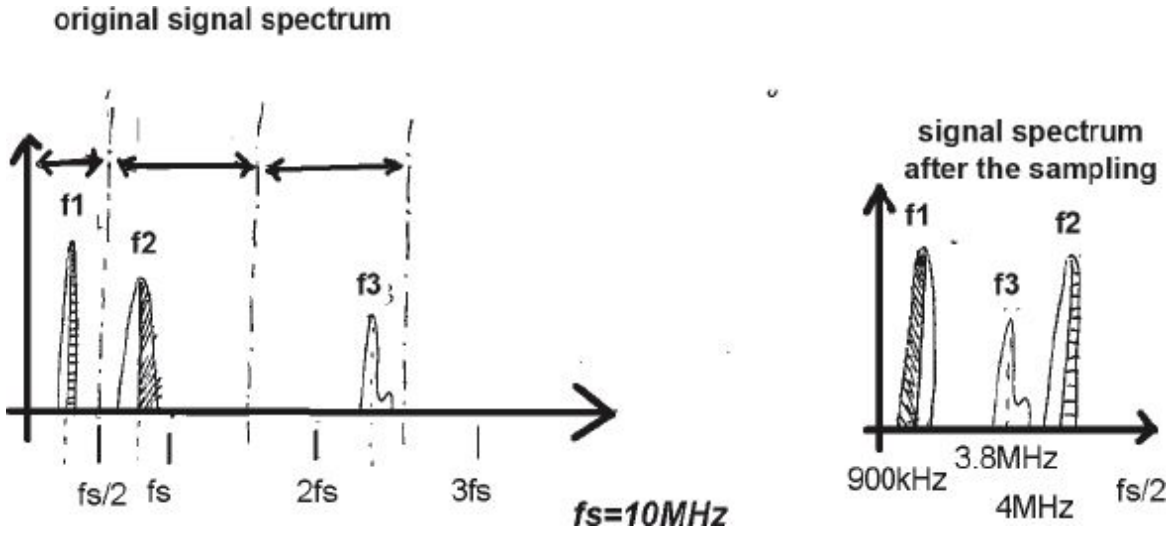


Fig. 10.3: Folding effect for the signal spectral component after the sampling.

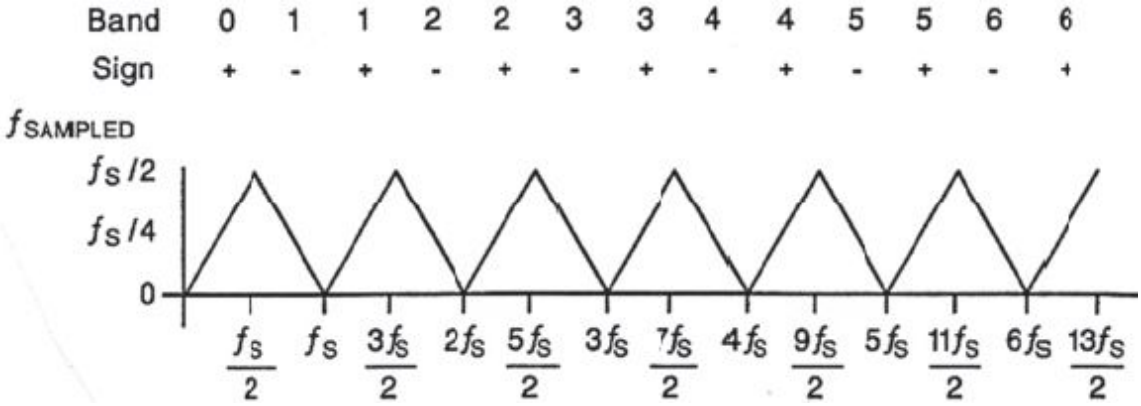


Fig. 10.4: Images mirrored around integer multiples of f_s.

To translate the input signal spectra into the baseband, they must be included in one of the sub-bands (marked in Fig. 10.4 by 0+, 1-, 1+, 2- …); otherwise, in the case of larger bandwidth, we could have a distortion and a loss of information. To reconstruct the useful signal, the spectral components have to be included in one of the sub-bands.

To demonstrate the effectiveness of the under-sampling, we could consider a converter model employed in this type of application; it is the device ADC614 produced by Burr-Brown. Salient features for this commercial converter are 14bits resolution and a sampling frequency up to 5MSps. The amplitude of its analog bandwidth (FPBW: Full Power input Bandwidth) is up to 30MHz and enables the converter for the under-

sampling. In this type of application, the input signal has very high frequency components (even higher than the maximum frequency f_s of the device), and therefore it is necessary that the input stage of the ADC can correctly perform the conversion without slew-rate limitations and introducing distortions.

From the following graphs, we can see how the device's performance is for both the low frequency (f_{in} $f_s/2$) processing and the under-sampling input signal. The graph in Fig. 10.5 shows the output spectrum of a sinusoidal signal with a frequency of 2MHz (the frequency close to the Nyquist limit for this device). The signal-to-noise ratio, measured by means of the SFDR (Spurious-Free Dynamic Range), is equal to 87dB. The measurement shown in the graphs of Fig. 10.5 and Fig. 10.6 are made by using a sampling technique fixed around 5MHz, with slight adjustments made in accordance with the processed input signal, to make a coherent sampling possible. To maintain the coherence between frequencies and phases, a phase-locked-loop (PLL) has been used.

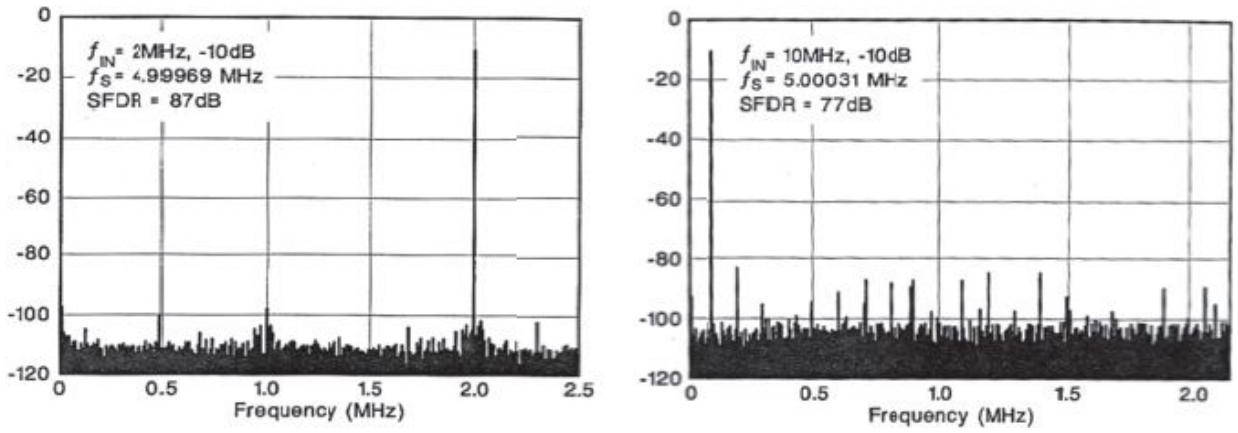


Fig. 10.5: Output spectrum obtained with a sampling of $f_s>2\cdot f_{in}$ (on the left) and with $f_{in}=5MHz$ close to f_s (on the right).

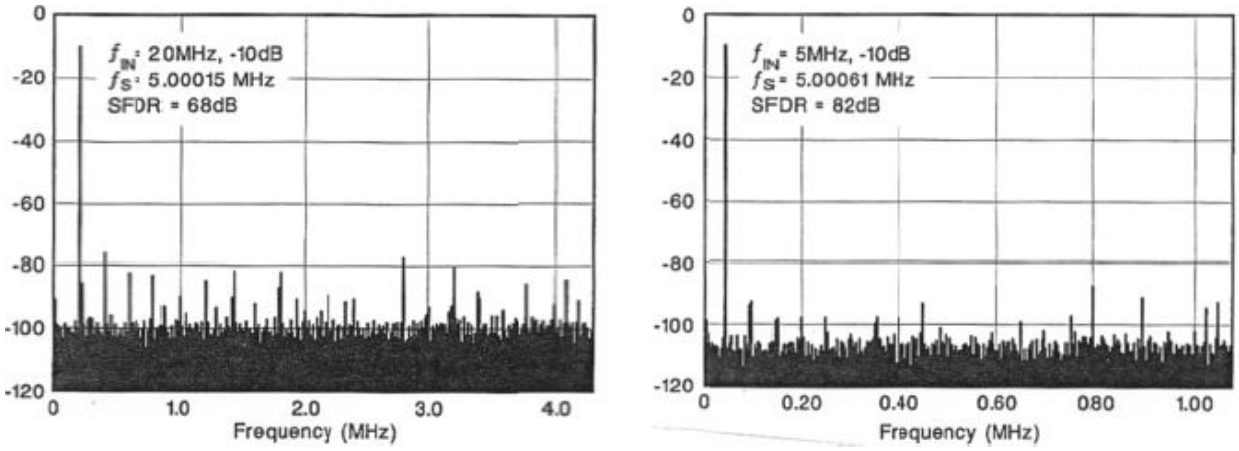


Fig. 10.6: Output spectrum obtained with a sampling of f_{in}=10MHz, double in respect of f_s (on the left) and with f_{in}=4f_s (on the right).

For example, by employing a sampling frequency of 5.0061MHz close to the input frequency (f_{in}=5MHz), an SFDR of about 82dB is obtained (Fig. 10.6). Doubling the input frequency, we can see a further SFDR reduction of 5dB up to 77dB. Finally, in Fig. 10.6, we can see the output spectrum with f_{in}=4f_s and obtain an SFDR=68dB.

For every input frequency doubling, we have a constant and low deterioration in the signal-to-noise ratio of about 5dB. We can understand how the under-sampling technique is an effective and cheaper method to convert high frequency analog signals, avoiding the use of very high resolution and speed, which is expensive.

To obtain such performance levels, the converters employed in under-sampling applications must satisfy some important requirements. Because the input signal bandwidth is many times higher than the sampling frequency, ADCs for under-sampling must have input bandwidth wider than the typical sampling frequency and good linearity characteristics. About this last aspect, we can consider the static indicators DNL and INL. The former, indicating the variation in the code width in respect of the ideal value of 1LSB, takes into account the increment in the quantization noise and the increment in the noise-floor. The INL, instead, shows the deformation in the converter characteristic in respect of the ideal trend (straight line), generating spurious harmonics. Nevertheless, for under-sampling applications, it is very important to pay attention to the dynamic characteristics described by parameters, for example the SiNAD (usually in

the range 40÷60dB, slightly lower than the SNR), the ENOB (usually 1 or 2bits less than the nominal resolution of the ADC), and the SFDR (typical values in the range 50dB÷80dB). Other indicators useful to know the ADC non-linarites are the THD, which refers to the first 5 spurious harmonics generated, and the IMD (between –70dB and –90dB).

Because the input signal has very high bandwidth, it is very important to have a very low aperture jitter for the input S&H. This jitter makes a phase modulation of the sampling moment, limiting the maximum signal frequency allowed (derivative dV/dt). For example, to maintain the error less than $^1/_2$LSB, the maximum input frequency is brought by the following relationship:

Vendor	Device	Bits	Sample rate (Msamples/sec)	Bandwidth (MHz)	Power	SNR* (dB)	SINAD (dB)	SFDR (dBc)	THD (dBc)Ø	IMD (dBc)	Price
Analog Devices	AD9042 AD876	12 10	41 20	 150	5V/575 mW 5V/160 mW	70	 51	–80 –65	 –60	–90	$200 (1000) $10 (1000)
Analogic Corp	ADC3120	14	20	80	±15, +5, –5.2V,4W	75		–90	–82		$3500 (10)
Burr-Brown Corp	ADS7819 ADS605	12 12	0.8 10	1.5	±5V/225 mW ±5V, 220 mW	70 66	70 63	–77 –63	–82 –70		$22.15 (100) $125 (100)
Comlinear Corp	CLC949	12	30	100	5V/400mW	65		–72		–70	$98 (1000)
Datel Inc	ADS-946 ADS-945	14 14	8 10	10 50	±5V/19W ±15, -5, –5.2V/4.2W	74 78	70 74		–73 –80	–82 –84	$381 (100) $866 (100)
Harris Semi-conductor Corp	HI1386 HI5702	8 10	75 40	150 250	–5.2V/580 mW 5V/390mW	57	41 56	–63	–60	–59	$48 (1000) $35 (1000)
Linear Technology Corp	LTC1410	12	1.25	20	+5V/160 mW		68		–82	–84	$23 (1000)
Maxim Integrated Products	MAX153 MAX100	8 8	1 500	1 1200	5V/40 mW 5, –5.2V/5.2W	45	45	–50	–50		$6.63 (1000) $265 (100)
Micro Linear Corp	ML6401	8	20	80	5V/235mW	44		–50	–46	–46	$6.75 (1000)
National Semi-conductor Corp	ADC12062 ADC12662	12 12	1 1.5		5V/75 mW 5V/200 mW	72 70	71 70		–82 –80	–80 –80	$29 (1000) $34 (1000)
Signal Processing Technologies Inc	SPT7750 SPT7840	8 10	500 10	500	5V/5.5W 5, –5.2/100 mW	44 57	37 56	–41 –63	–38 –59		$250 (100) $9.60 (1000)
Sony Electronics Inc	CXA1866Q	6	140	210	±5V/325mW	32					$20 (1000)

Tab. 10.1: Comparison table for the most common ADCs.

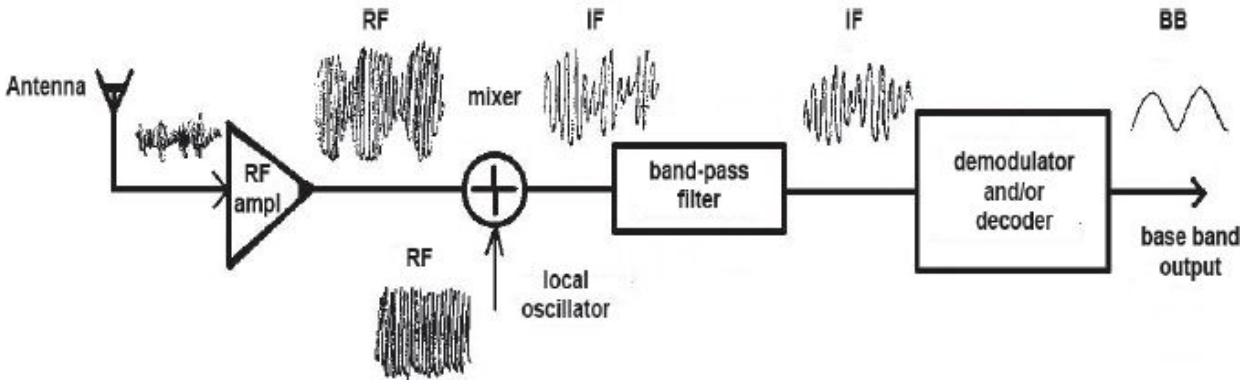


Fig. 10.7: Armstrong super-heterodyne detector.

$$f_{inMax} \leq \frac{1}{t_{aperturetimejitter_{rms}} \cdot 2\pi \cdot 2^{n+1}}$$

The jitter effect on the SNR causes, in fact, degradation of:

$$SNR_{jitter} = -20\log(2\pi \cdot f_{in} t_{aperturtimejitter})$$

Therefore, in a 12bits ADC with an input signal whose bandwidth is 1MHz, a jitter of $20ps_{rms}$ causes a worsening of the SNR, which is equal to 15dB.

It is important to highlight that all the specified parameters must be evaluated for input signals with high bandwidth, even up to the whole FSR. On the contrary, an evaluation made for only small signals could erroneously provide results that seem to be better than they really are because the converter is not fully exploited.

Tab. 10.1 shows the main data for some ADCs used for under-sampling RF communication applications.

10.1.1 Under-sampling Applications

The under-sampling can be used in RF communications. In analog radio, in order to receive a radio-frequency it was common to use the Armstrong super-heterodyne equipment. The scheme in Fig. 10.7 shows such a structure for receiving HF broadcasts (*High Frequency*, also called ‘long’ waves, LW, in the 3÷30MHz range) in AM (Amplitude Modulation). For example, if we desire to demodulate an input signal of 3.5MHz, we will “beat” the captured signal (opportunely amplified) with a local oscillator with a frequency of

3.955MHz. Therefore, the mixer output will be a signal at a difference frequency of 455kHz, named the intermediate frequency IF, and the component at the sum frequency is eliminated by means of proper filtering around the IF. At this point, the signal can be demodulated in accordance with its characteristics: in the case of AM modulation, we will use a diode detector or a multiplying detector for a signal SSB (Single Side Band); instead, for a signal FM, we will use the Foster-Seeley discriminator; or, finally, in the case of digital modulation, the signal will be opportunely decoded.

The transition to the intermediate frequency is usually performed to simplify the realization of filters and amplifiers, separating the often conflicting requirements of RF selectivity, sensitivity, and demodulation purity. In some cases, we even use a double conversion on two intermediate frequencies gradually decreasing, doubling the required circuitry (mixer, oscillator, and filter). For example, in the case of VHF (Very High Frequency) detection, for short waves in the range 30÷300MHz, we can employ a unique IF at 10.7MHz or a double conversion at 30.85MHz and then at 455kHz.

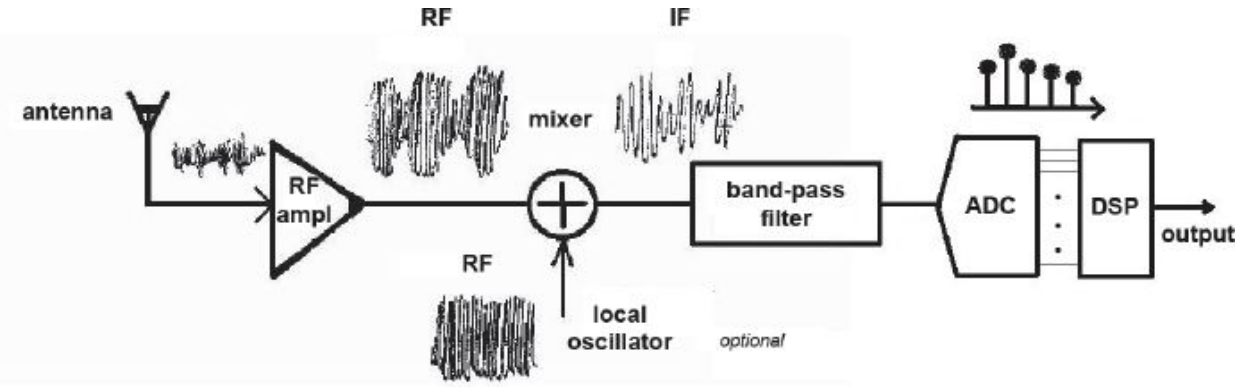


Fig. 10.8: "Baseband sampling" architecture.

The use of an IF often allows resolving conflicting requirements of RF selectivity, sensitivity, and demodulation purity Today, we try to substitute this architecture, constituted entirely by analog components, with a system that realizes a great part of the signal processing chain by employing digital devices and using analog devices only in the interface between the two "worlds" (close to the antenna). This technique is known as "Digital IF",

“baseband sampling”, “IF sampling”, or “baseband digital conversion”. Fig. 10.8 depicts the basic architecture of this apparatus.

In this way, we try to move the electronic project to only three integrated circuits: the analog front-end with the antenna, one digital circuit to convert from the RF (or IF) to the baseband which employs an ADC (with the associated digital circuitry), and finally a processor (or DSP) to extract and demodulate the desired signal.

The basic idea of the A-to-D conversion is the under-sampling to bring the high frequency to the baseband (the lateral bands contain the information). Depending on the RF frequency of interest, a first down-conversion from RF to IF can be necessary, but the actual conversion is made by the under-sampled ADC. The task of the DSP is to extract the starting signal from the under-sampled digital signal. In this “digital” approach, therefore, the front-end and mixer optimization for the RF is not needed because it is the ADC that digitalizes the whole RF bandwidth (with all channels). To recover the original information, the converter operates by dynamically discriminating single channels and implementing the correct codification. This operation is performed entirely by the DSP. It is not necessary to modify the hardware (as for the super-heterodyne device) for different types of modulation, but we act directly via software (managing the information processing steps from the antenna to the last speaker).

Notice that what has just been presented is the “software radio” that digitalizes a wideband and that is very different in respect of the single channel wideband down-conversion. In fact, in this last case, we have a wideband containing many independent signals (various analog radio stations, ETACS phones, digital radio, GSMs, digital protocol transmission ...) which must be demodulated simultaneously. The DSP must be extremely sophisticated in order to execute multiple couples of the “tune” and “decode” algorithms for all the signals which must be recovered. In both cases, the ADC must have a very high linearity (maintaining a low IMD) reducing the intermodulation distortion which causes beats between different processed frequencies. For this reason, ADCs used in under-sampling applications are tested with eight tones (not only two).

The under-sampling technique is advantageously employed to visualize extremely fast waveforms, as made in some *sampling oscilloscopes*. In these equipment, the under-sampling effect is used to convert downwards the

frequencies of the pulse whose trend we want to trace. Practically, the sampling oscilloscope visualizes an attenuated and slowed version of the original pulse. Obviously, for a correct operation, we need a series of pulses with the same trend. The sampling oscilloscope action is exemplified in Fig. 10.9, in which the periodic signal under measurement, the pulses train with a variable period, and the output signal are shown. As we can see, the output signal is a slowed version of the single period of the original signal.

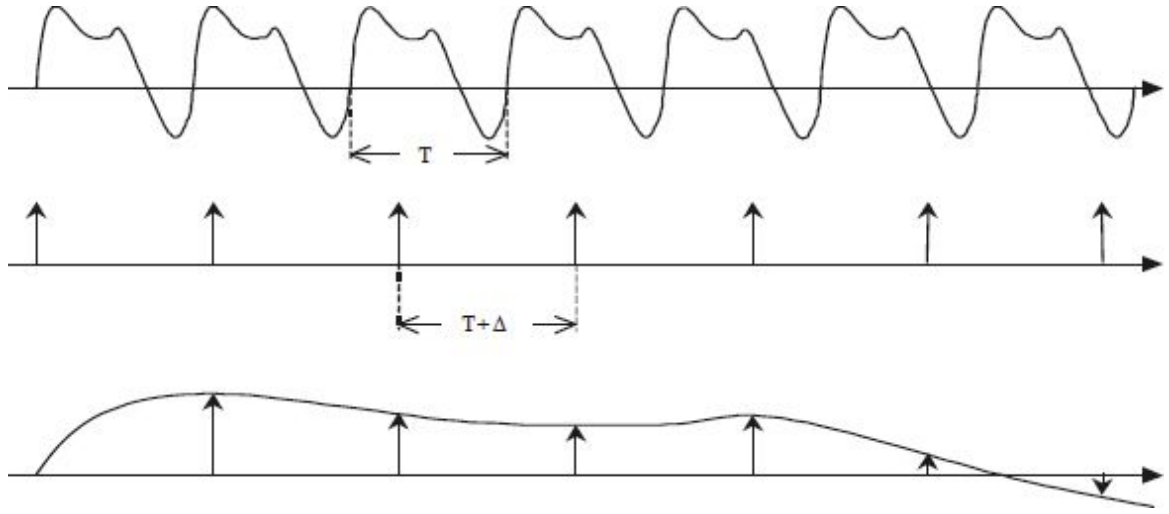


Fig. 10.9: Input signal, sampling pulses, and output of the sampling oscilloscope.

10.2. Over-Sampling

On the opposite side, the over-sampling technique, using a sampling frequency much higher than the Nyquist limit ($f_s >> 2f_{Ny}$), determines a clear separation in the spectral replicas of the sampled signal, reducing the aliasing problem during the reconstruction. Fig. 10.10 shows a signal sampled with a frequency $f_s=2f_0$ equal to the Nyquist limit (b) and an example of over-sampling with a frequency $f_s=4f_0$ (c).

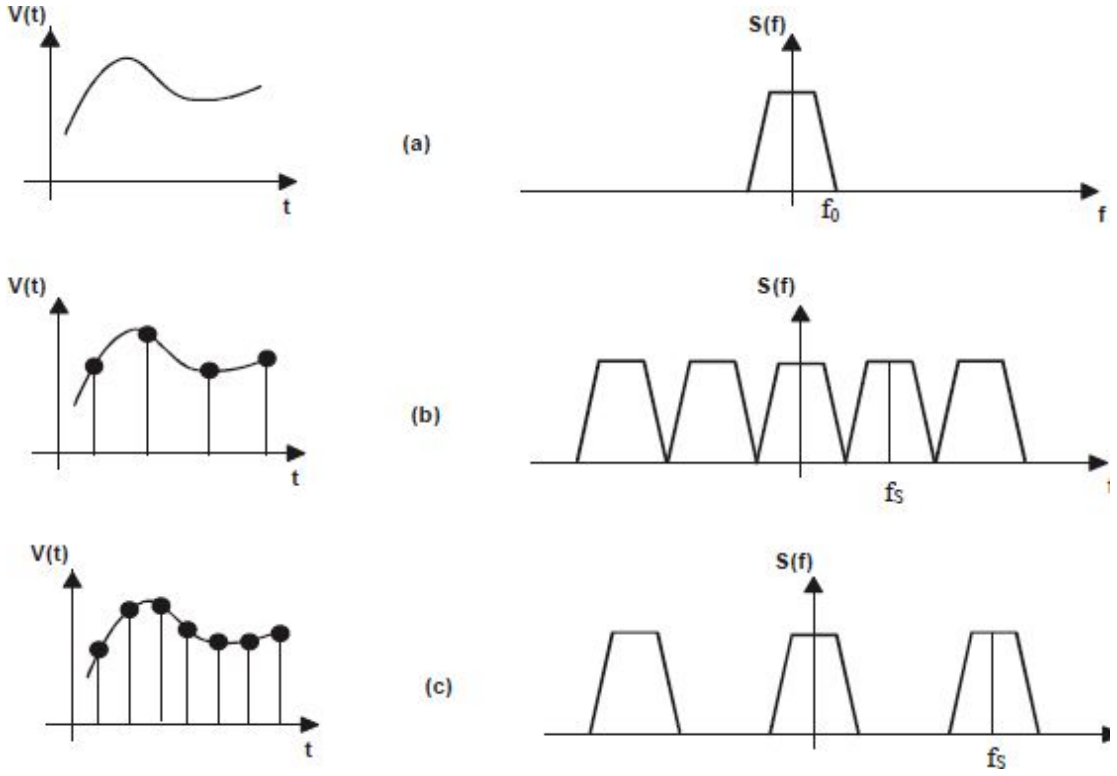


Fig. 10.10: Original analog (a) and sampled with $f_S=2f_0$ (b) and $f_S=4f_0$ (c) spectra.

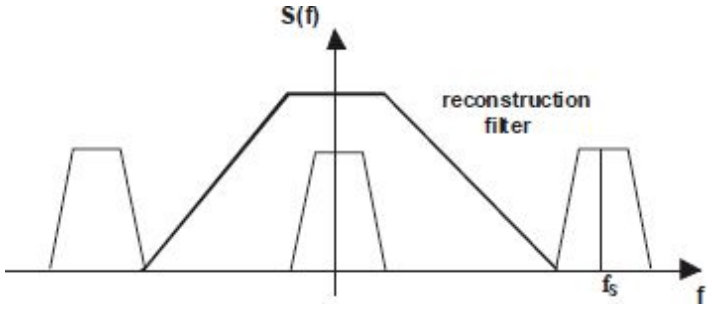


Fig. 10.11: Selectivity requirements for a reconstruction filter for a signal sampled with $f_S=4f_0$.

In both cases, the useful information is contained in the range of frequency between $-f_0$ and $+f_0$; nevertheless, in the example of Fig. 10.10c, if you desire to reconstruct the original signal from the quantized samples, you must use a reconstruction filter (downstream the sampler) with selectivity requirements not very stringent, as shown in Fig. 10.11, without distortion or aliasing problems.

We need to remember that, in the case of over-sampling, we need to apply an anti-aliasing filter upstream the sampler to abolish the disturbance

components outside the useful bandwidth; these components could be superimposed on the baseband signal (distorting it) because of the frequency folding effect.

The analog-to-digital conversion looks like the series of the different functions depicted in the block scheme in Fig. 10.12. Upstream the sampler, the signal undergoes an anti-aliasing low-pass filtering needed to define the bandwidth. The real sampling operation, which "picks" the signal at fixed intervals, follows the quantization phase. This operation introduces an important distortion of the original signal because it enforces the approximation of each input sample with the closest level allowed.
Finally, there is the sample coding phase with a fixed number of bits (word).

With a high resolution sampling of a non-constant input signal, the quantization noise is well approximated with a white noise superimposed on the signal and with a power of:

$$\sigma_q^2 = \left(\frac{FSR}{2^n\sqrt{12}}\right)^2 = \frac{LSB^2}{12}$$

For example, with a Full Scale Range of 5V, a 12bit conversion will have a noise of $\sigma_q^2 = (352\mu V)^2$ due to only the quantization noise. If this noise is uniformly distributed on the range $0 \div f_S$, we would get a constant power spectral density (for quantization noise) equal to: $Q_q(f) = \frac{\sigma_q^2}{f_S/2}$. At 20kHz sampling, the value would be $Q_q(f)=(3.52\mu V/\sqrt{Hz})^2$.

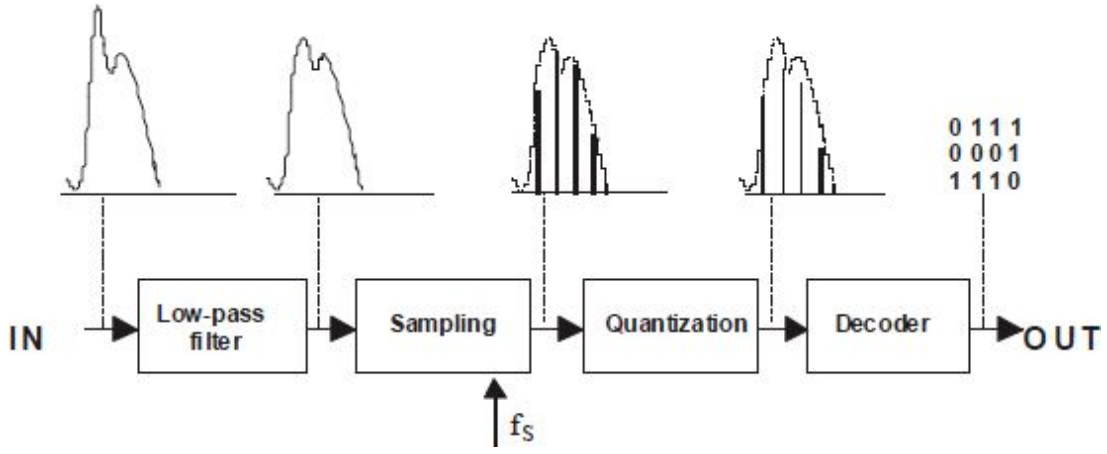


Fig. 10.12: A-to-D block conversion scheme.

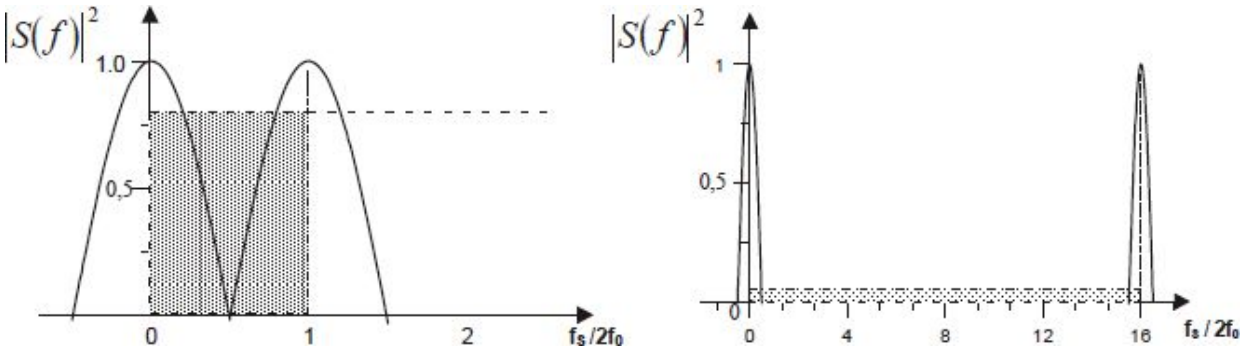


Fig. 10.13: Signal and noise quantization spectra in the case of Nyquist sampling (a) and over-sampling by a factor OS=16 (b). The dashed regions have the same area and represent the quantization noise power σ_q^2 while the amplitude is the density Q_f^2 (f).

If f_0 is the upper limit of the signal bandwidth, the quantization noise power superimposed on the useful information depends on the sampling frequency; therefore, with a variation in f_s, the noise power σ_q^2 is constant while the spectral density will change, as shown in the graphs in Fig. 10.13 (the dashed regions are equal).

If the input signal has the same bandwidth, after the quantization, to obtain the signal, we will use a simple low-pass filter. Employing a selective reconstruction filter whose bandwidth is equal to that of the useful signal f_0 (the maximum frequency of the input spectrum), the total noise power superimposed on the reconstructed output information (due to only the quantization noise of the A-to-D conversion) is obtained integrating $Q_q(f)$ on the interval $0 \div f_0$ and depends on the ratio $2f_0/f_S$ in accordance with the following relation:

$$\sigma_{out}^2 = \int_0^{f_0} Q_q(f)\,df = \sigma_q^2 \cdot \left(\frac{2f_0}{f_s}\right)$$

Having defined the ***oversampling factor*** (OS) as the ratio '$(f_s/2)/f_0$' between the used sampling frequency and the lower limit fixed by the Nyquist theorem (f_S/f_{Ny}), we try to highlight how many times we oversample in respect of the needed '$2f_0$' sampling frequency. We see,

moreover, a noise power reduction superimposed on the signal of a factor $\sigma_{out}^2=\sigma_q^2/OS$. Because the oversampling reduces the r.m.s. value of the quantization error within the bandwidth of the OS ratio, in every sampling frequency doubling, we see a 3dB noise quantization reduction. In the case of the Nyquist limit sampling ($f_s=2f_0$ with OS=1), we do not have noise attenuation. Thanks to the over-sampling (OS>1), the noise within the useful bandwidth (0-f_0) is reduced and distributed uniformly in the range 0-f_s.

This can be said in terms of SNR. Assuming that the quantization noise is the only noise contribution and applying a sinusoid whose amplitude is equal to the maximum dynamic

(FSR/2), we have: $$SNR_{theoretical}=\frac{V_{eff}}{q}=\frac{FSR/2\sqrt{2}}{\Delta/\sqrt{12}}=\frac{FSR/2\sqrt{2}}{FSR/2^n\sqrt{12}}=2^{n-1}\sqrt{6}$$

Translating the expression into dB, we have:

$$SNR_{theoretical\,dB}=20\log_{10}\left(2^{n-1}\sqrt{6}\right)=6,02n+1,76$$

In the case of over-sampling, the SNR improvement is equal to:

$$SNR_{theoretical\,dB}=20\log_{10}\left(2^{n-1}\sqrt{6}\right)=6,02n+1,76$$

We see an improvement of a factor √2 (equal to 3dB) for every doubling of the sampling frequency f_s or, equivalently, an increase in the resolution corresponding to a ½bit; quadrupling f_s corresponds to reach an SNR improvement of 6.02dB, i.e. an SNR equivalent to that of an ADC with 1bit more than the original one.

With a sinusoid with a frequency f_0 and sampling with $f_s=2f_0$, we should have the trend shown in Fig. 10.14a. To reconstruct the signal, we should filter opportunely. Assuming that we use a low order filter (one pole) in f_0, we obtain the qualitative reconstruction shown in Fig. 10.14b; indeed, using the same filter with a pole at f_0 and over-sampling of a factor 10, we should obtain Fig. 10.14c with its reconstruction in Fig. 10.14d. With the filter

frequency at $f_s/2$, we will not have any advantage (Fig. 10.14e). Further improvements can be obtained by properly using the digital data with post-processing operations on the digital frequency.

We saw that f_s is higher than the minimum needed to correctly reconstruct the original waveform. This means that the digital data streaming has a very high frequency. The over-sampling operation produces an amount of information higher than the minimum to reconstruct the original signal. Such additional data can be employed in the following post-processing phase of the digital sequence to enhance the conversion accuracy by means of simple (either analog or digital or both) filtering.

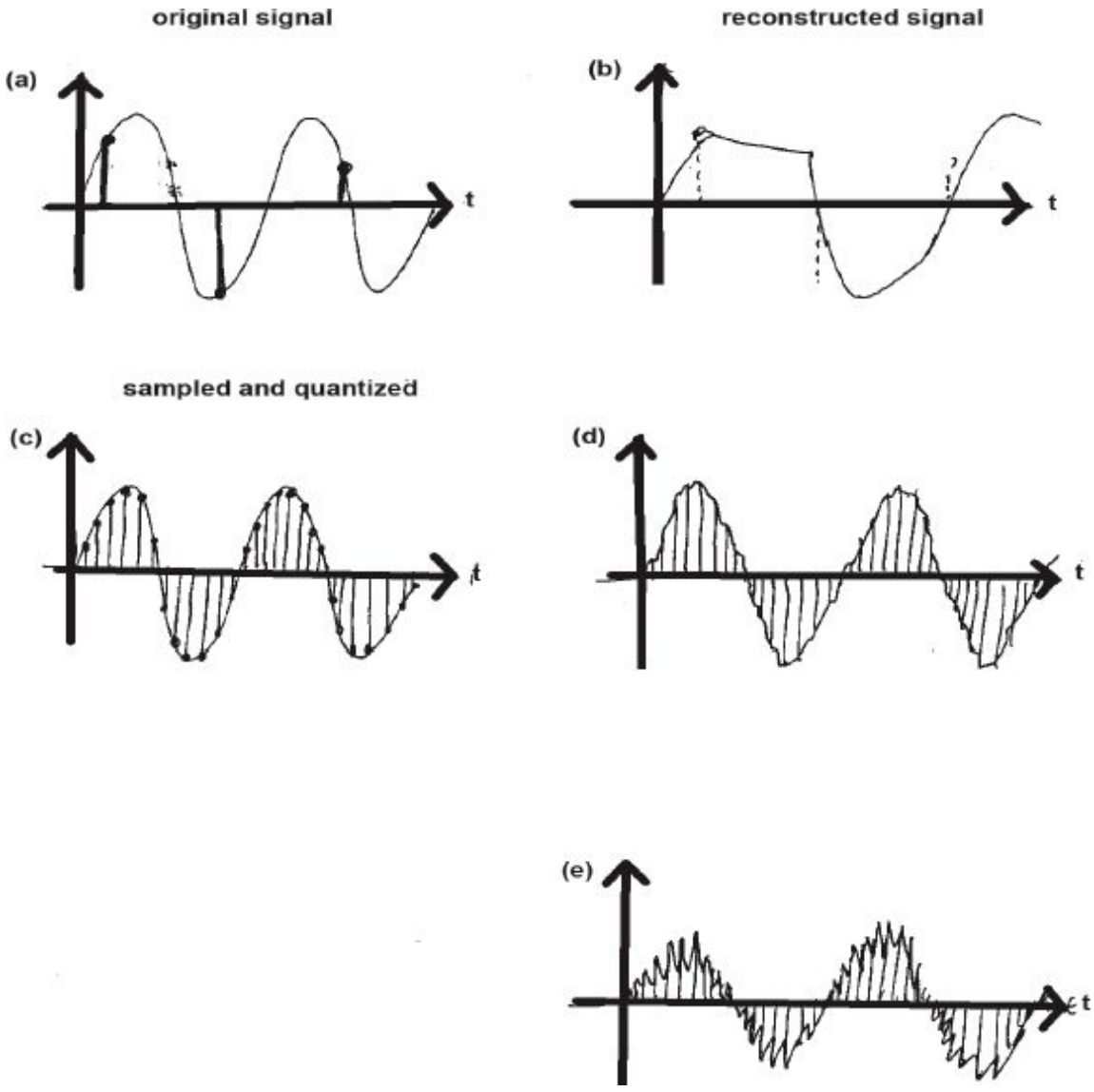


Fig. 10.14: Qualitative description of the advantages of sampling with proper filtering. (a) and (b) Signal sampled with $f_s/2f_0$ and reconstruction with a filter with a pole at f_0. (c) and (d) Signal oversampled with a ratio of 10 and its reconstruction. (e) Reconstruction with a filter with a cut-off frequency equal to $f_s/2$.

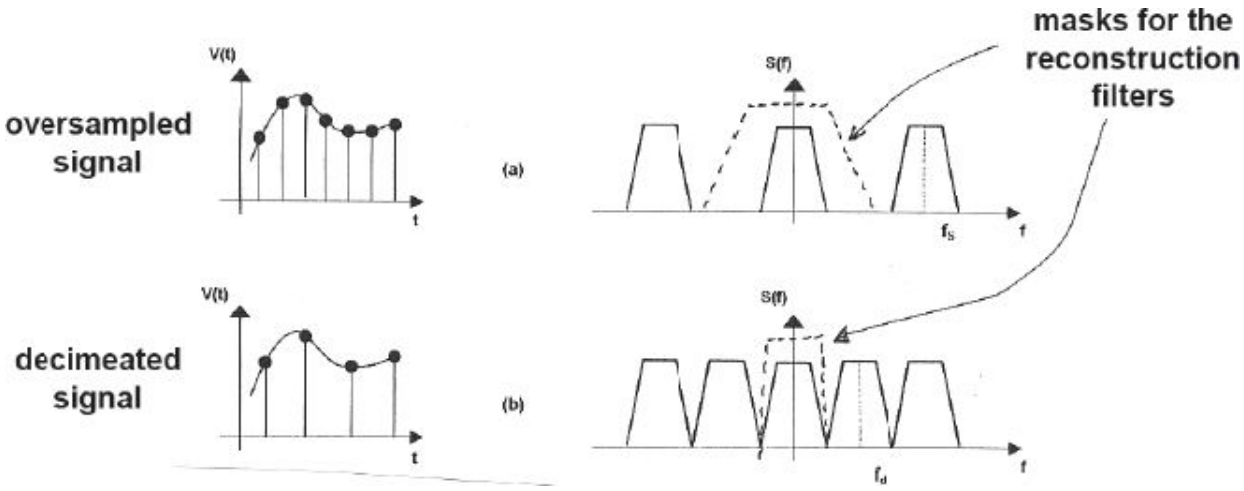


Fig. 10.15: Signal and oversampled spectrum (a) and then decimated (b); the dashed line represents the mask for the reconstruction filters.

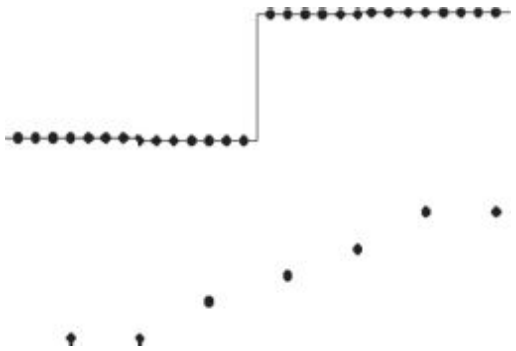


Fig. 10.16: Signal sampled and quantized with only 2 levels (at the top) and average with 4 samples (at the bottom).

After an OS=2 (Fig. 10.15), we could use, on the signal (a), a high-order filter or eliminate every other samples. his second operation (b) is a multiplication in the time domain with a comb 101010 or a convolution of the spectrum in (a) with a δ-comb around the frequency $f_s/2$, obtaining the spectrum (b).

Therefore, it is smarter to use over-sampled data and perform a digital processing. For example, we can make an average for two adjacent digital data and provide it to the output with a reduced frequency. Fig. 10.16 shows an example of this: decimation with a factor 4 is performed (one bit every 4) on a signal sampled and quantized with only two levels (at the top in the figure), employing only an ADC with one bit. At the decimator output, the samples follow one another with a frequency lower than the input one (at the bottom in the figure), but the resolution is increased with more levels because the resolution is improved by a factor of 4 (2 bits).

One can appreciate this effect on a sinusoidal signal over-sampled, but with a low number of bits. Fig. 10.17 shows both the starting analog signal (for example, a 1Hz sinusoid with FSR=10V and f_s=32Hz, i.e. OS=16) which is the 3bits quantized signal (8 levels).

Fig. 10.18, instead, shows the over-sampled signal with a frequency equal to 16 times the Nyquist one, which is then quantized (averaged) with a 3bits resolution device. Downstream the decimation operation, the obtained sinusoid is similar to the original signal with 32 allowed levels, corresponding to a 5bits conversion. From Fig. 10.19, we can see the phase delay introduced by the digital processing because the digital filter makes an average on the input data and must wait a number of samples (in this case 16) to provide the correct processed output. This is a simple signal delay, i.e. the filter introduces a linear delay on the signal.

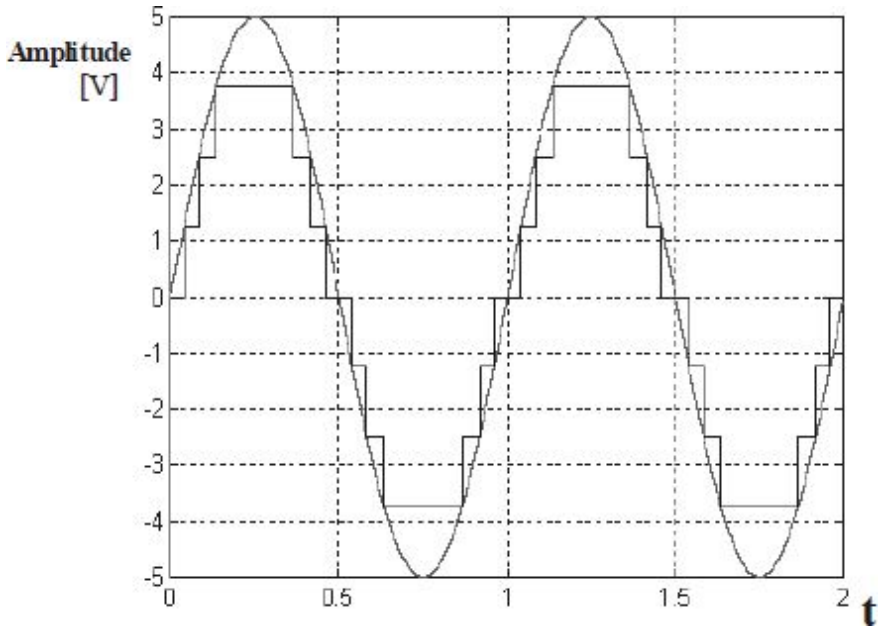


Fig. 10.17: Sinusoidal signal sampled with a 3bits quantization.

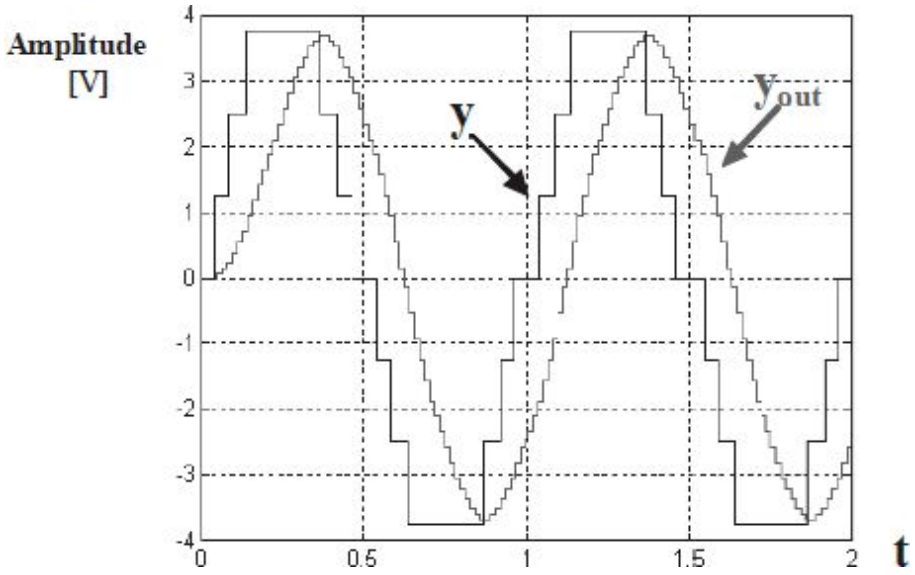


Fig. 10.18: Sinusoid sampled at $f_s=32f_0$ with a 3bit ADC and decimator output signal.

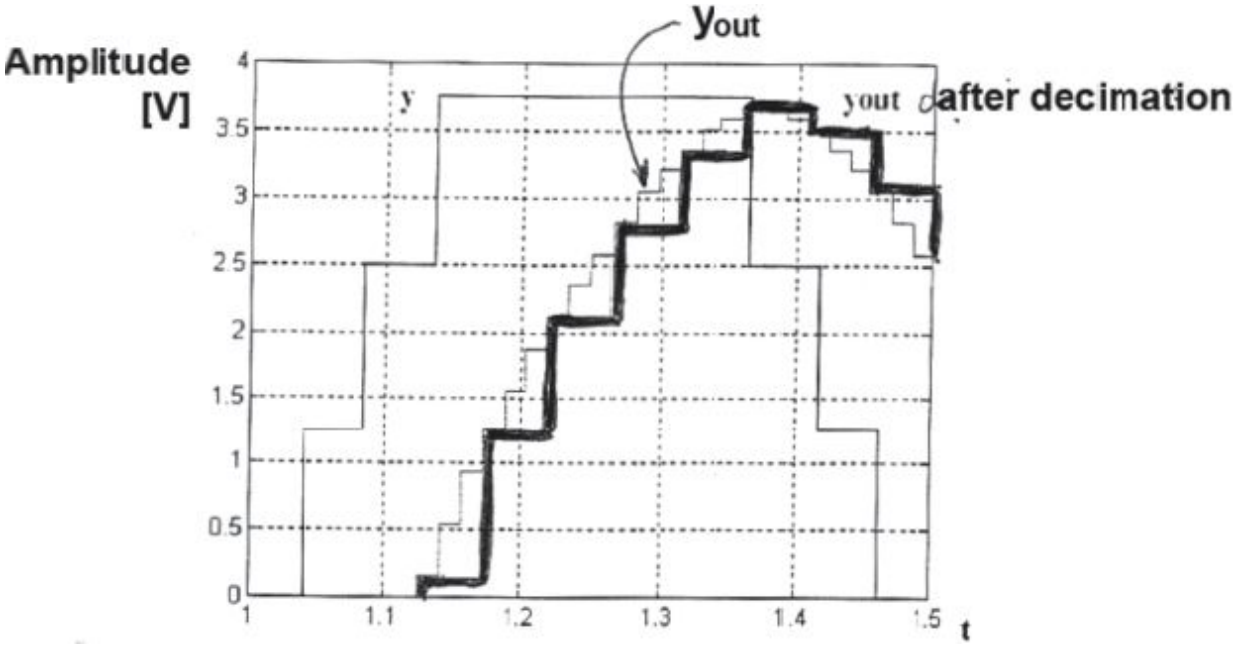


Fig. 10.19: Detail of the sampled sinusoid of Fig. 10.18.

The described technique is not efficient; in fact, to obtain a resolution improvement of just 2bits, we need to employ a sampling frequency 16 times higher than the Shannon limit. Moreover, sometimes the quantization

noise can not be considered as “white”. To cope with these problems, we need a new architecture able to get better SNR improvements at every doubling of f_S and able to uncorrelate the signal to the superimposed noise: this is effectively realized by the ∑Δ modulator.

10.3. ∑Δ MODULATOR

The idea at the basis of the differential modulation (or Δ modulation) is the following: with the transfer function of the signal, one can transmit the difference between two following samples, instead of their values, with a saving in the number of sent bits. In the simpler case, we can use only one bit, which can be either 0 or 1 if the preceding sample is lower or higher than the current one, respectively. Fig. 10.20 shows the basic scheme of the Δ modulator with X=analog input, Y=digital output (if X>W, Y=high, otherwise, Y=low), and W=reference analog signal.

For the system to work properly, one needs to operate at a sampling rate higher than the Nyquist limit in order to obtain a strong correlation between adjacent samples (the oversampling $f_s >> f_{Ny}$ is needed). The system will only transmit the sign bit, either positive or negative, of the difference between two adjacent samples. In the scheme of Fig. 10.20, we use a simple comparator on the forward branch with a quantization of only one bit while the feedback loop forces the integrator output W (shown as an adder) to follow the input. This corresponds to approximating the analog signal with a series of steps with constant amplitude Δ and transmitting the corresponding binary sequence; it is similar to the tracking ADC, in which the whole digital code is not transmitted, but only Y=high (rise) or Y=low (fall), converting n=1 bit.

What was said until now is shown in Fig. 10.20 (on the right) by means of the comparator and integrator output signal. As we can see, the series of output bits transmitted by the Δ modulator is related to the derivative of the original signal: if the input derivative is positive, the output is a series of ones; if the input derivative is negative, the output is a series of zeros; and if the input derivative is zero, the output has an oscillatory behavior. While requiring a very high operating frequency, the Δ modulator has important

advantages. The most important of these is the possibility of using only two quantization bits, simplifying the circuitry of the modulator and the receiver.

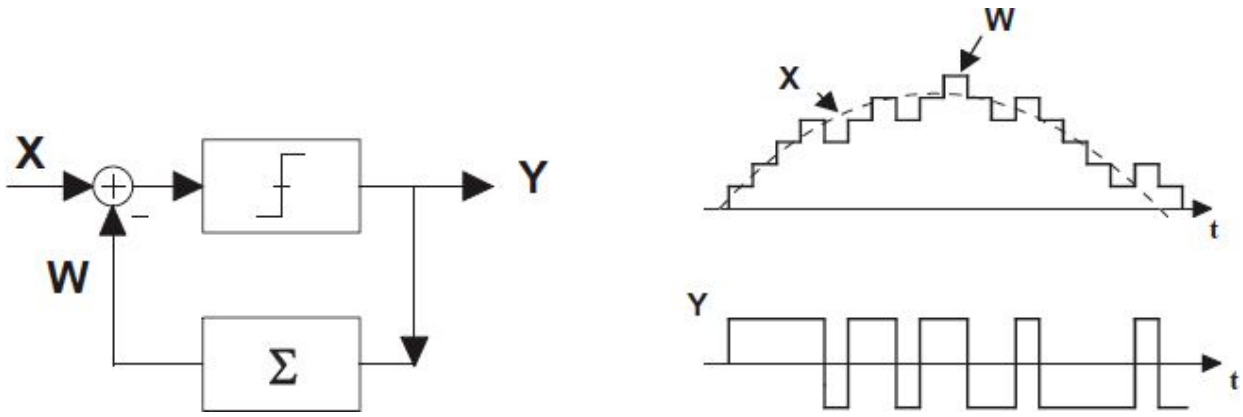


Fig. 10.20: Block scheme of the Δ modulator (on the left) and integrator and comparator output signals in a Δ modulator (on the right).

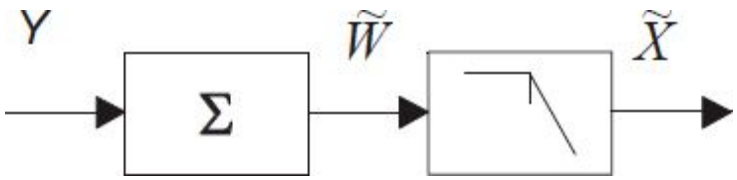


Fig. 10.21: Block scheme for a Δ modulation receiver.

The 'Σ' block is constituted by an accumulator (a circuit which accumulates the output Y pulses) providing an analog signal W. Also the receiver will be very simple (Fig. 10.21), which is constituted by the same adder used in the modulator loop (which, adding the received increments, provides the transmitted digital code) with a low pass filter connected at the output. In this way, we can find the starting analog signal X.

Note that the feedback loop eliminates the correlation between the signal and the quantization noise, introducing a constant signal; in fact, the output Y is a square-wave oscillating at half of the sampling frequency. We can think that, introducing a continuous signal, the output, and therefore also the quantization noise, is not constant.

Starting from the scheme for both the transmission system with Δ modulation (Fig. 10.20) and the receiver (Fig. 10.21), you can see the simple evolution towards the basic structure of the 'ΣΔ' modulator. The scheme of Fig. 10.22a is constituted by the simple union of the described blocks,

(modulator and receiver). Exploiting the linearity of the integration operation, it is possible to put the Σ within the chain without altering the signal transfer function (actually, we must slightly modify the circuit because the "digital" Σ must be modified in an "analog" ∫, as shown in Fig. 10.22b).

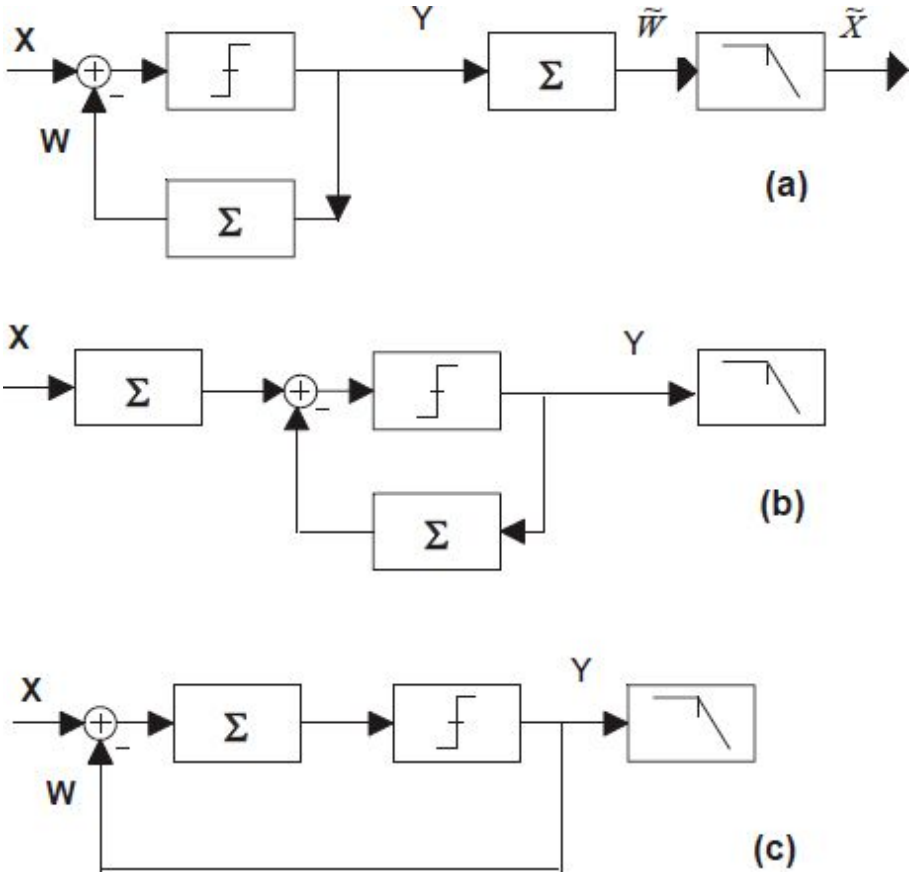


Fig. 10.22: a) Transition between the Δ differential modulation and the ΣΔ modulator. b) Δ modulator with a reconstruction filter (integrator and low-pass filter). The integration filter anticipates the quantization phase. c) ΣΔ modulation.

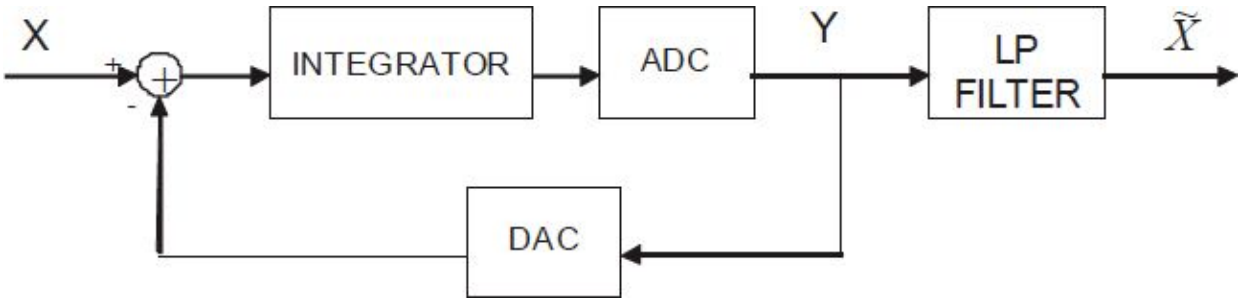


Fig. 10.23: Block scheme for a ΣΔ converter.

The adder inside the loop can be viewed as an integrator applied to a DAC converter. Because the two integrators converge to the same input differentiator node, they can be placed in the same point within the loop, obtaining the scheme shown in Fig. 10.22c, which represents the ΣΔ modulator. We must highlight that the steps taken to evolve the circuit in Fig. 10.22a to the final configuration are correct only if the blocks are perfectly linear.

In the final ΣΔ modulator configuration, the input signal is processed by a quantizer after the integration is performed by the block Σ. The quantized output is reported to the input by means of feedback and subtracted from the input signal; subsequently, the low-pass filter performs the reconstruction of the original signal. The feedback forces the quantized signal *Y* to follow the mean value of the input *X*. Moreover, the loop contained in the circuit works as a modulator of the quantization noise because it reduces its presence within the band, shaping the spectrum to the higher frequencies (this effect is named *noise shaping*). In addition to the modulation, the circuit exploits the advantages of quantization noise reduction (within the bandwidth) given by the over-sampling, obtaining a good A-to-D conversion technique.

There are other circuitry solutions for the ΣΔ modulator in which, instead of a simple comparator, an ADC converter with many quantization levels is used, as shown in Fig. 10.23. In this case, the project is complicated because one needs to put a D-to-A converter on the feedback branch whose resolution is equal to or higher than the modulator resolution. Moreover, such a device, because of the position within the loop, requires a great effort in terms of design because there is not the feedback reduction in the non-idealities.

The comparator output is a coarse quantized signal that is different from the input signal. The error between the input and the output signals is indicated with ‘e’. Assume that the comparator is linear and has a gain equal to 1. The integrator is an adder with this transfer function $1/(1-z^{-1})$ and with a delay z^{-1} on the forward path, as shown in Fig. 10.24.

Avoiding the non-linearity problems in multi-thresholds devices, in the ΣΔ modulator, there are only two quantization levels. Therefore, in the following mathematical analysis about the ΣΔ modulator, we are going to

use the model depicted in Fig. 10.22c, in which the quantizer is a simple comparator. Because it is a discrete-time circuit, the integrator is constructed by the adder and the forward path delay while, for the comparator, we will base our analysis on the simple hypothesis of unity gain, representing it with only the introduced error *e*. The scheme is shown in Fig. 10.24. By using the Z-transform, it is simple to obtain the transfer function of the signal X(z) and the quantization noise e(z) as:

$$Y(z) = X(z)\frac{1}{1+\frac{1-z^{-1}}{z^{-1}}} + e(z)\frac{1-z^{-1}}{1+\frac{1-z^{-1}}{z^{-1}}} = X(z)\cdot z^{-1} + e(z)\cdot\left(1-z^{-1}\right)$$

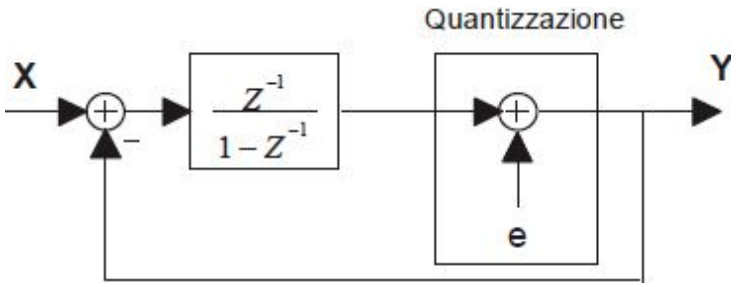


Fig. 10.24: ΣΔ modulator linear model.

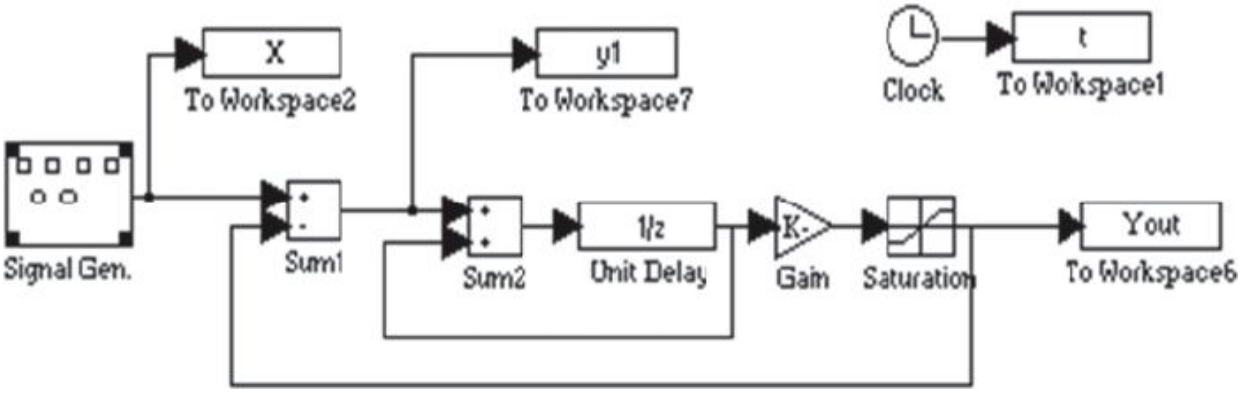


Fig. 10.25: Simulink implementation of the ΣΔ modulator.

As one can see starting from the preceding relationships, the signal is transmitted to the output (except for the propagation delay) unchanged while the quantization noise undergoes attenuation (derivative operation: current error E(z)·1 minus preceding error E(z)·z^{-1}): there is a rise in the high

frequency components and attenuation in the low-frequency components. To obtain the linearized model in Fig. 10.24, we supposed that the noise had a uniform spectral density and was uncorrelated with the signal. Such approximations are not addressed analytically because of the presence in the loop of a non-linear block, the comparator.

To evaluate the deviations introduced in the linearized model and verify the approximations, one could perform a mathematical analysis and compare it with the results of a circuit simulator (which can evaluate the effect of the non-linearity). We can employ the simulation tool *Simulink* of MATLAB. The block scheme which models the ΣΔ modulator is shown in Fig. 10.25. In the implemented model, various waveforms are made by the special block connected to the input (*Signal Gen.*). After that, there is the loop that implements the modulator: *Sum1* allows the comparison between the analog input signal and the digital output; the blocks *Sum2* and *Unit Delay* implement the integrator while *Gain* and *Saturation* make the comparator model. The amplification introduced by the block gain *Gain* is necessary to elevate the signal over the maximum dynamic allowed by *Saturation* so as to always have two quantization levels.

We can evaluate the processing results with a sinusoidal signal, whose amplitude is 300mV, which has a frequency of 0.2Hz, and which is simulated with MATLAB in the model depicted in Fig. 10.25. For the comparator, a dynamic of 1V is chosen while the sampling frequency used in the modulator is 16 times than the Nyquist one (i.e. $f_s=16\cdot2\cdot0.2=6.4$ Samples per second). The obtained output is shown in Fig. 10.26, on the left, in which one can note the highest density of commutation around the input signal's highest slope regions (highest derivative).

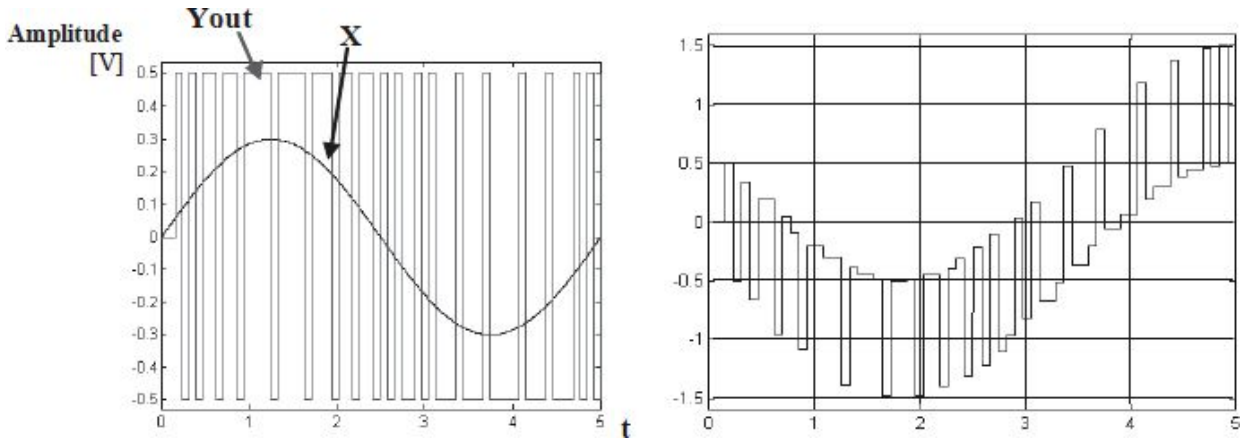


Fig. 10.26: Input and output signals (left) of the modulator shown in Fig. 10.25 and quantization noise superimposed on the output signal (right).

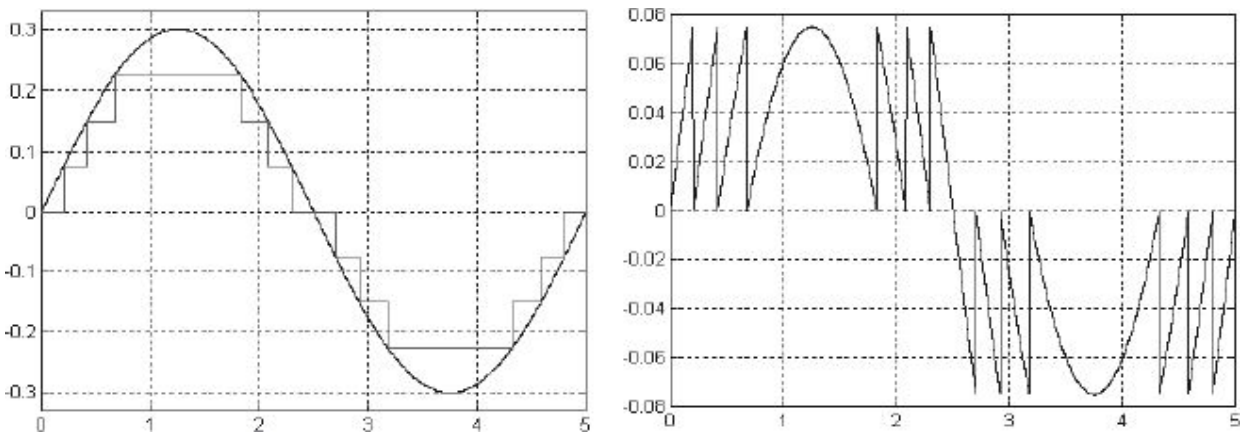


Fig. 10.27: 3bits quantization (OS=16) for a sinusoidal signal (left) and its quantization error (right).

In Fig. 10.26 (on the right), the trend of the error made on the output signal due to the quantization is shown. The noise is more intense, around 1V, but has fast components (close to the sampling frequency and so much higher than the original signal components). This feature allows considerably reducing the intensity of this disturbance on the low-frequency output (there is not quantization disturbances) with a simple filtering operation.

To understand the effectiveness of the differential modulator, one can compare the graph in Fig. 10.27a with that in Fig. 10.27b. The latter shows the error made with an over-sampling factor equal to 16 (OS=16), but using a simple 3bits quantization (3bits non-ΣΔ ADC). In this case, the quantization error is more intense because of the conversion with higher resolution (3bits instead of only one), but has frequency components uniformly distributed on the range $0 \div f_s$, which are indistinguishable from the useful signal in Fig. 10.27a. Using the ΣΔ modulator, there are more abrupt transitions, particularly at high frequencies, around $f_s/2$ (Fig. 10.27b).

We graphically saw how the quantization error is derived if compared to the original spectrum E(z). To understand the delay effect on the signal X(z), we can remove the effect due to the quantizer (comparator), closing the loop directly with the integrator output, observing the output waveform $Y(z)=X(z)z^{-1}$. With the new circuit shown in Fig. 10.28, the trend for the

input and output signals X and Y are those shown in Fig. 10.29. As we can see, Y_{out} is the sampled input sinusoid with a delay equal to δ. Because the ratio between the sampling frequency and the Nyquist one is equal to 16 in this case, the delay introduced by the modulator processing is equal to 1/32 of the sinusoid period T.

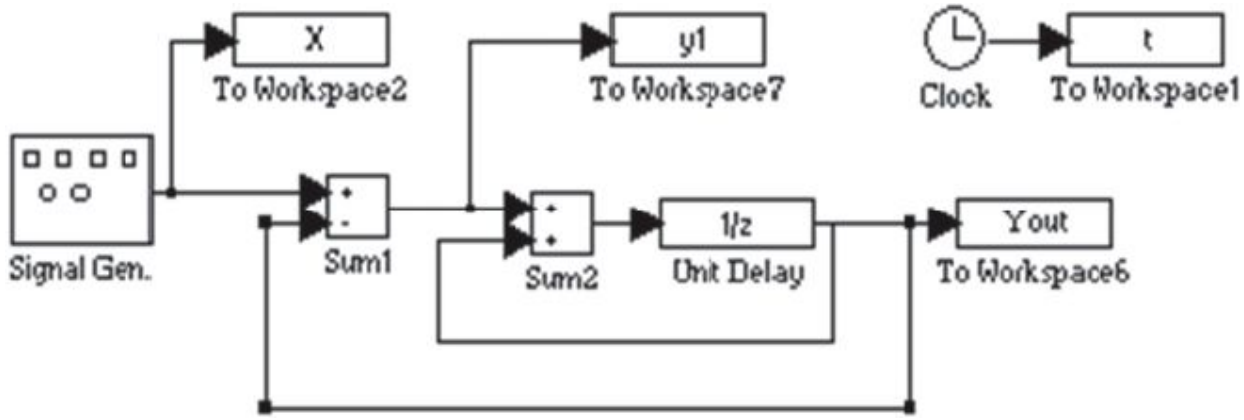


Fig. 10.28: ΣΔ modulator model in which the quantization effect is removed (to verify that the modulator works properly only on the signal).

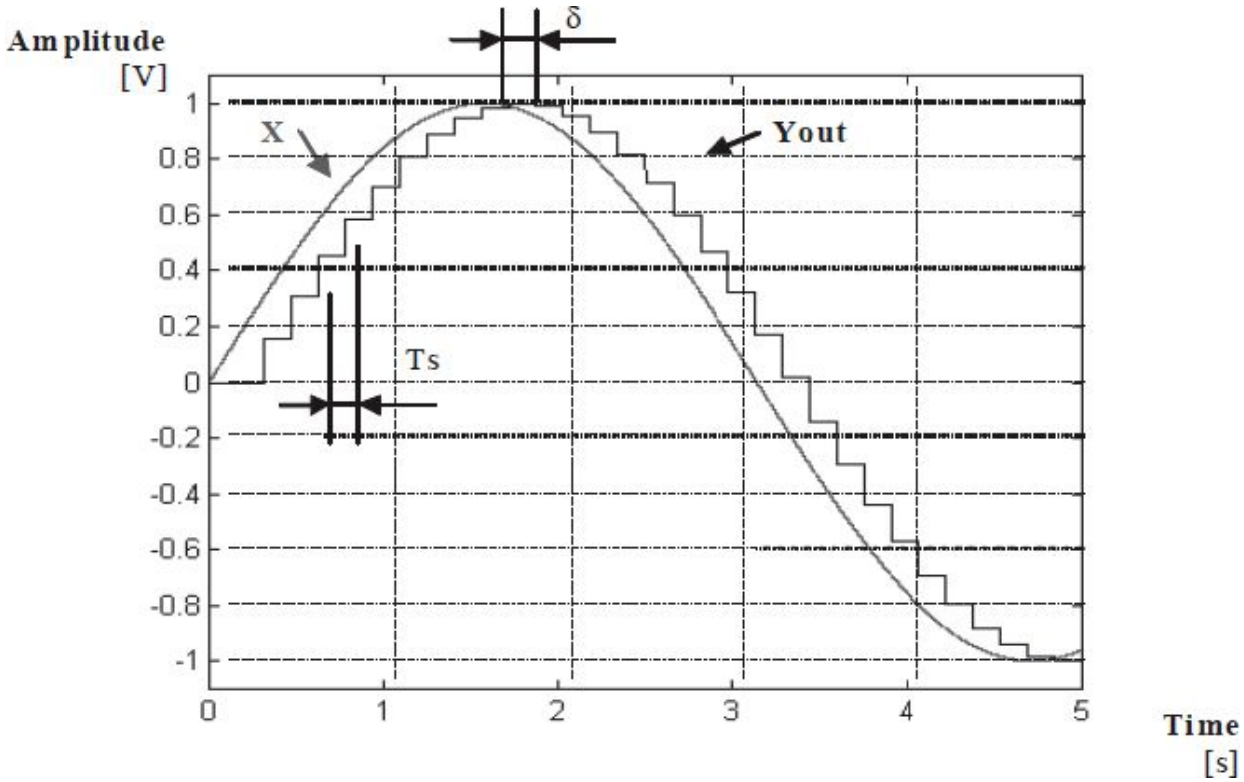


Fig. 10.29: Delay and discretization on the input analog signal, which is realized by the modulation loop, without the error introduced by the quantization.

The introduced delay δ (equal to 1) is equal to the sampling time $T_s=1/f_s$. Choosing, in this example, OS=16, there are 32 samples per period and therefore δ=T/32=1/32f=156ms.

We can quantitatively analyze the effect due to the ΣΔ modulation on the quantization noise. Assume that the noise E(z) is white with a constant power spectral density $Q_p(f)$. Converting to the frequency domain by substituting $z=e^{j2\pi fT}$, we obtain the output noise spectral density of the ΣΔ modulator $E(z)\cdot(1-z^{-1})$, as shown in the equation (*). The *'noise shaping'* effect made by the modulator on the quantization error is more evident, taking into account the output noise spectrum. Assuming that the quantization noise spectral density is constant in frequency (this hypothesis true in most cases), its frequency domain trend at the modulator output is the following:

$$Q(f) = Q_0\left|1 - e^{j2\pi fT}\right|^2 = 4Q_0 \sin^2(\pi fT) \qquad (*)$$

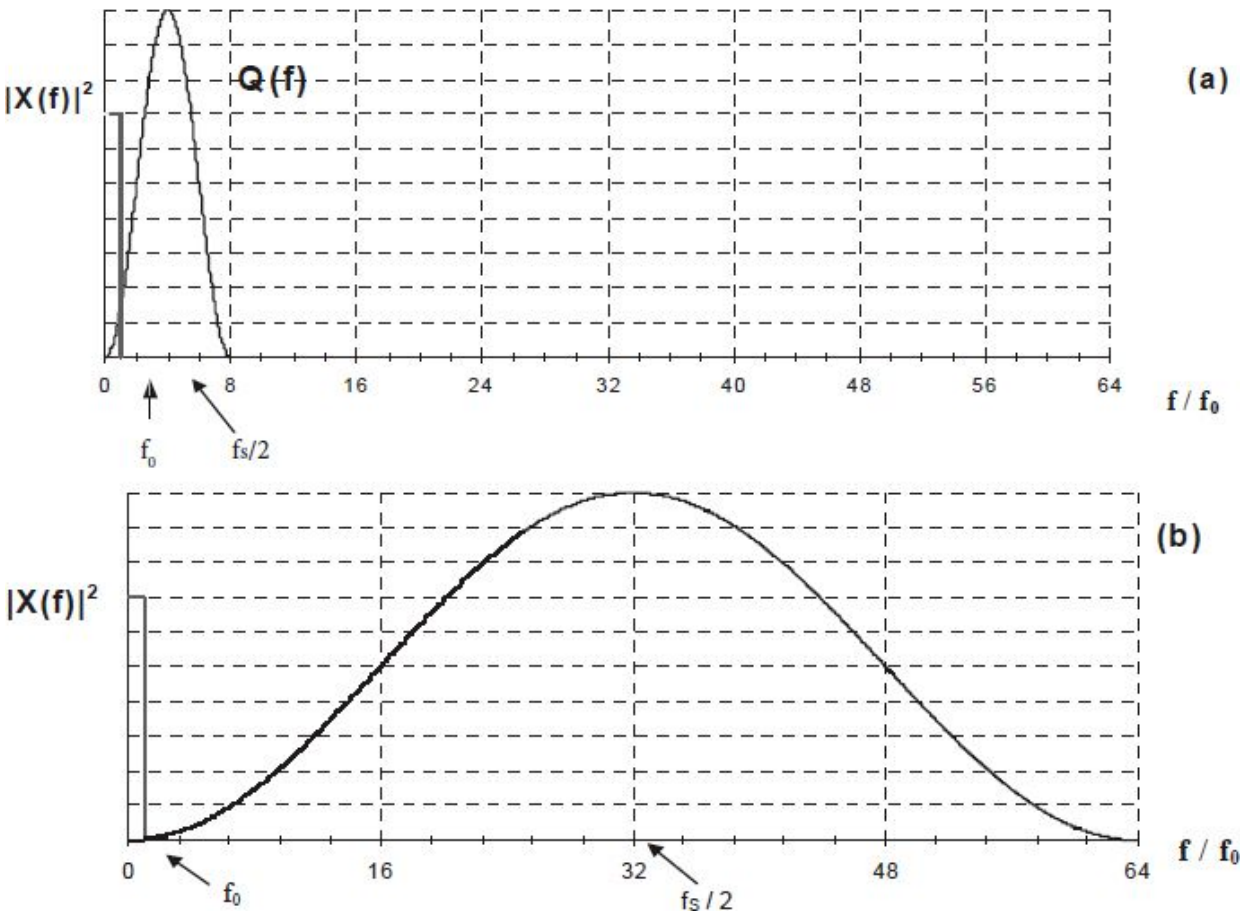


Fig. 10.30: Spectrum of the signal /X(f)/ at the ΣΔ modulator output for two different sampling frequencies (OS=4 above and OS=32 below). There is a minimum for f=0 and for f=f_s and a maximum for f=f_s/2 (the area of Q(f) is constant in both cases).

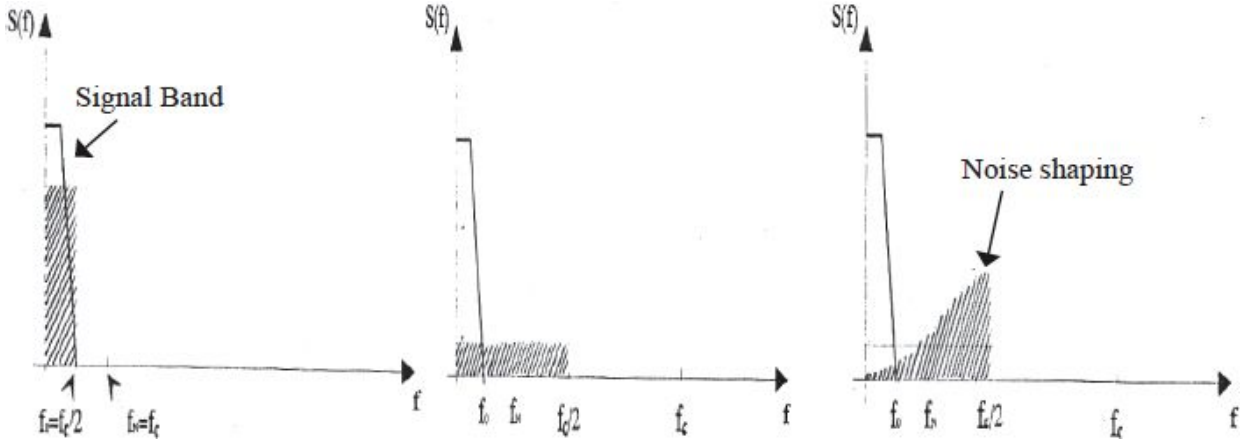


Fig. 10.31: (a) Nyquist sampling. (b) Over-sampling (see the noise reduction in the desired band). (c) Over-sampling and ΣΔ (see the out-of-band noise shaping).

Fig. 10.30 shows the power spectral density for two different over-sampling factors: OS=4 (a) and OS =32 (b). We can appreciate how the feedback reduces the noise presence at low frequencies, increasing it at high frequencies (with a minimum at f=0 and f=1/T_s= f_s). There is a maximum at f_s /2 in the noise shaping. Note that, increasing the over-sampling factor, the quantization error is pushed towards the high frequency (in respect of the low frequency).

From the point of view of the quantization noise, the ΣΔ modulator improves the performance levels obtainable by the over-sampling, thanks to the out-of-band noise-shaping. One can see in the diagrams of Fig. 10.31 that the noise is translated towards the high frequencies, increasing the oversampling factor.

From the preceding graphs, we see how the in-band ($0 \div f_0$) quantization noise is much lower than the value $Q_0(f) \cdot 2f_0$ (for a simple Nyquist sampling) and smaller for a higher sampling frequency. Assuming that the quantization noise is white with a spectral density equal to: $Q_0 = \frac{\Delta^2}{12}\frac{2}{f_s}$ (con Δ=LSB).

(con Δ=LSB).

To effectively evaluate the noise power superimposed on the signal within the useful bandwidth, we can integrate Q(f) between 0 and f_0 with the following result:

$$q^2 = \int_0^{f_0} Q(f)df = \int_0^{f_0} \frac{\Delta^2}{12}\frac{8}{f_s}\sin^2(\pi fT)df \approx \frac{\Delta^2}{12}\frac{\pi^2}{3}\frac{1}{OS^3}$$

Now, we can see the evident advantages of the ΣΔ modulation; in fact, for every sampling frequency doubling, there is in-band noise reduction of 9dB, i.e.:

$$SNR_{out} = SNR_{in} \cdot \frac{3}{\pi^2} \cdot OS^3$$

Obviously, to fully exploit the noise reduction, we must employ a digital low-pass filter downstream the modulator with a cut-off frequency f_0.

For a more general discussion of the effects of differential modulation, we have to consider the ADC on the forward branch as a '*c*' bits quantizer (not a simple comparator). This condition helps increment the obtainable resolution because, with increasing 'c', the LSB is reduced to $FSR/2^c$ (where the FSR is the input signal maximum dynamic), reducing the quantization noise power.

With a sinusoidal signal whose amplitude is equal to FSR/2 as the converter input, the SNR is increased up to the value:

$$SNR = \frac{FSR/2\sqrt{2}}{\frac{FSR}{2^c}\frac{\pi}{(OS)^{3/2}\cdot\sqrt{36}}} = \frac{\sqrt{18}}{\pi}\cdot(OS)^{3/2}\cdot 2^{c-1} \qquad (**)$$

The dB result is: $SNR_{dB} = 20\log_{10}(2^c) + 20\log_{10}\left[(OS)^{3/2}\right] - 3.41 = 6.02(c+1.5L) - 3.41$ where we chose an OS multiple of 2, i.e. $OS=2^L$. Therefore, the increase in the number of bits for the ADC causes an improvement in the SNR of about 6dB while the sampling frequency doubling increases the SNR by 9dB.

At the output modulator, we have only *c* bits. And then, what is the SNR_{out} dB higher than the usual 6.02·c dB? Downstream the modulator, we need a digital filter which, starting from *c* bits (the minimum 1 for the case seen until now) with a frequency f_s, can provide to the output with digital *words* with *n* bits.

Therefore, the length of the digital filter output word will be:

$$b = \frac{SNR - 1,76}{6,02}$$

which is the final resolution of the whole ΣΔ converter, i.e. the equivalent number of bits which can describe that signal. If the number of bits n>b, it means that only the first *b* will be the most significant (there are many unnecessary bits) while, if n<b, I have not fully exploited the ΣΔ advantage (there is a bad filter, and we obtain a step sinusoid).

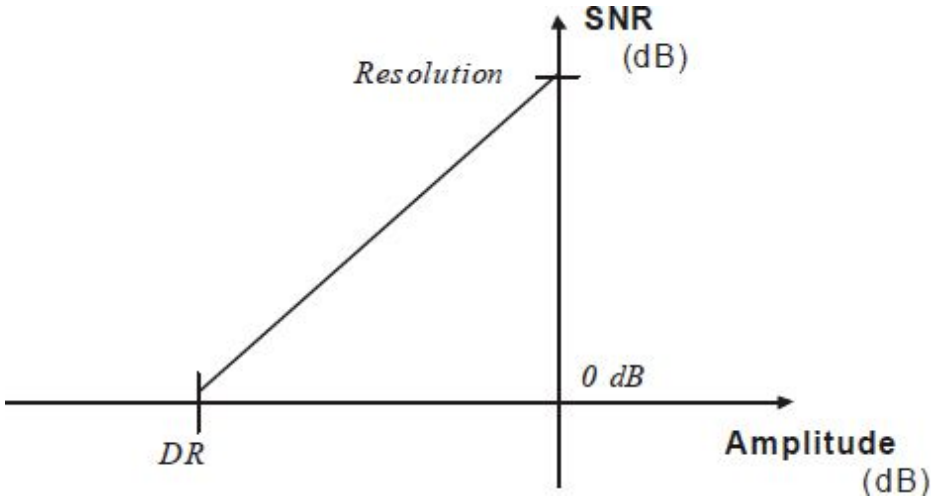


Fig. 10.32: Ideal modulator output SNR vs. the input sinusoidal amplitude.

We have an increase in the resolution in respect of a common over-sampling conversion not only because, with an equal SNR, we obtain a higher *b* value, but also because the SNRs in the formulas are different. Particularly, the SNR obtained with the ΣΔ modulation is higher (grows by 9 dB for every sampling frequency doubling, compared with the 3dB increase with a simple oversampling). Moreover, with the ΣΔ modulation, we can act on the ADC (contained in the feedback loop) to increase the overall resolution. In fact, we obtain a resolution increase of 1,5+c bits for every OS doubling, where *c* is the ADC resolution. Instead, the normal oversampling allows obtaining only ½ bit for every OS doubling.

For example, using a 2bits ADC (4 discrete levels) and an OS equal to 16 (L=4), we obtain an SNR of 44.75dB, corresponding to 8bits resolution. The excellent performance levels are due to the high working frequencies employed. With an input signal whose bandwidth is 20kHz, we need to use a ΣΔ converter with a f_s of $16 \cdot 2f_0$=640kHz while, with a simple converter, we could work at f_s=40kHz with an 8bits ADC (256 discrete levels). The ΣΔ modulator is very simple because it uses only 2bits, instead of 8bits of a classic cheaper ADC, and simpler compared with a 16bits ADC for high resolution. Moreover, the ΣΔ modulator (high error at high frequencies) is the alternative to obtain resolution otherwise cannot be obtained.

The results given here were obtained in ideal operating conditions, i.e. neglecting the presence of any noise contribution (for example electronic noise, charge injection of the devices, etc.), except for the quantization effect

which degrades the SNR. In literature, the ΣΔ modulator's performance levels are shown in a bi-logarithm graph which represents the SNR (experimentally measured) versus the input signal amplitude. For the ideal modulator, the graph is a straight line with a unity slope (Fig. 10.32). In fact, this trend is obtained by analytically substituting Vin/2 to the numerator in the (**), where Vin is the input sinusoidal signal dynamic.

$$SNR = \frac{\frac{Vin}{2}}{\frac{V_{fs}}{2^c}\frac{\pi}{(OS)^{3/2}\cdot\sqrt{36}}} = Vin\cdot\frac{\sqrt{36}}{\pi}\cdot\frac{(OS)^{3/2}\cdot 2^{c-1}}{V_{fs}}$$

Therefore, in the ideal working conditions of the modulator, the maximum SNR is reached with the maximum input signal. Such a value is the device *resolution* and is the value computed in the (**). On the abscissa axis, it is possible to obtain the minimum signal that the system can convert, which conventionally is the amplitude of the sinusoid that makes the SNR equal to 0dB.

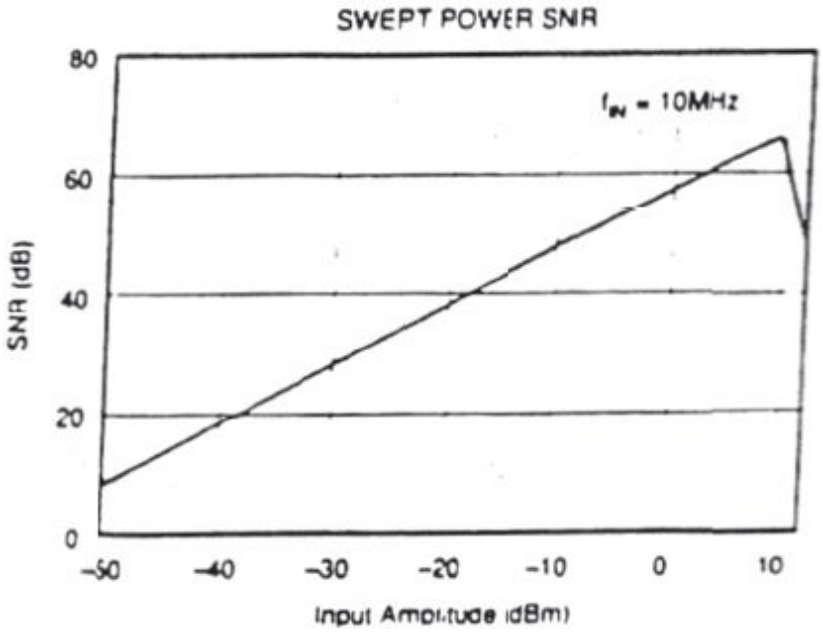


Fig. 10.33: Ads801 converter SNR vs. amplitude.

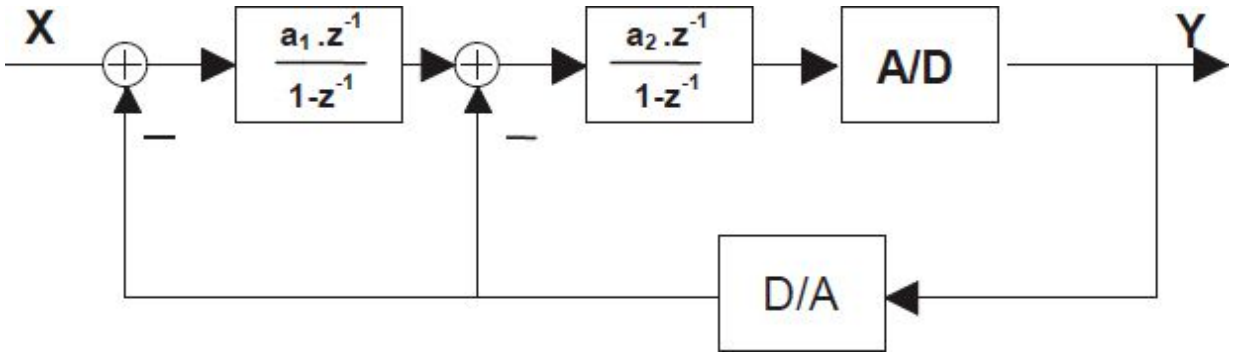


Fig. 10.34: Block scheme of a second order ΣΔ modulator.

The ratio between the maximum and the minimum *dynamic* amplitude is the DR (Dynamic Range). For the aforesaid reasons, within the ideal modulator, the dynamic and the resolution are numerically equal (SNR_{out}dB=DR) while, actually, the resolution is the maximum for amplitude values lower than that of the maximum input admitted. Fig. 10.33 shows the relationship between the SNR and sinusoidal amplitude.

10.4. II ORDER ΣΔ MODULATOR

Until now, we implicitly treated the first order modulator because the order is defined by the order of the feedback integrator. The scheme for a second order system is shown in Fig. 10.34. The presence of the second feedback loop allows obtaining advantages as the voltage dynamic limitation in the output nodes and a further decrease in the quantization noise. With two integrators a_1 and a_2, respectively equal to 1 and 2 (considering the stability of the whole architecture), the signal and noise transfer functions are equal to:

$$Y(z) = X(z) \cdot z^{-2} + e(z) \cdot \left(1 - z^{-1}\right)^2$$

The signal is unchanged, but it has a phase shift equal to twice the sampling period, while the noise is shaped with a second order derivative which further decreases the in-band noise power. In fact the output error sees two integrations and it is shaped as a $\sin^4$, as the following formula shows: $Q(f) = 16Q_0 \sin^4(\pi fT)$

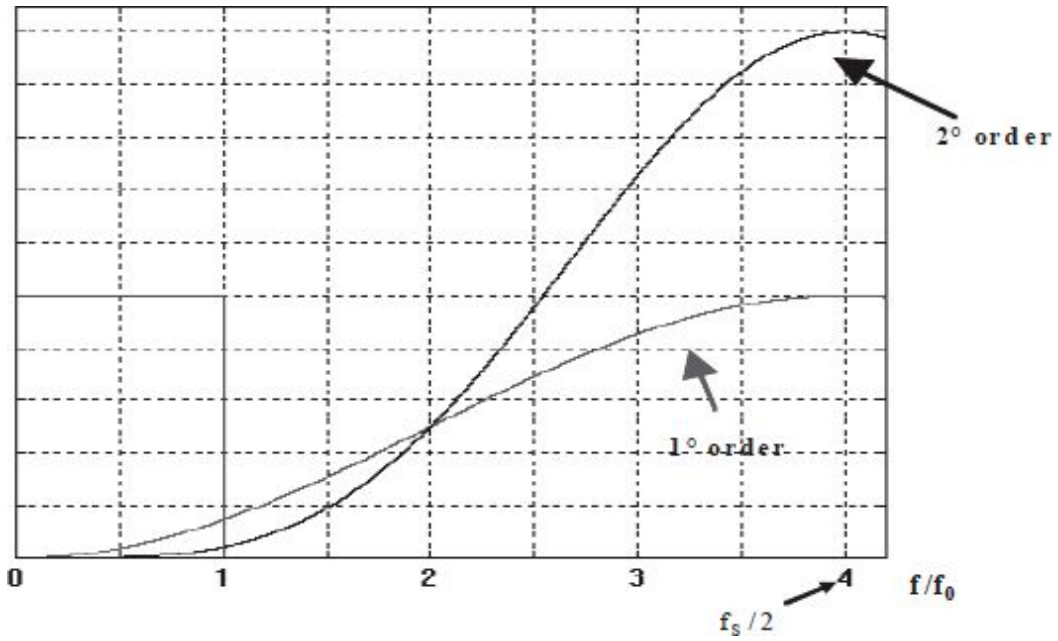


Fig. 10.35: "Noise shaping" for first and second order modulators.

The in-band noise power is reduced and is equal to:

$$q^2 = \int_0^{f_0} Q(f)df \approx \frac{\Delta^2}{12}\frac{\pi^4}{5}\frac{1}{OS^5}$$

With a sinusoidal signal with a high dynamic and supposing to employ (in the forward branch of the modulator) a *c* bits ADC, we can obtain a higher SNR equal to:

$$SNR = \frac{V_{fs}/2\sqrt{2}}{\frac{V_{fs}}{2^c}\frac{\pi^2}{(OS)^{5/2}\cdot\sqrt{60}}} = \frac{\sqrt{15}}{\sqrt{2}\pi^2}\cdot(OS)^{5/2}\cdot 2^c \qquad (***)$$

Expressed in dB, substituting OS with the expression 2^L, we have:

$$SNR_{dB} = 6.02(c + 2.5L) - 11.14$$

i.e. for every f_0 doubling, we have an increase of 15dB in the SNR, i.e. equal to 2.5bit.

The final resolution for a second order ΣΔ modulator is equal to:

$$b = \frac{SNR - 1.76}{6.02}$$

Such a modulator is more complex, but has a $\sin^4$ noise-shaping (the higher the frequency, the higher the noise), as you can see in the comparison in Fig. 10.35.

As for the first order modulator, for every additional bit, the total system resolution is increased. Nevertheless, in the case of the second order modulator, doubling the sampling frequency f_s, the noise power is reduced by a factor of 32, corresponding to a gain of 2.5bit. This higher resolution, due to the decrease in the in-band quantization noise, can be viewed also in the Fig. 10.36, where we see the noise-shaping made by the modulator. In this case, we used an OS factor equal to 4.

The use of two integrators within the feedback loop allows reducing the in-band residual quantization noise, increasing the resolution (2.5bits more for every OS doubling). Nevertheless, with a modulator order higher than 1, it is possible to have a problem due to signal divergence of the integrator output for some input values, as, for example, in the case of constant signals.

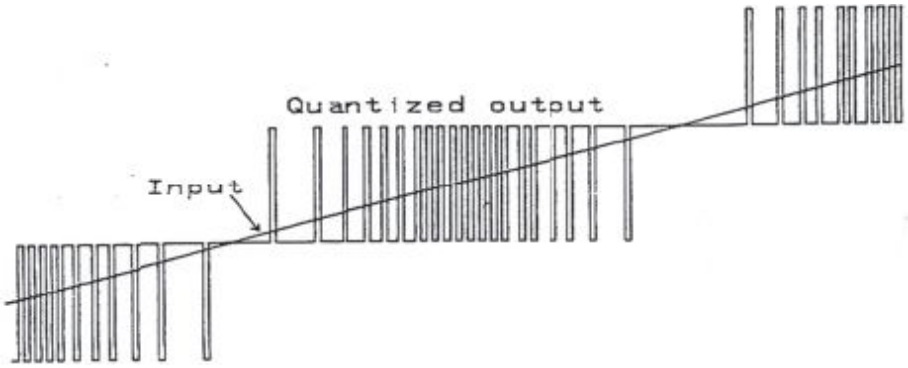


Fig. 10.36: I order ΣΔ modulator response for a ramp signal.

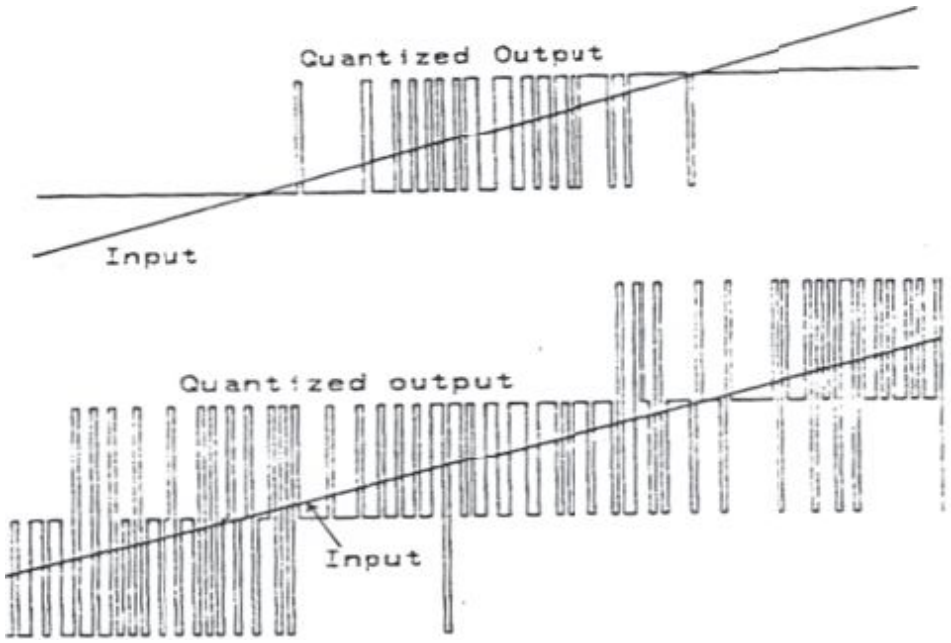


Fig. 10.37: 1bit (at the top) and 2bits (at the bottom, II order ΣΔ response for) for a ramp input signal.

```
0 2 1 2 1 | 2 1 1 2 0 2 2 0 2 2 0 2 1 2
1 1 1 2 1 | 2 0 3 0 2 1 1 2 1 1 2 1 1 1
2 1 2 0 2 2 0 2 1 2 0 3 0 2 1 1 1 2 2 0
2 1 1 2 1 | 2 1 1 1 2 1 2 0 2 1 2 1 1 2
0 2 2 0 2 2 0 2 1 2 1 1 1 2 1 1 2 1 1 1
2 1 1 2 1 | 2 1 1 1 2 1 2 0 2 2 0 2 1 2
1 1 1 2 1 | 1 2 2 0 2 1 1 2 1 1 2 1 1 1
2 1 2 0 2 2 0 2 1 2 0 2 2 1 1 1 1 2 1 2
1 1 1 2 1 | 2 1 1 1 2 1 1 2 1 1 2 1 1 1
2 1 2 0 2 2 0 2 1 2 1 1 1 2 1 1 2 0 3 0
```

Fig. 10.38: ΣΔ modulator response for a constant input of 1.3V.

For example, we can show the typical 2bits I order ΣΔ modulator output in Fig. 10.36 (4 quantization levels: 00 01 10 11 at the output) during the linear ramp sampling. In Fig. 10.37 is shown the II order 1bit ΣΔ modulator (at the top) and 2bits (at the bottom). With 2bits with the same frequency f_s, we have much noise, but it is uncorrelated with the signal (steps from the level 0 to the level 2 2, 0-2, 0-2 …), pushing it to a high frequency. Also with a 1bit, we have an improvement.

Fig. 10.38 highlights the 2bits II order ΣΔ output modulator, in which, since the input has a value equal to 1.3V, the output moves on 4 possible

levels to ensure the desired mean value (precisely 1.3V). Now, we are going to explore in depth the behavior of the I order modulator with constant input amplitude of 0.3V (while the comparator maximum dynamic is equal to V_{FS}=0.5V) and a sampling frequency of 640kHz. The output signal is not constant, but will oscillate between $+V_{FS}$ and $-V_{FS}$ to have the correct mean value 0.3V.

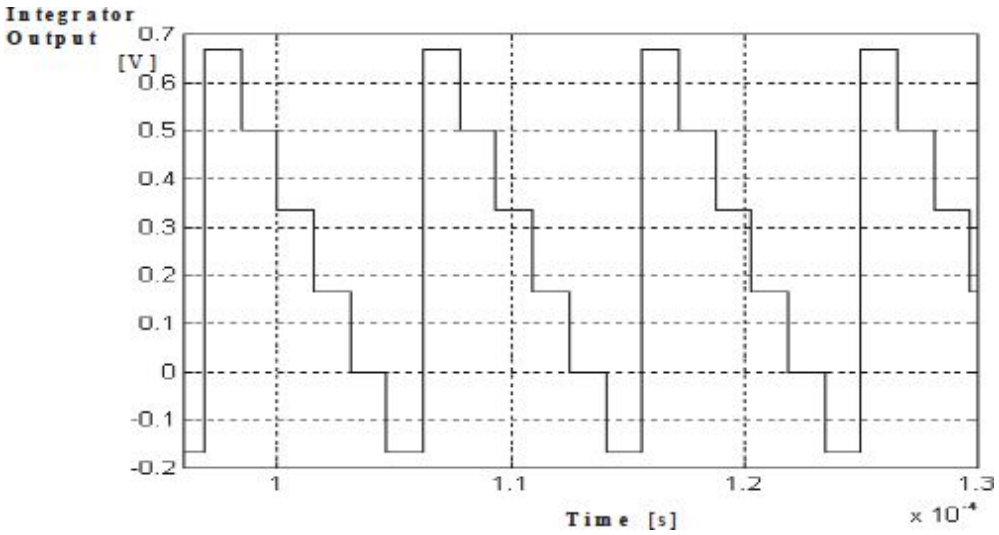


Fig. 10.39: Integrator output signal of a I order modulator.

Analyzing the integrator output shown in Fig. 10.39, we can notice that it will exceed the value V_{FS}. Because the voltage actually cannot overcome the supply voltage, in these cases, we will incur the voltage saturation phenomenon (clipping) with distortion in the original information. To avoid such occurrences, we must limit the input signal amplitude well under the maximum dynamic allowed by the power supply.

The system's working conditions must be studied in depth to ensure the "stability" for every input signal, avoiding oscillations (higher than the power supply dynamic) for the integrator output. Nevertheless, the analysis of stability is complex because of the non-linearity of the system. In any case, to do this analysis, we must substitute the quantizer with a linear block, using a strong approximation. The stability problem is addressed with high level simulations which verify the modulator performance levels for every working condition.

To avoid the integrator saturation, we must limit the input signal dynamic by means of the value a_1. Studying the linearized model for a second order circuit, we can verify that the closed loop poles are stable (within the unity circle) only if $a_2 > a_1$. Nevertheless, from the simulation results, we see that using $a_2 > 1.25 \cdot a_1$ is the best solution for a single bit modulator; otherwise, we can have very high oscillations.

As an example, we can implement with Simulink a second order modulator model, as shown in Fig. 10.40, where the single block function is similar to that of the first order modulator.

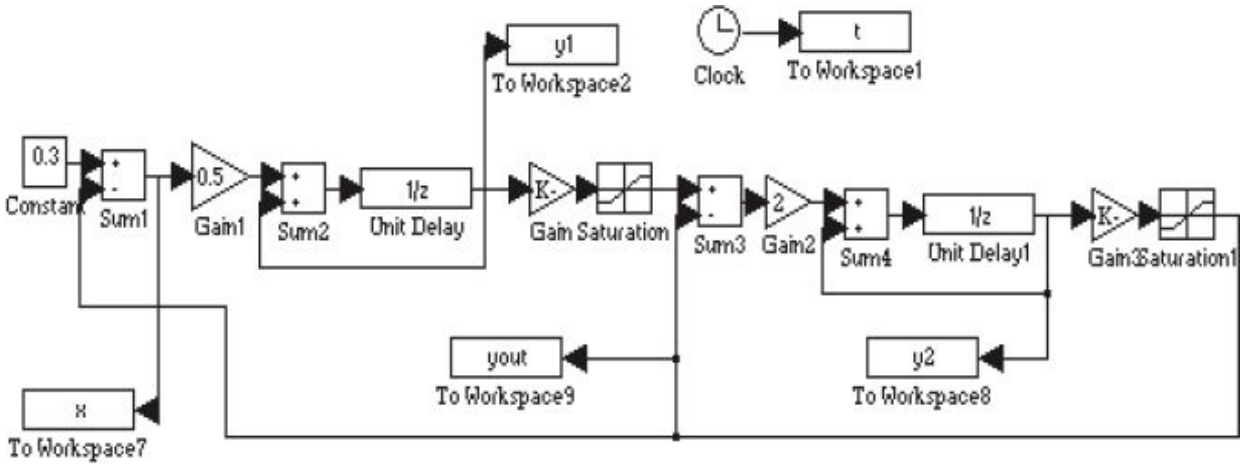


Fig. 10.40: Simulink model of a second order modulator.

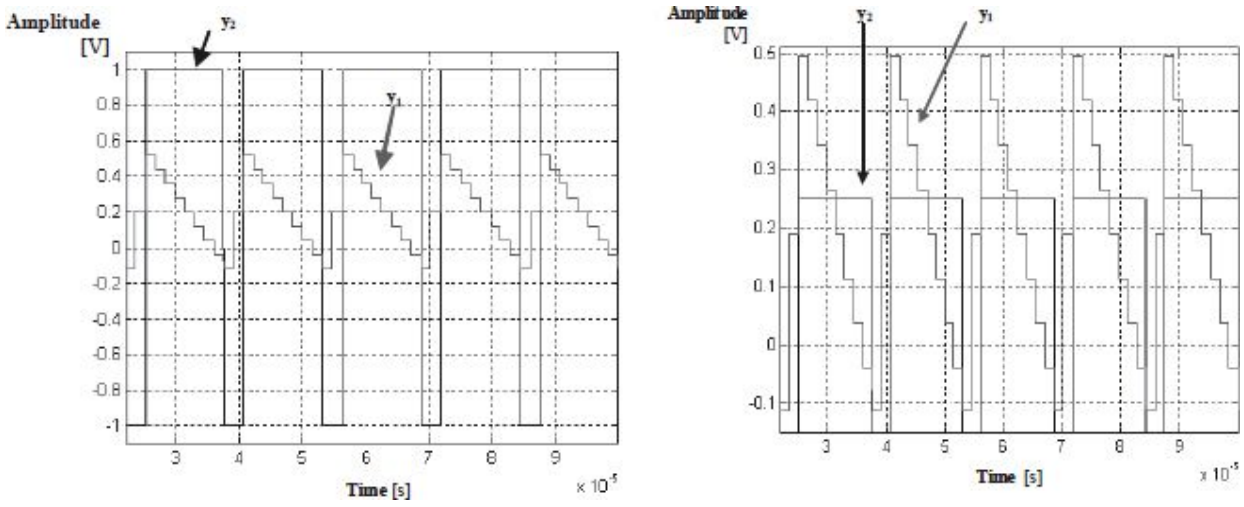


Fig. 10.41: Integrator output voltages of a second order modulator (on the left) and with a gain equal to 2 (on the right).

Connecting a constant signal whose amplitude is 0.3V, the output voltages' trends are shown in Fig. 10.41. One can see that the dynamic of the first integrator is lower than the first order circuit dynamic for an equal input signal; nevertheless, the output of the second integrator must overcome the available dynamic.

In fact, a second order modulator with a comparator as an ADC has the same stability problems for a first order architecture, and therefore the gains must be modified (in respect of the values a_1=0.5 and a_2=2) to ensure the whole architecture stability. Particularly, reducing the a_2 gain value, we will obtain a decrease in the comparator amplitude, as can be observed in the graph Fig. 10.41, on the right (with a_2=0.5).

Obviously, we can use noise-shaping order greater than the second; however, these modulators have great stability problems. Stability is ensured with ancillary circuitry, but the price is higher area consumption. For example, we can analyze the III order modulator with the linear model, obtaining the closed loop poles' positions, deducing that the system is unstable for some integrator gain values.

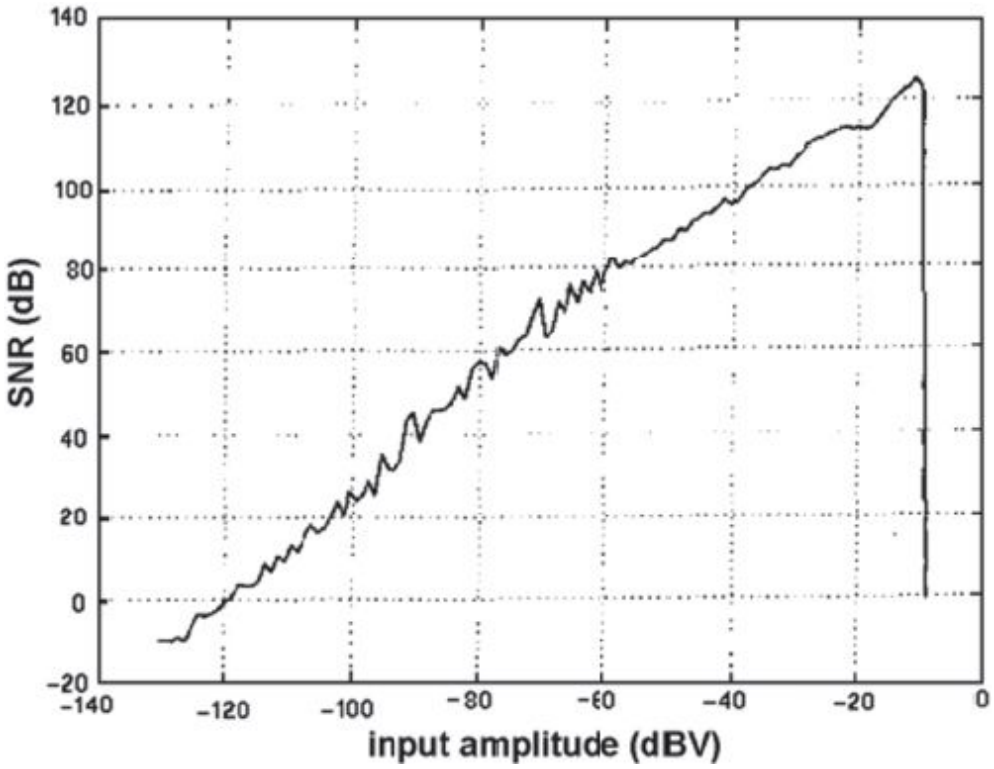


Fig. 10.42: II order ΣΔ output SNR.

To sum up, the integrator dynamic, and equivalently the dynamic of the quantizer, in the first and second order modulators must be greater than the maximum dynamic of the input signal. Otherwise, when the signal approaches to the dynamic limits, the quantization noise will increase (the input signal amplitude will increase, but not the quantizer output because it saturates). For this reason, as regards the graph in Fig. 10.42, for a two levels second order system, the last 6dB of the dynamic presents a “vertical” drop in the SNR.

10.5. ΣΔ Problems

10.5.1 Pattern Noise

One can study in depth the case of a constant signal connected to the input of the modulator; this will not provide a constant output. In such a condition, the output will oscillate between two quantization levels (as shown in Fig. 10.38), maintaining its mean value equal to the input one. The output is switching with a frequency, depending on the input signal level. In the simplest case (input equal to zero and 1bit modulator), the output will oscillate between +1 and -1 with a spectral component with a frequency $f_s/2$ and therefore out of band (we are working in the over-sampling case).

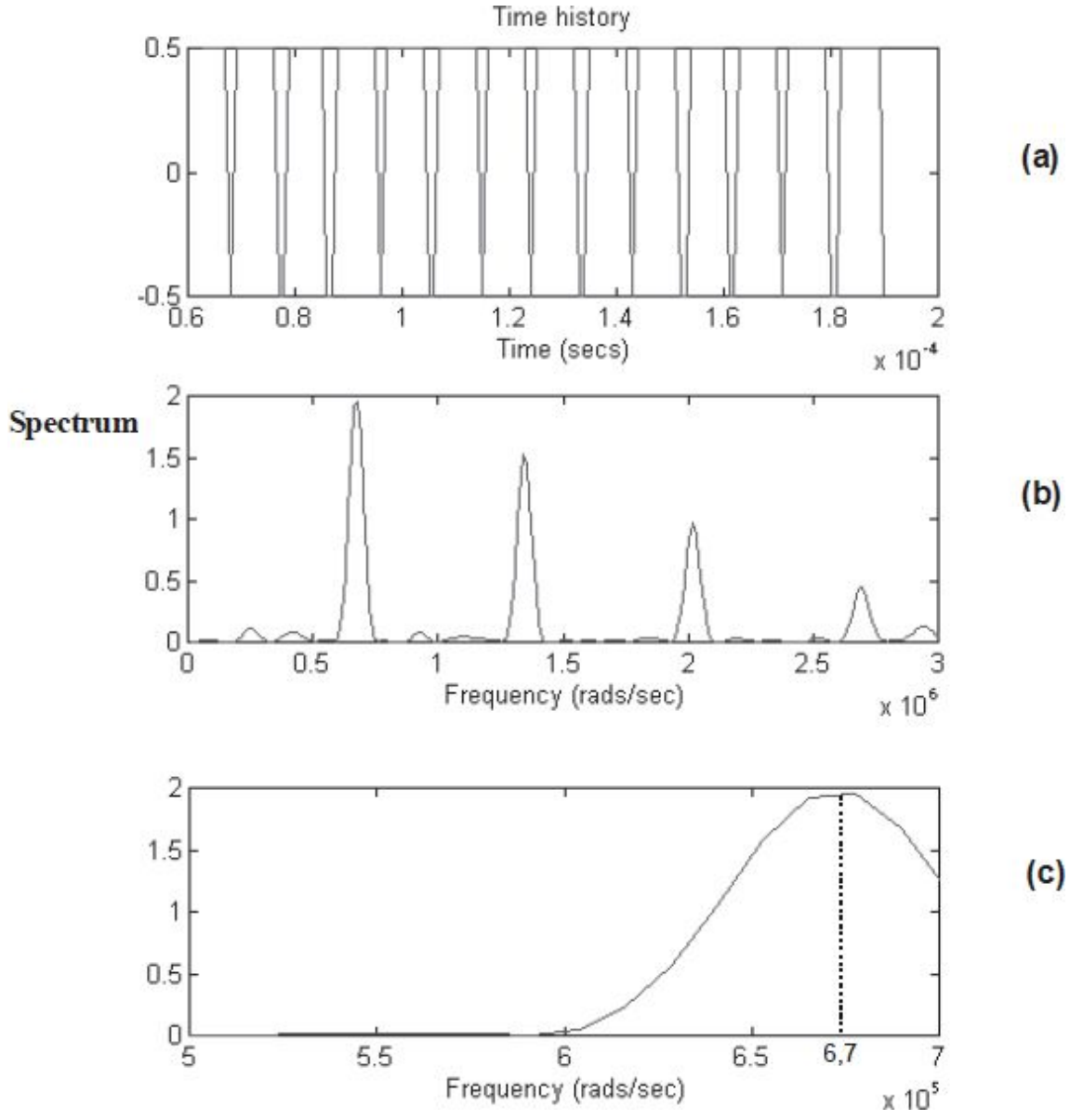


Fig. 10.43: Modulator output with a constant input signal with amplitude $^1/_3V$ (a), output signal spectrum with "tones" (b) and detail on the first tone (c).

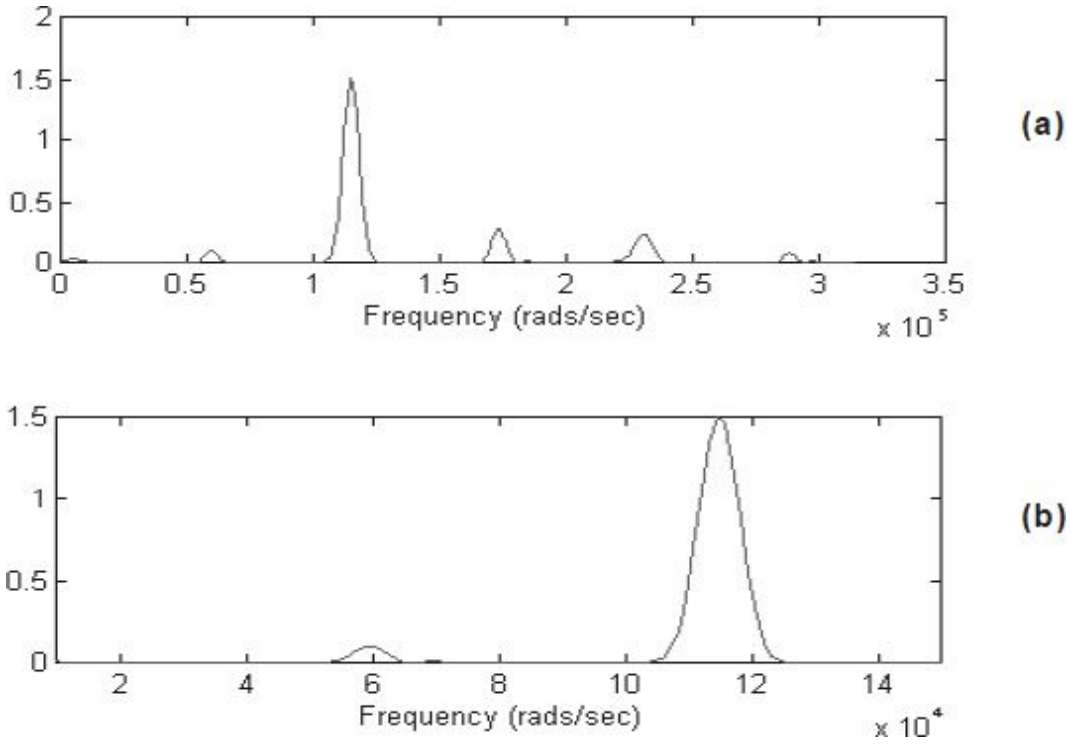


Fig. 10.44: Output spectrum for the modulator with Vin=$^1/_{17}$V (a) and detail of the first 'tone' (b).

Moreover, with a signal assuming the rational values 1/(2q+1), the output oscillates between +1 and –1, but, every 2q+1 samples, we have an additional +1 which gives a spectral component with a frequency $f_s/(2q+1)$. As we can see, for values of *q* high enough, such a frequency can fall within the bandwidth, causing an unexpected increase in noise. These components are called 'tones' and cause significant degradation in the final resolution of the modulator.

A typical example is shown in Fig. 10.43. We used an input signal with constant input with amplitude $^1/_3$V and a sampling frequency of f_s=320kHz. In the graphs, we can see the signal both in the time domain and the frequency domain. The spectrum has various peaks (the so-called 'tones') around multiples of the angular frequency $6{,}7\cdot10^5$rad/s (i.e. at a frequency around 107kHz), as foreseen by the theory $f_s/3$ (320kHz/3=106667Hz). If the input signal bandwidth is the audio band (up to f_0=20kHz), and therefore 125.6krad/s, then not even the lowest frequency tone causes any signal distortion.

However, with a continuous signal with lower amplitude, equal, for example, to $^1/_{16}$V, the first tone will have a frequency (fs/16) 360kHz/16=20kHz approximately, degrading the overall signal-to-noise ratio (SNR_{out}).

Using an input of 1/17V (Fig. 10.44), we can find, for example, a ‘tone’ at 18.8kHz (118krad/s) which remains unchanged towards the following filters and therefore is added to the in-band quantization noise penalizing the final resolution of the converter. We can evaluate the quantization noise power by considering the presence of these tones. The quantization noise depends on the input signal, and this effect is named ‘pattern noise’.

The results of the analysis for a first order ΣΔ modulator are shown in the ‘pattern noise’ graph in Fig. 10.45. We can notice how the SNR depends on the input signal level; in fact, there are peaks in correspondence of the rational values of the applied signal. So, we can see that, for some input values, the noise overcomes the theoretical values computed with the Eq. (***). In fact, if the noise mean value is equal to the theoretical limit evaluated previously, we can verify noise peaks whose amplitude and width are inversely proportional to the over-sampling frequency.

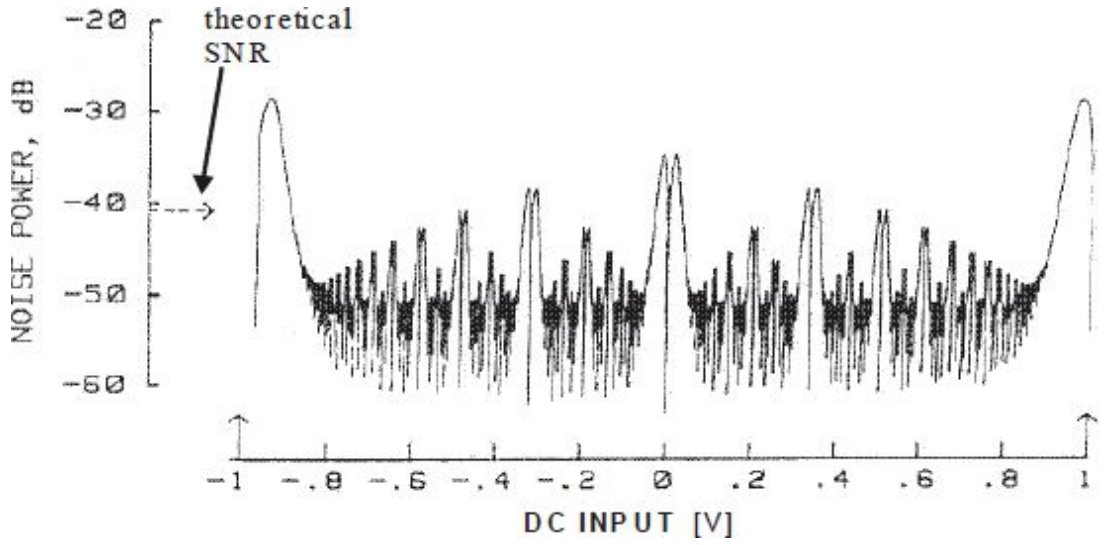


Fig. 10.45: “Pattern noise”: in-band quantization noise as a function of the input signal amplitude for a I order modulator.

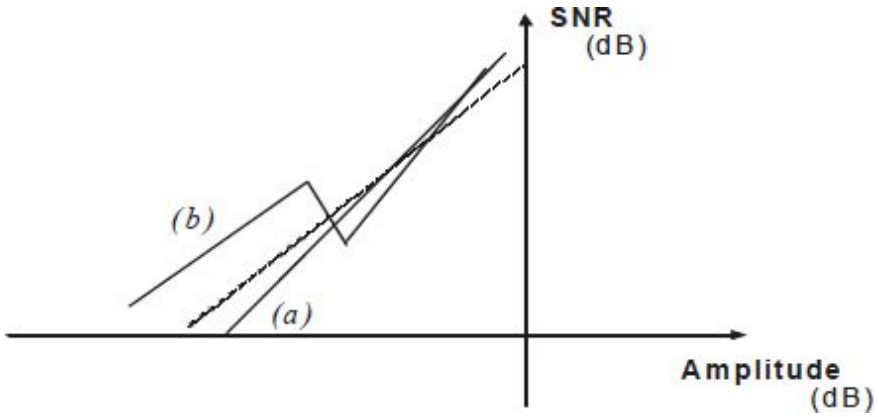


Fig. 10.46: Pattern noise effect for a first order modulator with sinusoidal signals as an input.

With a simple time integration to obtain the SNR by the noise power, we can report the results obtained on a bi-logarithm graph SNR vs. Input amplitude to verify the performance levels of the modulator itself. In the graph in Fig. 10.46, we show the curves for a pure sinusoidal signal as an input (a) and for a sinusoid with an offset of Δ/64 (b).

The pattern noise phenomenon degrades also the processing of sinusoidal signals because they are very slow compared to the over-sampling frequency (although they are not constant). Therefore, we can see that, for some input sinusoidal amplitude values, the SNR obtained is greater than the ideal value (the dashed line) because the input is far from the values which correspond to the noise peaks. On the other side, when the sinusoid amplitude is such that it causes to produce the maximum for the quantization noise, the performance in terms of SNR is degraded. The tones' generation as an input function is one of the main reasons for which this kind of modulation is not used.

To eliminate the pattern noise, we can try to avoid the autocorrelation of the quantization noise. A possible solution is the following: we can add to the noise a high frequency noise (named dither), and therefore out-of band, to allow a simple cancellation performed by the low-pass filter downstream the modulator. Its presence allows eliminating the correlation between the noise and the signal. The more effective solution is the one given by the modification of the architecture towards a second order (or more) modulator. The higher is the modulator order, the higher is the number of successive integrations for the quantization noise, decreasing the in-band presence.

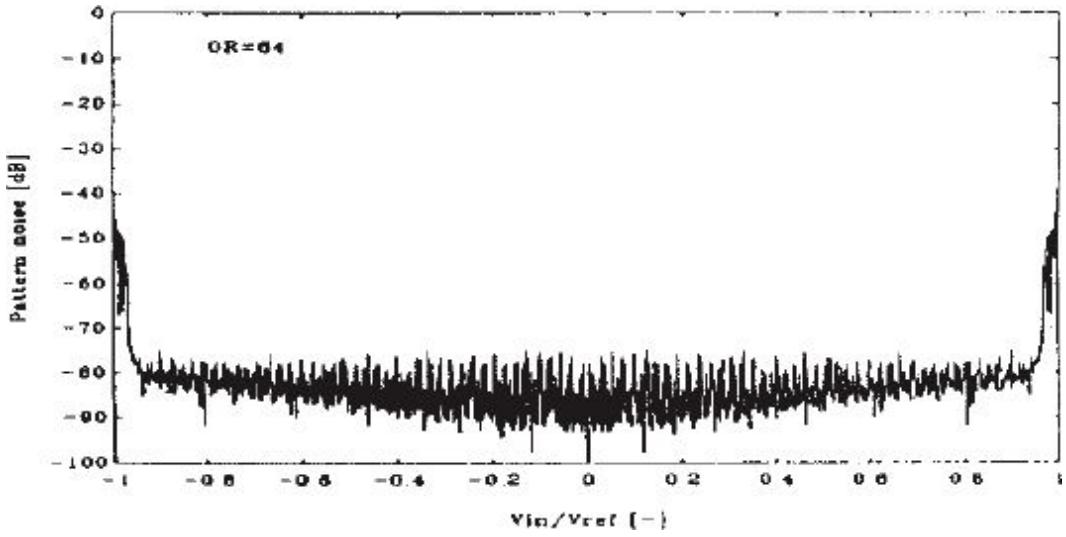


Fig. 10.47: "Pattern noise": in-band quantization noise as a function of the input signal amplitude for a II order modulator.

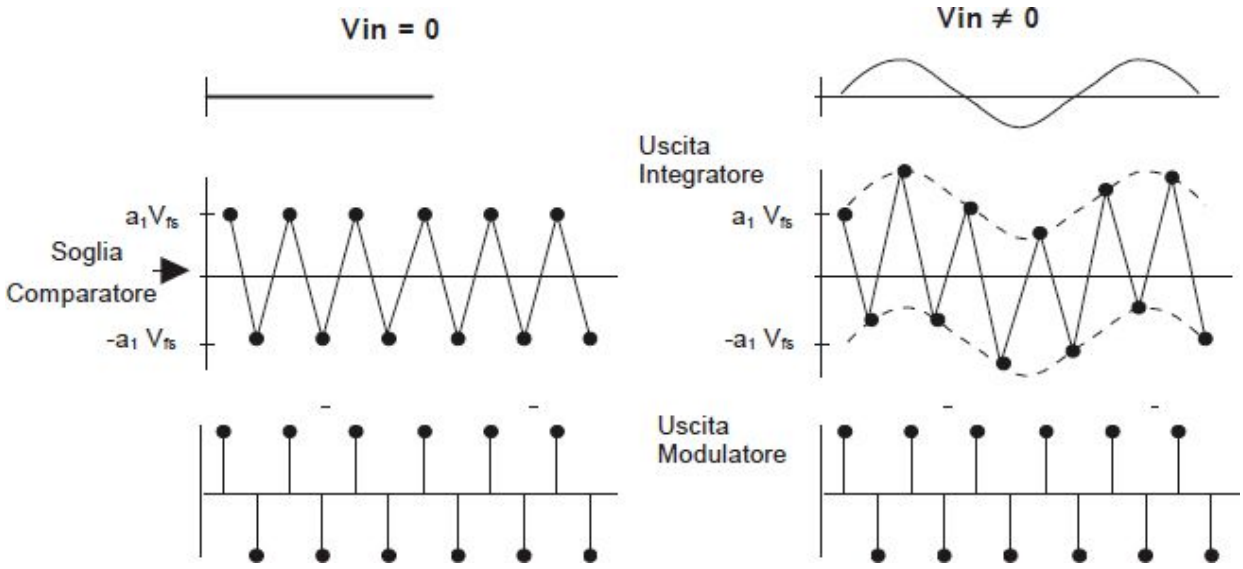


Fig. 10.48: "Dead zone" effect on a I order modulator: the output is unchanged for different input signals.

Fig. 10.47 shows the in-band noise for a second order modulator as a function of the DC input signal. We can notice that the tones' amplitude are much lower than those of the first order modulator (in Fig. 10.45) and that how the double integration loop can avoid the correlation more effectively.

Finally, we can demonstrate that the amplitude and the width of the peaks due to the pattern noise are inversely proportional to f_s.

10.5.2 “Dead zone” or “Idle channel”

There are some ‘dead zones’ in the converter, in which the comparator does not switch. Therefore, an interval of input amplitude values exists, for which the comparator of the ΣΔ modulator does not change its “decision”, limiting the resolution.

To analyze the ‘idle channel’ phenomenon, we can consider the modulator linear model. Consider a constant signal as an input, for example a null signal. As seen before, the comparator output will periodically vary between $-V_{FS}$ and $+V_{FS}$. The integrator output will vary between $-V_{FS}$ and $+V_{FS}$, too (in the more general case, with a DC gain equal to a_1, between $-a_1V_{FS}$ and $+a_1V_{FS)}$. Suppose now to superimpose a variable signal on the sinusoidal input with appropriate amplitude: the integrator output will vary, but the comparator “decisions” are not affected. The output sequence is identical to the preceding one with $v_{in}=0$, which is shown in Fig. 10.48. The range containing various integrator output values which do not modify the comparator “decisions” is named the “dead zone” or the “idle channel”.

In such a case, the integrator output will assume the trend shown in Fig. 10.48b, but it can happen that the “decisions” of the comparator are unchanged. It is therefore possible to shift the integrator output without changing the sequence provided by the modulator in correspondence of the range value named the dead zone. The position and the extension of this region depend mainly on the value of the input signal, and we can construct a correspondence between the noise peaks (due to the pattern noise) and the idle channel. This is obviously a qualitative discussion of the phenomenon. To be more precise, we can quantify the input sinusoid which does not modify the output sequence. If the transfer function of the integrator is the following:

$$H(z)=a_1\cdot\frac{z^{-1}}{1-z^{-1}}$$

and the output of the integrator will be a signal whose amplitude is $a_1H(f)\cdot Vin$. In Fig. 10.48b, we see how the amplitude is lower than a_1V_{FS} and that the sign of the output samples does not change. And because the comparator output is given by the samples of the integrator, we understand how the sequences provided in these two cases are equal. This phenomenon takes into account the resolution of the ΣΔ modulator. To produce an output variation in correspondence of an input variation, we must satisfy the relationship $|a_1H(f)Vin|>a_1V_{FS}$. The dead zone is avoided if:

$$a_1 \cdot Vin \cdot \left| \frac{z^{-1}}{1-z^{-1}} \right| > a_1 \cdot V_{fs}$$

which, making the substitution $z=e^{i2\pi f}$, becomes: $Vin\frac{1}{\sin(2\pi fT_s)} > V_{fs}$

The worst case is verified when the integrator gain is the minimum, i.e. for $f=f_0$. Therefore, we will determine a lower sensitivity of the modulator for the input signal variations in correspondence of the points with the minimum integrator gain. In the range $f_0T_s<<1$, approximating $\sin(2\pi fT)$ with $2\pi f_0T_s=\pi/OS$, we avoid the dead zone for:

$$\frac{Vin}{V_{fs}} > \frac{\pi}{OS}$$

Indicating with *n* the resolution in terms of number of bits of the whole modulator, we have the relationship: $Vin=V_{FS}/2^n$, which, substituted in the preceding inequality, gives:

$$\frac{1}{2^n} > \frac{\pi}{OS} \qquad \rightarrow \qquad OS>\pi 2^n$$

We see that we can enter the idle channel even for low over-sampling factors.

10.6. Digital Filtering

The modulator output signal will contain out-of-band components in addition to those in the in-band. The former are due to the input signal, the electronic circuitry noise, the quantization error, or possible interferences. We need a low-pass filtering to avoid that, reducing the output signal frequency down to the Nyquist one, a noise increment (due to the aliasing phenomenon) is verified. This can be translated into performance levels and final resolution worsening for the ΣΔ modulator. Such a filter will obviously be digital because it works on the samples series produced by the modulator.

A digital filtering can be very advantageous in respect of its corresponding analog filtering. In fact, the latter implements the required equation by means of reactive components (capacitors and inductors) whereas a digital filter implements the same equation using two simple mathematical operations: sum and product. For this reason, the digital filters are very stable and have a great repeatability in the transfer function; moreover they are more resistive to aging and temperature variations.

For the ΣΔ modulator, the filter cut-off characteristics (for the quantization noise) are easier to be achieved because the noise is pushed towards the high frequencies. We need to cancel the out-of-band spectral components (with a frequency greater than f_0) to avoid the aliasing phenomenon during the decimation. We can make a filter (shown in Fig. 10.49) with bandwidth $2f_0$ capable of a net cut.

Operating at high frequencies (for a high OS factor), it is very challenging and expensive to satisfy the frequency selectivity requests (work close to f_s with bandwidth equal to $2f_0$) to make a selective digital filter.

To overcome to this problem, we prefer to divide the filtering operation in a cascade of two or more phases: the first stage is a filter capable of cancelling the quantization noise (and all the undesired high frequency components) and, at the same time, of reducing the sample frequency between the over-sampling and the Nyquist frequencies. The second filter will work with a lower frequency, to cancel the out-of-band signal components (note that it is simpler and cheaper to have a very selective filter at low frequency than at high frequencies). After this filtering, we can decimate the output sample frequency down to the final value, i.e. the Nyquist frequency '$2 \cdot f_0$'. Along the filtering chain in Fig. 10.50, the resolution of the output signal increases gradually.

The filter in the two stages can be made as *FIR* ('Finite Impulse Response') or *IIR* ('Infinite Impulse Response') filters, depending on the form of the *z* transform (made by only zeroes or poles and zeroes, respectively). The target of the filtering (for both kinds of filters) is to remove the quantization noise. The ΣΔ modulator "shapes" the quantization noise towards the high frequencies, decreasing the in-band presence. The high frequency noise is removed by the digital filtering, and the in-band noise reduction is equivalent to an increase in resolution for the digital output signal. As observed, a further noise reduction is obtained by over-sampling the digital signal. The resolution increase is obtained reducing the output samples' frequency (by means of an average of more samples) decimating it.

The two filters are combined opportunely to obtain the desired filter (with two stages): use of an FIR for the I stage because it is a high frequency filter with little selectivity and decimation while the IIR is used for the decimation while the IIR is used for the II stage for a more selective low frequency filtering (after the decimation). Fig. 10.51 shows the blocks of the two filters.

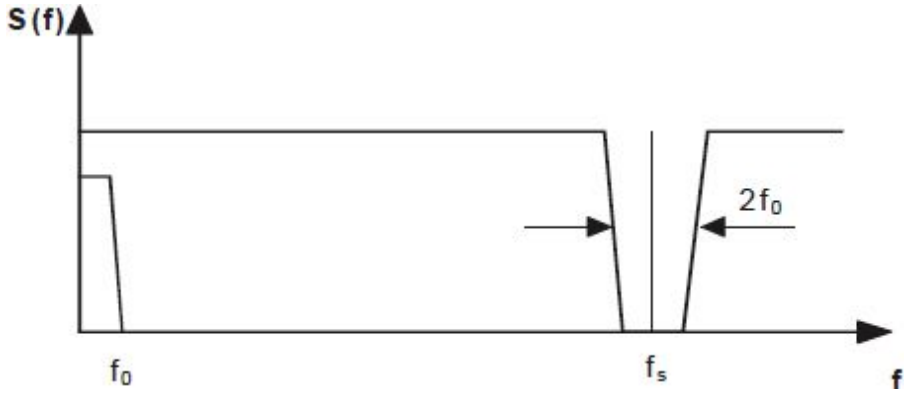


Fig. 10.49: Possible filter to cancel the frequencies over f_0.

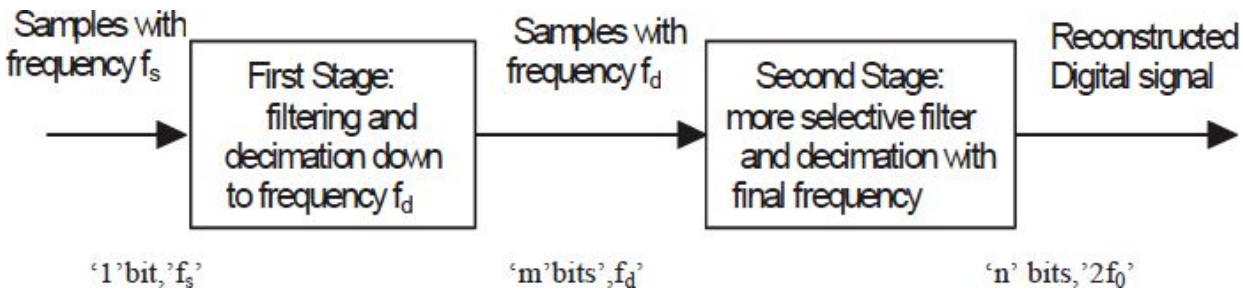


Fig. 10.50: Two stages filter block scheme.

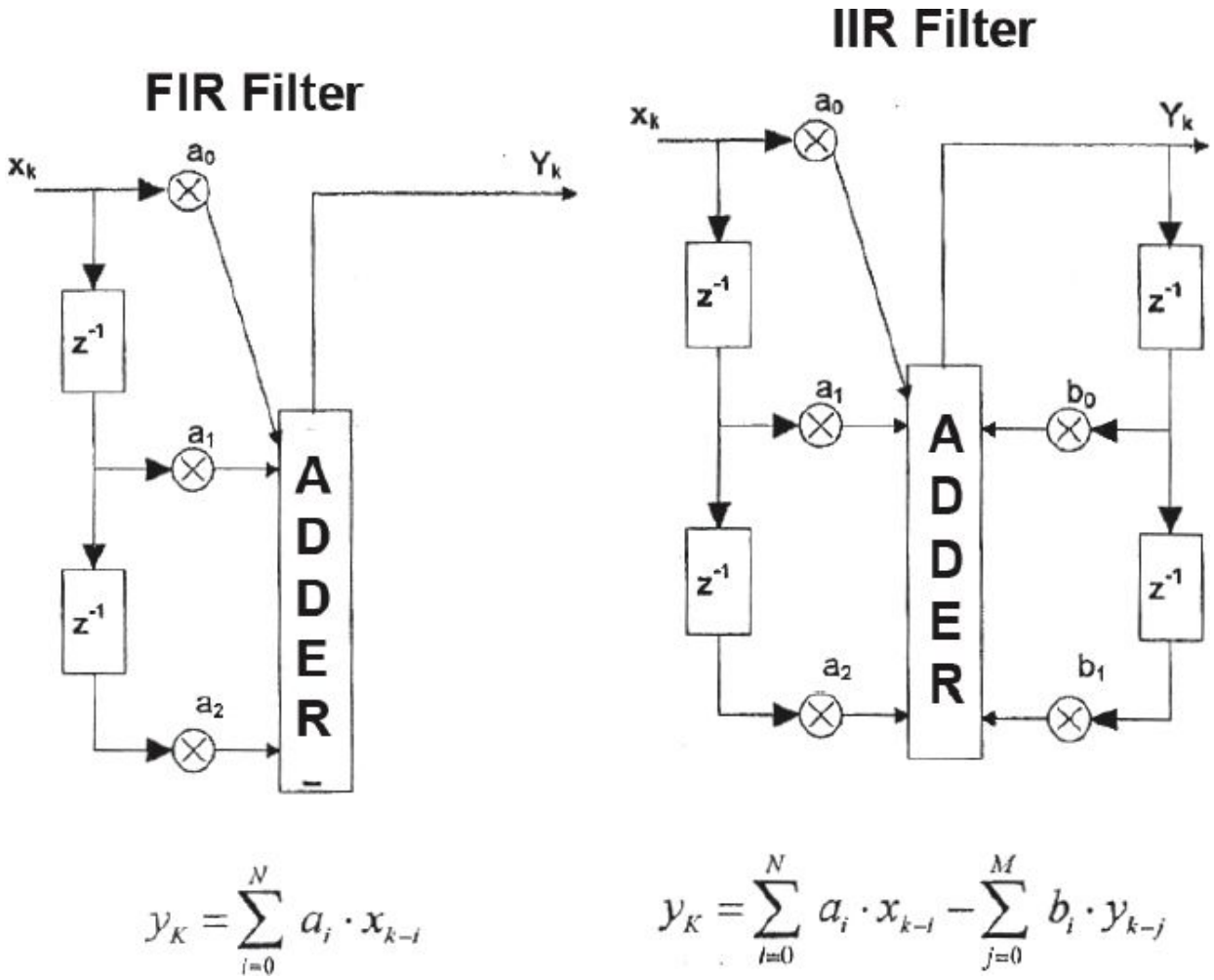


Fig. 10.51: Block scheme of the filters FIR and IIR.

10.6.1 FIR

An FIR filter can simultaneously perform the decimation and low-pass filtering. These filters use a finite number of input samples to compute an output sample. A simple FIR example is the average of N samples with a weight 1/N making the decimation (with an N factor) of the input data.

More generally, an FIR filter can have a weight function which “treats” the samples differently. If the input samples x_i have a frequency f_s (= over-sampling frequency), and each is weighted with the factor a_i, while the

output samples y_k have a frequency f_d (the intermediate frequency after the first decimation), we have:

$$y_k = \sum_{i=0}^{N} a_i \cdot x_{k-i}$$

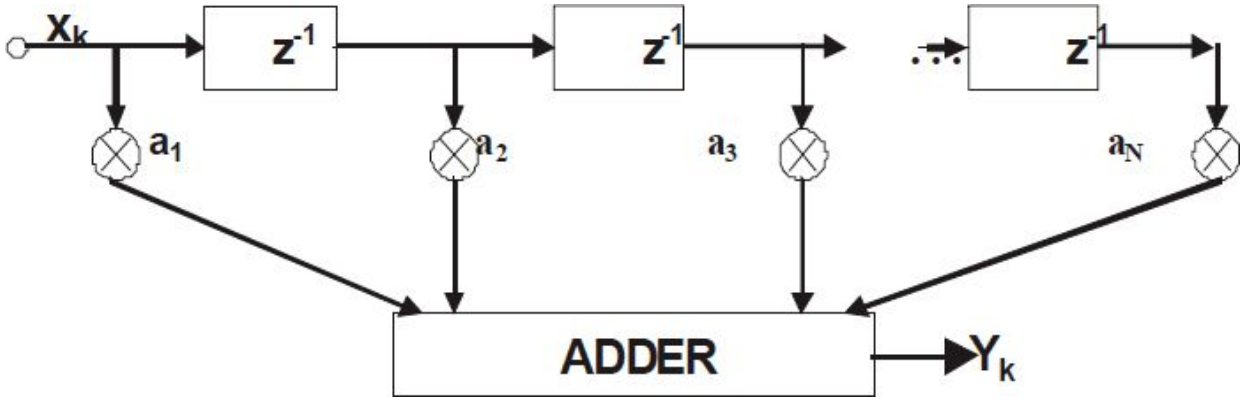


Fig. 10.52: Block scheme of a generic FIR filter with an order N.

where $N = f_s / f_d$. The samples y_k can be obtained with the circuit in Fig. 10.52, in which there are adder, digital multiplier, and blocks z^{-1} which delay the sample with a sample time $1/f_s$. For the filter, we perform a simple average, obtaining:

$$y_k = \frac{1}{N} \sum_{i=0}^{N} x_{k-i}$$

In this simple case, the realization of a filter does not need multiplier circuits because the coefficients of every x_i (within the summation) are equal to 1. Adopting this solution, we can simplify the circuit realization, a requirement that is a must when working at high frequencies. On the other side, the transfer function that can be obtained causes signal “distortions” because of the non-constant gain within the base-band. The following filtering block, operating at a lower frequency, can have more circuit complexity, which can be exploited to compensate the “distortions” introduced by the preceding stage, as well as for greater selectivity.

We can obtain the transfer function as:

$$H(z)=\frac{Y(z)}{X(z)}=\frac{1}{N}\sum_{i=0}^{N}x_{k-1i} \qquad \Rightarrow \qquad H(z)=\frac{1}{N}\frac{\left(1-z^{-N}\right)}{\left(1-z^{-1}\right)}$$

and substituting $z = e^{i2\pi fT}$:

$$H(f)=\frac{\sin c(\pi fNT)}{\sin c(\pi fT)} \qquad (x)$$

i.e. a sampled *sinc* function. The zeroes of the H(f) can be employed to reduce the undesired noise. To choose the decimation factor f_d/f_0, we need to note that the filter H(f) distorts the signal within the useful band f_0; particularly, for a given f_d, we have that the signal out of the band f_0 is attenuated. Thus, we cannot choose f_d to be too low. At the same time, we cannot choose f_d to be too high; otherwise, the decimation is not effective, and the samples will have a very high frequency. For this reason, observing the attenuation of the components f_0 in the graph shown in Fig. 10.54, usually, we choose $f_d/2f_0=4$.

Ideally, the transfer function for this filter should be a rectangle (in the frequency domain), non-implementable with actual circuits. For this reason, a *sinc* function is used, which is shown in Fig. 10.53. The main lobe is the part which approximates the rectangle function; there are also zeroes with a frequency multiple of f_d, which can be used to improve the converter performance levels 'preparing' a *low noise zone* for the decimated signal. In fact, after the decimation, the signal spectra are close to each other and can lie where the noise is very high (for a disturbance with this new frequency), lowering the performance improvement obtained with the over-sampling.

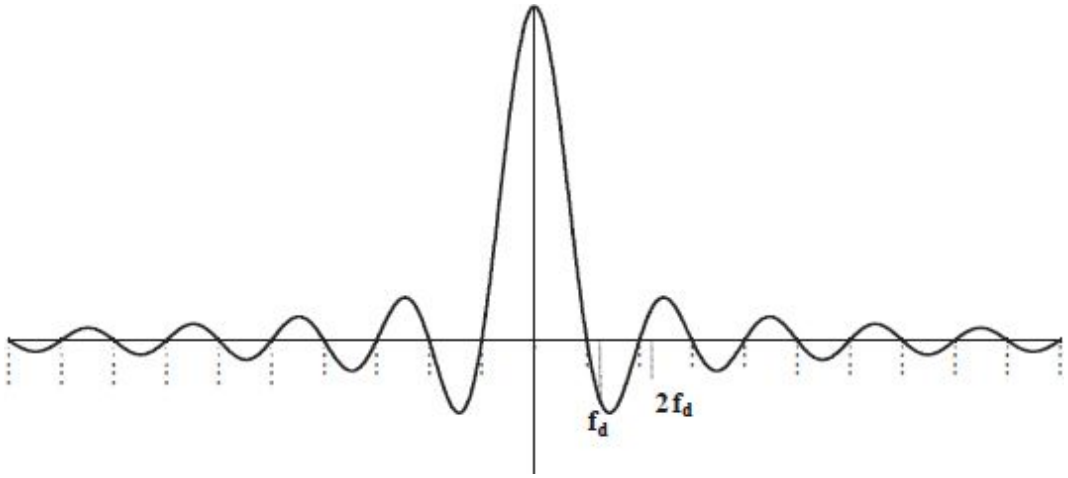


Fig. 10.53: Representation of the transfer function which makes the average on N samples.

If the digital filter introduces zeroes in correspondence of central frequencies (around which the spectrum is repeated after the decimation), we can attenuate or eliminate the undesired noise.

Suppose, for example, to have the spectra of the signal sampled at f_s with a disturbance close to f_d, the frequency at which the decimator output samples will be given. With a simple frequency lowering at which the samples are given (without filtering, but taking a sample every N), the spectrum will be superimposed on the noise already present, distorting the contained information. This can be avoided with a filter placed before the decimation stage. This filter must attenuate the frequencies around which the spectra after the decimation will be centered. This can be obtained with a sinc transfer function. If the frequencies around which the spectra after the decimation will be centered are f_d and its multiples, the first zero of the FIR filter transfer function is exactly at f_d. Obviously, the other zeroes have a frequency multiple of f_d.

One can think of using a transfer function as, for example, $sinc^2$ or $sinc^3$, to improve the efficiency for the frequency "cleanliness" around f_d because, for these functions, the higher is the exponent, the lower is the amplitude of the secondary lobes. At the same time, the in-band noise due to the aliasing phenomenon can be increased. We can demonstrate that the optimal filtering condition is obtained for $sinc^{L+1}$ where L is the modulator order.

At this point, we have to choose the value of the decimation factor for this first block, or better, what has to be the value of f_d (first zero frequency for the *sinc* function), compared to the value of the Nyquist frequency. The first parameter that must be taken into account is the filter response cut-off for frequencies close to f_0. In fact, because this transfer function is not a rectangle, we have attenuation for the signal components with a frequency less than f_0. What matters is compensating the attenuation for the signal components at the edges of the bandwidth; this compensation is not convenient if higher than 3dB and is shown in Fig. 10.54 (for N=4) as a function of the intermediate over-sampling, i.e. for the ratio between the intermediate frequency f_d and the Nyquist frequency $2f_0$. In the graph, we can see that the attenuation is contained within 2.75dB, but it rapidly increases, decreasing the ratio of the intermediate OS factor. For this reason, 4 times the Nyquist frequency is the minimum value of the f_d intermediate frequency.

Another problem is the noise increase during the decimation with a function $sinc^{L+1}$ due to the noise quantization aliasing. Until the frequency decreases down to 4 times the Nyquist one, the noise increase is limited within 0.14dB, but, increasing the decimation factor, it grows very rapidly.

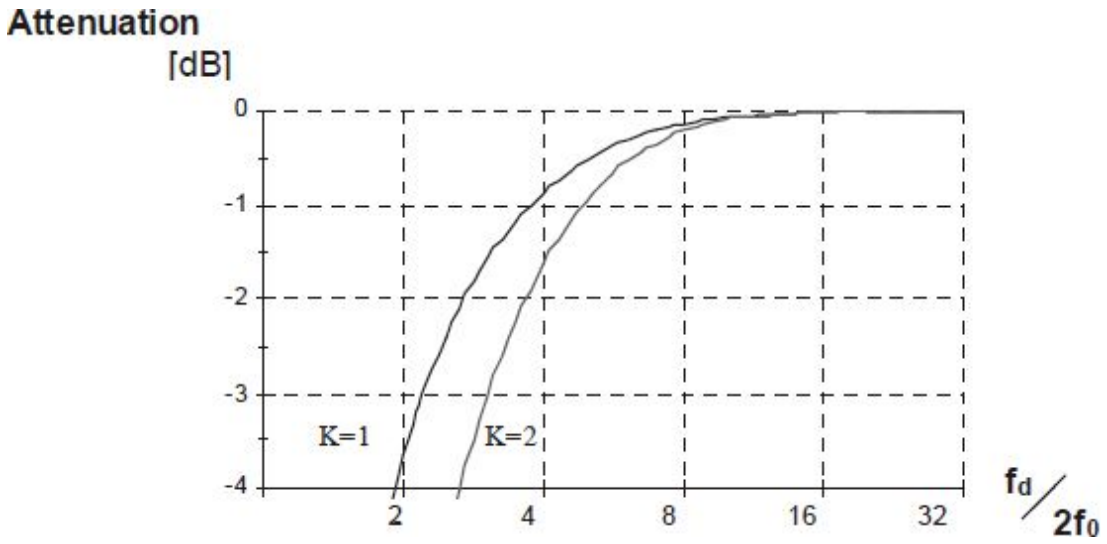


Fig. 10.54: Attenuation representation of the function $sinc^K$ for frequencies close to f_0, as a function of the intermediate oversampling factor.

At the same time, it is not opportune to increase f_d above this value. In fact, if the OS is too low, the output samples have a frequency too high, and the following IIR filter may not have the selectivity characteristics necessary to reach the theoretical resolution expected for the whole modulator. For this reason, generally, we choose an intermediate decimation factor which allows us to have $f_d / 2f_0 = 4$.

A final comment is about the fact that the decimation allows improving the resolution with which the samples are provided. If we indicate with *n* the quantizer number of bits (n=1 for a comparator), the levels at which the samples arrive to the FIR are 2^n. The output resolution, with a transfer function equal to $sinc^K$ and a decimation factor *N* is:

$$b = \log_2\left(2^n \frac{N^K}{\sqrt{N}}\right) \qquad (XX)$$

For example, starting from n=2bit, a decimation frequency $f_d / 2f_0 = 4$, and a third order *sinc* (k=3), we obtain 7bits at the output.

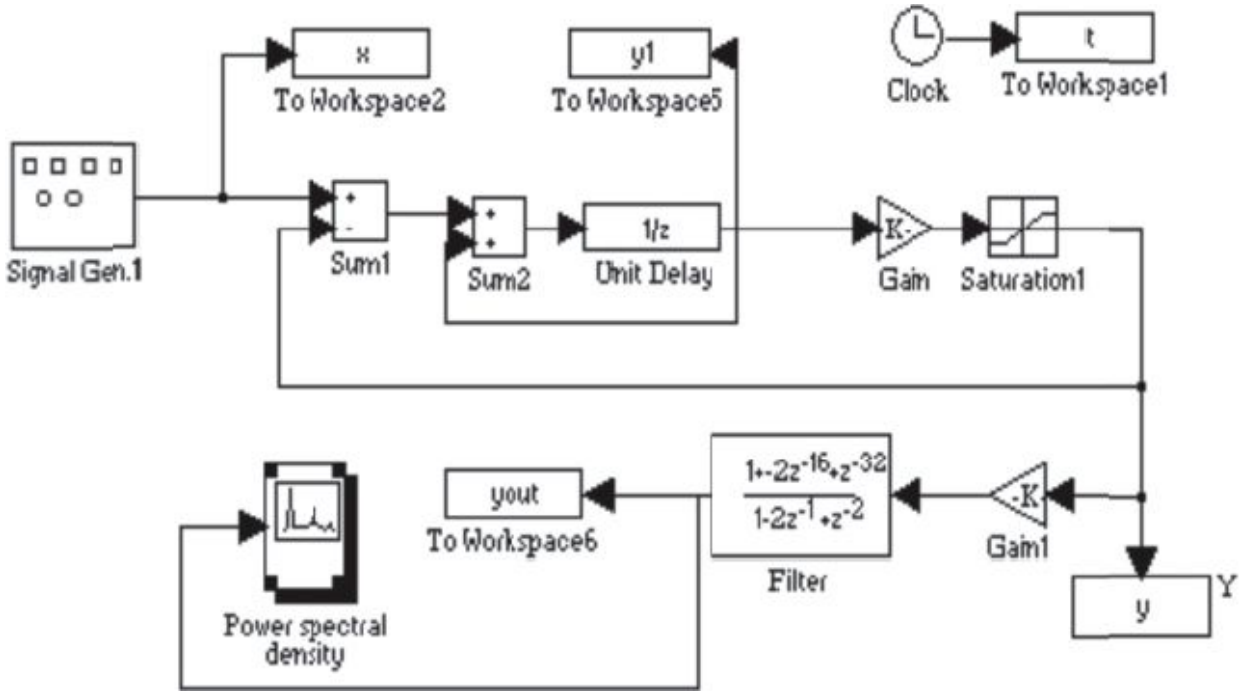


Fig. 10.55: Simulink modeling of the ΣΔ converter and the following decimation of the output sequence.

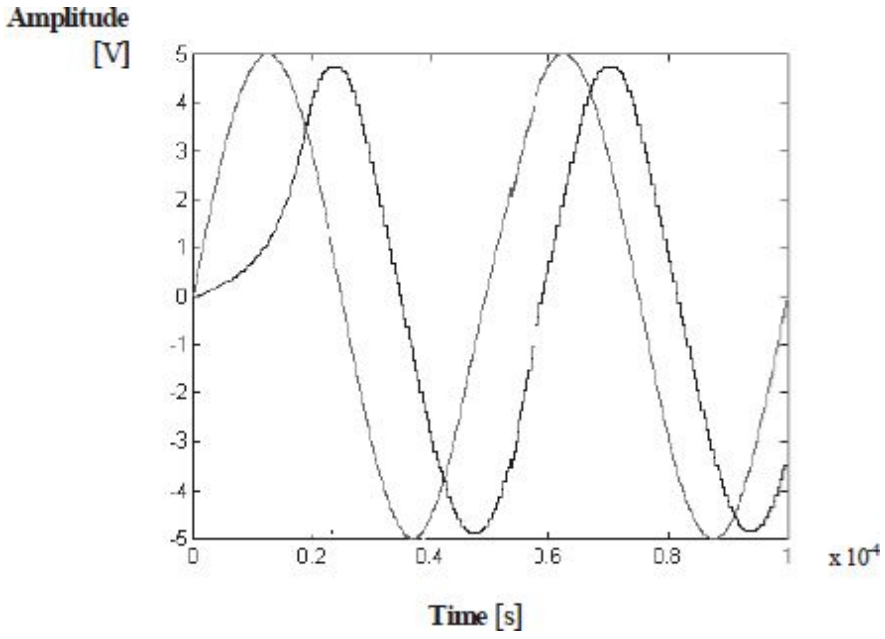


Fig. 10.56: Input and output ΣΔ converter signals for oversampling and decimation =16, for a first order modulator with a 1 bit ADC.

One can see a MATLAB simulation in Fig. 10.55. Compared to the preceding model (Fig. 10.25), there are the blocks *Gain1* and *Filter*, which, together, implement the Eq.(x), i.e. the average with N samples with N=16, and therefore a sampled *sinc*, which reduces the output samples frequency by a factor of 16, in respect of the block input frequency (a sample every 16, decimation).

We are considering always a first order modulator (i.e. it has only one feedback loop), and therefore K=2, which uses the simplest ADC: the comparator, i.e. n=1. From the preceding equation, we can obtain that the output converter resolution should be 7bits, i.e. y_{out} will be a discrete signal with 2^7=128 allowed levels. As the input signal, we use a sinusoid with amplitude 5V, i.e. the maximum allowed because of the dynamic of 10V, and a frequency 20kHz. The obtained result is shown in Fig. 10.56. In addition to the higher number of levels which define the output signal quantization, we can observe the attenuation (about 1dB) produced by the final filter; in fact, y_{out} does not reach 5V, but its peak value is around 4.5V.

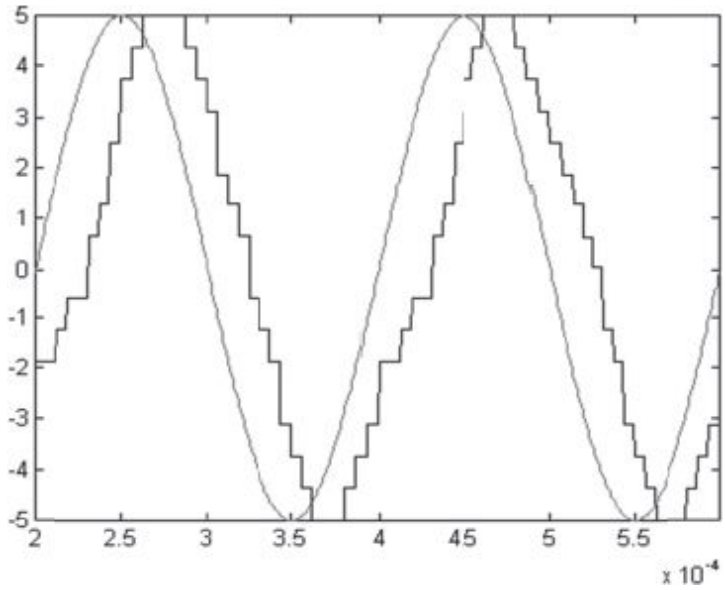


Fig. 10.57: Input and output ΣΔ converter signals for over-sampling and decimation with a factor of 4, for a first order modulator with a 1bit ADC.

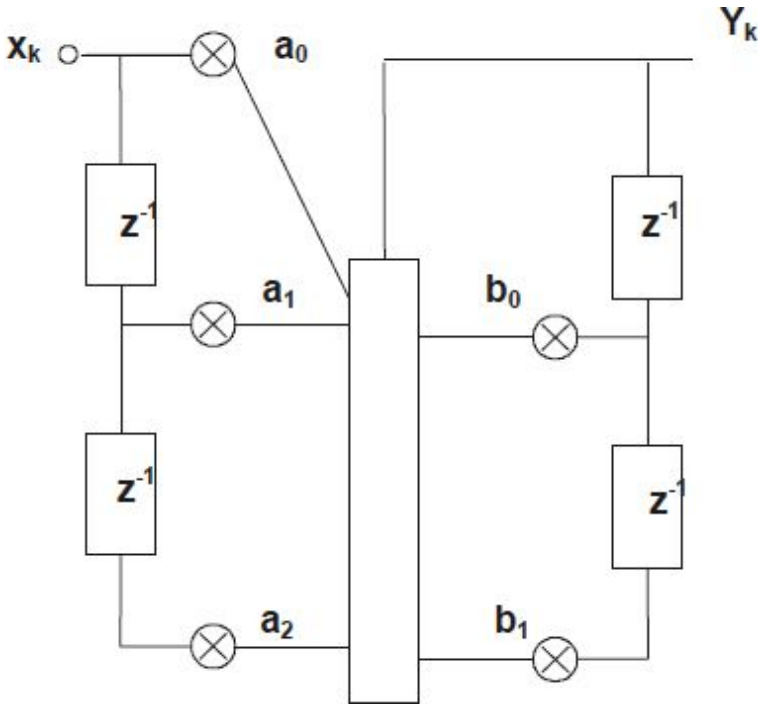


Fig. 10.58: Block scheme of a generic IIR filter.

Reducing the sampling frequency by a factor of 4 (i.e. N=4), leaving the other parameters (n=1, K=3) unchanged, what is obtained is shown in Fig.

10.57. We see how the discrete levels are decreased, and more precisely, the converter can provide 4bits as the output (16 levels), as foreseen by the theory.

An FIR is preferable for the circuit implementation (very important for the high frequencies because the input sequence has a frequency f_s) and for the possibility to clearly cancel the undesired spectral components with the zeroes of the transfer function (paying attention to the distortion introduced on the original spectrum because of $sinc^k$).

10.6.2 IIR

Unlike the FIR filters, which use a finite number of preceding pulses to compute the output, the IIR filters use all the preceding pulses. In this way, the filter can implement transfer functions which contain both zeroes and poles, which provide a 'mask' in the frequency domain with great slopes and, therefore, great frequency selectivity. Obviously, these better performance levels are paid with a higher circuitry complexity and therefore are actually obtainable only with signals with a lower frequency ($f_d << f_s$) because the FIR output samples frequency is decimated by a factor N. The transfer function for this filter is:

$$y_K = \sum_{i=0}^{N} a_i \cdot x_{k-i} - \sum_{j=0}^{M} b_i \cdot y_{k-j}$$

Now, there is a feedback term (the output for a given moment depends on its preceding value), which determines the better performance levels (at least for intermediate frequencies), but also the higher circuit complexity, as we can see in the block scheme in Fig. 10.58.

In Fig. 10.59 and Fig. 10.60, we see the effect of the decimation on the spectrum and the noise. We can see an example for a 1bit ΣΔ converter with a filter $sinc^2$ which samples and decimates of a factor 16. The Nyquist frequency is fixed at 20kHz. The over-sampling factor is OS=16 i.e. the modulator output samples have f_s=16.20kHz=320KHZ. The Simulink model is shown in Fig. 10.61.

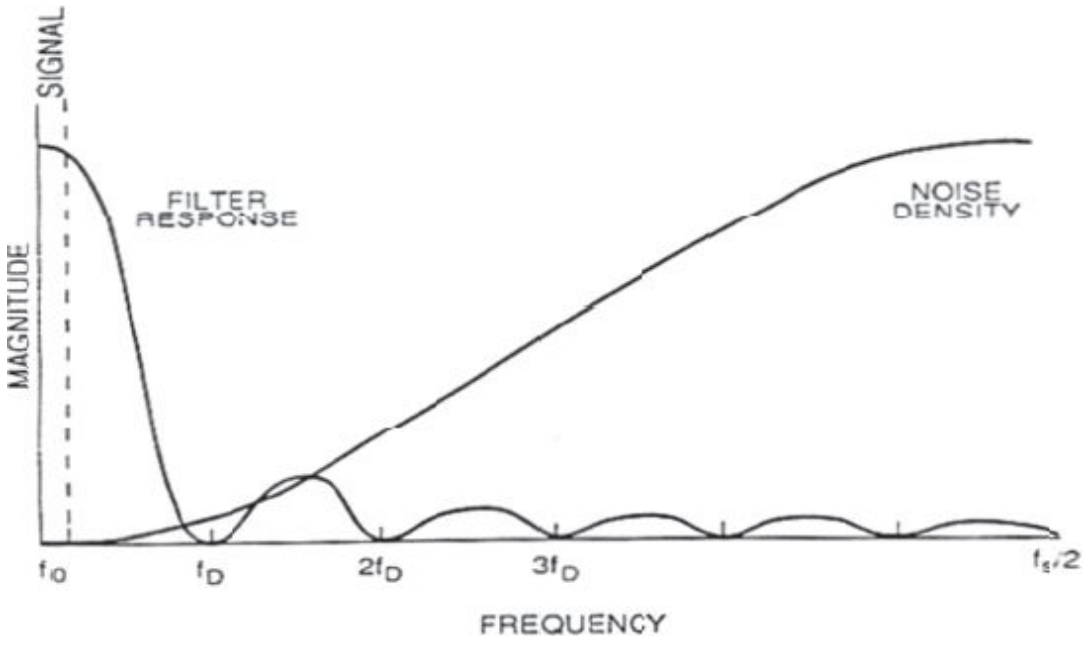


Fig. 10.59: Decimation by means of a filter with a response $sinc^2$: f_S is the sampling frequency, f_D the intermediate decimation frequency, and the useful bandwidth is in the range $0<f<f_0$.

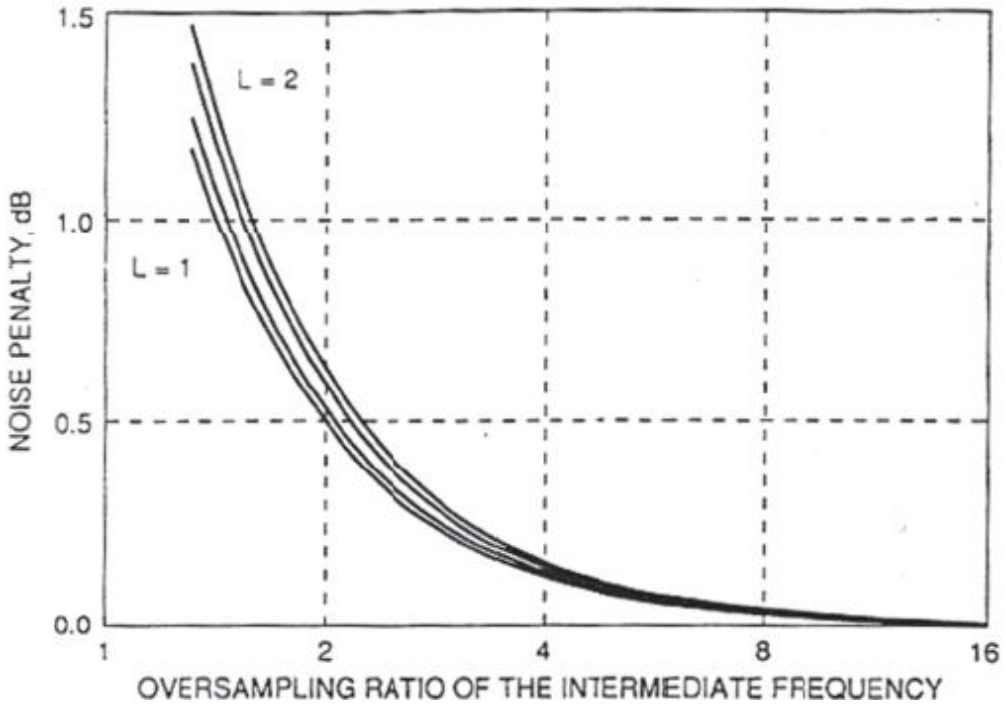


Fig. 10.60: Increase in the noise due to the decimation by means of the filtering $sinc^{L+1}$: $N=f_S/f_D$ is the decimation ratio. The results concern ΣΔ modulators for I (L=1) and II (L=2) orders.

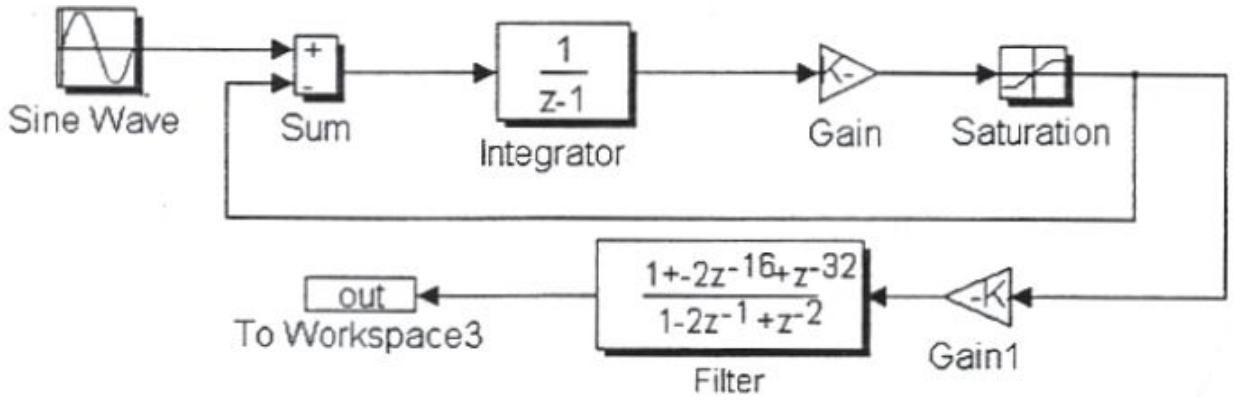


Fig. 10.61: Simulink model with a 1bit ΣΔ converter.

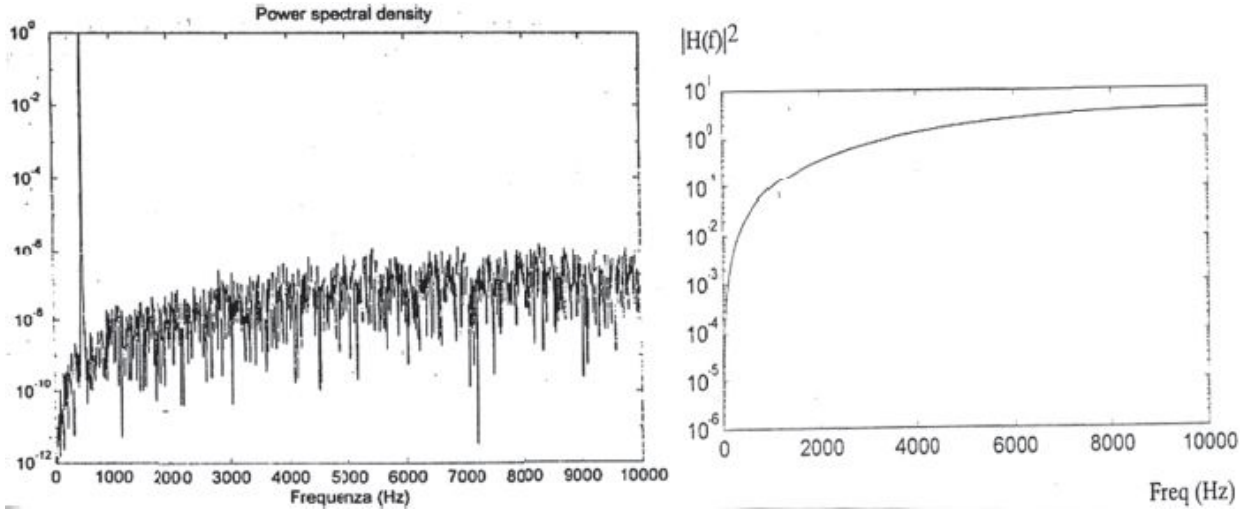


Fig. 10.62: Simulated and computed ADC output noise shaping.

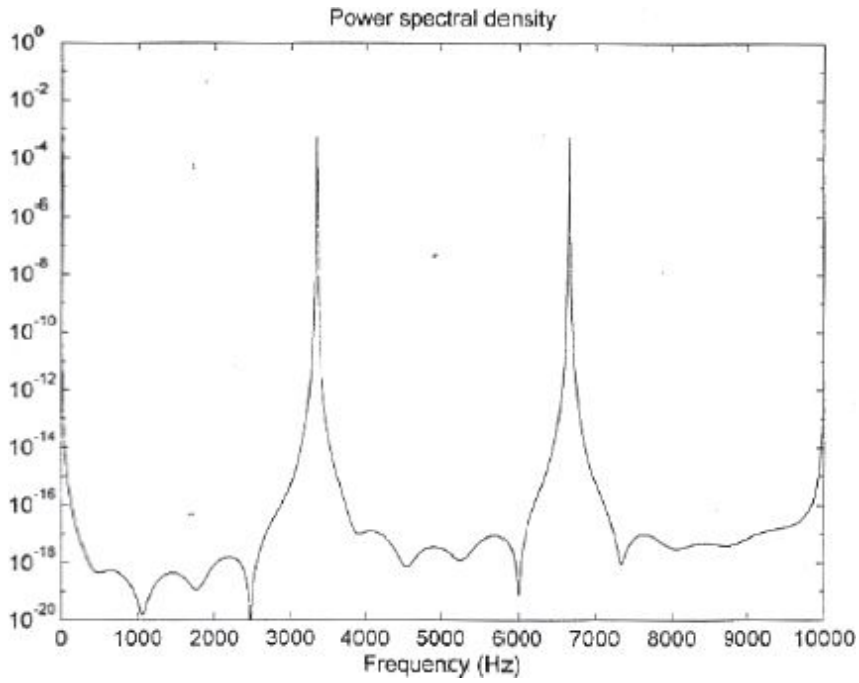


Fig. 10.63: Pattern-noise effect.

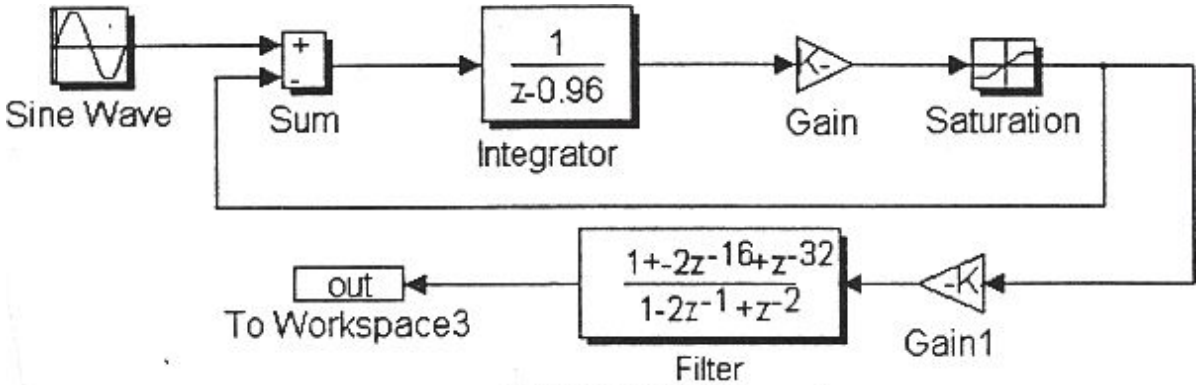
The expected SNR is SNR=6.02·(c+1.5L)-3.41=38.8dB because c=1 and L=4. The converter resolution is b=(SNR-3.41)/6.02=(c+1.5L)=7bit. Note how this resolution is obtained using a 1bit ADC (comparator) with other components, particularly digital (digital filter). The output spectrum with f_{in}=530Hz is shown in Fig. 10.62 with its simulated (on the left) and computed (on the right) noise-shaping. To see the pattern-noise effect, we can choose $v_{in}={}^{1}/_{96}$·FSR/2 and trace the noise quantization spectrum (Fig. 10.63). Note the intense peaks due to the tones at 3.33kHz, 6.66kHz, and 9.99kHz.

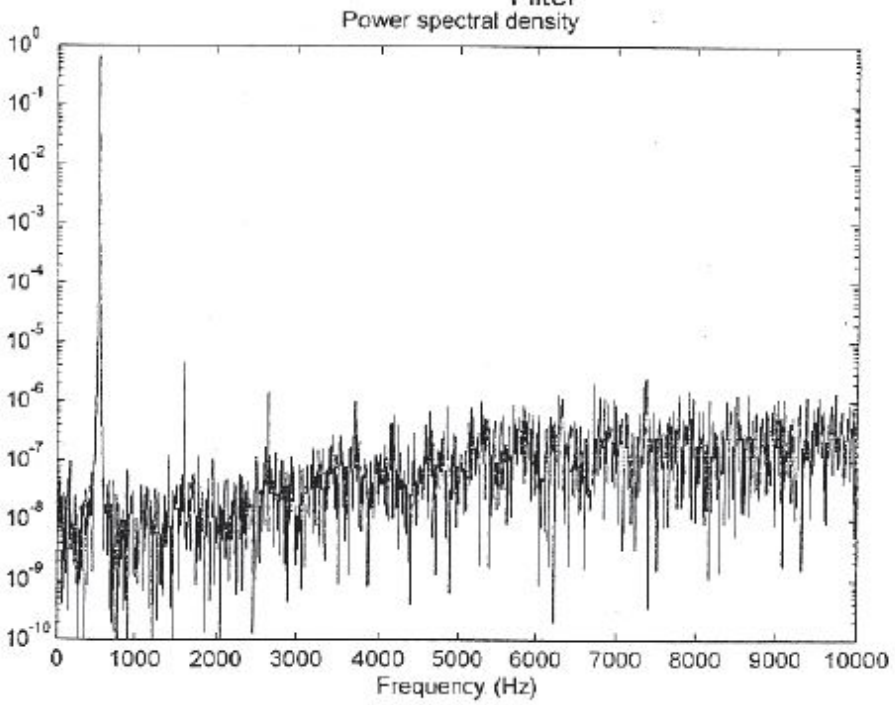
The tones separation is equal to $f_s/(2p+1)$=320kHz/96=3333Hz. Because the input signal is v_{in}=1/(2p+1)· FSR/2, we can compute the tones noise power equal to $E^2(f)=2[1/(2p+1)]^2(FSR/2)^2$. The number of tones in the bandwidth is: $(f_s/2.OS)/(f_s/2p+1)=(2p+1)/2\cdot OS$, therefore the total noise caused by the pattern-noise is: pattern-noise=$[OS\cdot(2p+1)]^{-1}\cdot(FSR/2)^2$ =-32dB (in the case shown in Fig. 10.63).

It is possible to consider a new ΣΔ non-ideality due to the fact that the real integrator does not have infinite DC gain, but only A_{DC}. This allows defining a leakage of the integrator defined as $L=1/A_{DC}$. The transfer

function is not $H(z)=1/(z-1)$, but is $H^*(z)=1/[z-(1-L)]$. Therefore, there is no more pole in the origin, but it has a frequency $f_p=L/2\pi f_s$. Because of this fact, in the frequency interval between 0 and f_p, the modulator loop cannot perform the expected noise-shaping.

Another problem is due to the small signal distortion because the idle tone happens for $v_{in}<FSR/|H(f)|$, and now H(f) is DC limited, but not infinite. Note such effects in the simulations shown in Fig. 10.64. The leakage introduces an in-band quantization noise increasing and, thus, a worsening of the SiNAD and of the ENOB. The tones worsen the SFDR and the SiNAD.


Sine Wave
Sum
1
z-0.96
Integrator
Gain
Saturation
out
To Workspace3
1+-2z-16+z-32
1-2z-1+z-2
Filter
-K
Gain1


Power spectral density
Frequency (Hz)

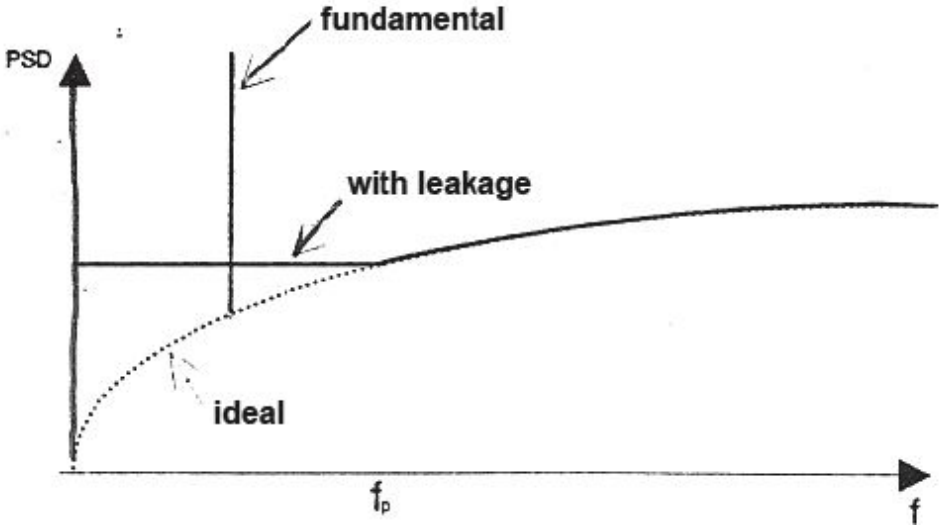

fundamental
PSD
with leakage
ideal
fp
f

Fig. 10.64: (a) Circuit with an integrator leakage; (b) idle tone spectrum; (c) leakage effect on the noise floor.

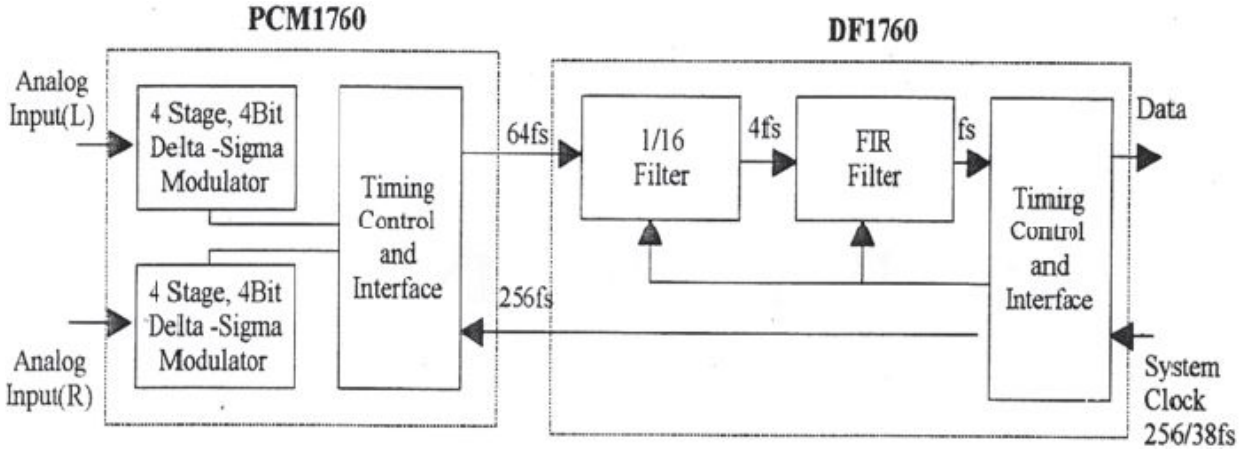


Fig. 10.65: Example of a modulator with a filter.

10.7. COMMERCIAL ΣΔ

Usually, they are constituted by two integrated circuits: one is the real modulator, and the other is the digital filter with decimation.

10.7.1 PCM1760 e DF1760

Consider, for example, the couple from the Burr-Brown (Fig. 10.65) constituted by a modulator PCM1760 p/v with a digital filter DF1760 p/v. The first is a IV order ΣΔ with a 4bit internal ADC and oversamples with a factor of 64. The digital filter is made by two FIR stages: the first decimates of 16 and the second of 4. The couple allows a global resolution of 20bits (output data channels).

The performance levels are provided differently than the traditional ADCs; see Fig. 10.66 in which the Gain error is specified in dB. Other dynamic parameters are reported in Fig. 10.67, as the Dynamic Range. The initials P and U refer to two models with dissipation of 580mW and 550mW while P-L and V-L refer to two models with maximum working temperatures of 260°C and 235°C, respectively.

PARAMETER	CONDITIONS	PCM1760/DF1760 MIN	TYP	MAX	UNITS
ACCURACY					
Gain Error			±0.5	±1.0	dB
Gain Mismatch				±0.5	dB
Bipolar Zero Error	V_{IN} = 0 at 20s Afeter Power-On			±0.4	%FSR[2]
Gain Drift	0°C to +70°C		±100		ppm/°C
Bipolar Zero Drift	0°C to +70°C		±20		ppm/°C

NOTES: (2) FSR means Full Scale Range, digital output code is from 90000H to 70000H, FSR=5V.

PARAMETER	CONDITION	TEMP	ADS801U MIN	TYP	MAX	UNITS
ACCURACY[2]						
Gain Error		+25°C		±0.6	±1.5	%
		Full		±1.0	±2.5	%
Gain Tempco				±85		ppm/°C
Power Supply Rejection Gain	Delta +V_s=±5%	+25°C		0.03	0.15	%FSR%
Input Offset Error		+25°C		±2.0	±2.5	%
		Full		±2.1	±3.0	%FSR%
Power Supply Rejection Offset	Delta +V_s=±5%	+25°C		00.5	0.15	%FSR%

NOTE: (2) Percentage accuracies are referred to the internal A/D Full Scale Range of 4Vp-p.

Fig. 10.66: Performance levels of the device under study compared with a normal ADC.

PARAMETER		CONDITIONS	PCM1760/DF1760 MIN	TYP	MAX	UNITS
DYNIMIC CHARACTERISTICS[4]						
TDH+N/(0dBFS)	P,U	f_{IN}=1kHz		-92	-90	dB
	P-L, U-L			-90	-88	dB
TDH+N/(-20dBFS)	P,U	f_{IN}=1kHz		-76	-70	dB
	P-L, U-L			-76	-70	dB
TDH+N/(-60dBFS)	P,U	f_{IN}=1kHz		-44	-42	dB
	P-L,U-L			-44	-42	dB
Dynamic Range	P,U	f_{IN}=1kHz, V_{IN}=-60dBFS	104	108		dB
	P-L,U-L		104	108		dB
SNR	P,U	V_N=-0. A Filter	108	110		dB
	P-L,U-L		106	110		dB
Frequency response		f_{IN}=20kHz		±0.1		dB
Channel Separation		f_{IN}=1kHz, A Filter	94	98		dB

NOTES: (4) Average response using a 20-bit reconstruction DAC with 20 kHz low-pass filter and 400 Hz high-pass filter.

Fig. 10.67: ΣΔ converter dynamic characteristics.

In this case, with an output of 20bits, we will have resolution of 5Volt/2^{20}=4.77μV, corresponding to 120dB. The value 106dB indicates that it is limited by the noise floor.

In the ΣΔ, the information about DNL and INL is missed. One can think that it is reasonable because a continuous input could originate typical pattern-noise ‘tones’. Such a phenomenon is surmountable using sinusoidal input signals (or ‘white’) and collecting the histograms for every output code. Therefore, this is not the reason for the lack of these specifications. From the Fig. 10.68, we can notice the noise-shaping effect.

The output data format is parallel-type, i.e. 4 lines for the 4bits of data. Other pins are provided to add external components to modify the characteristics of the integrated circuit. In the integrated digital filter, it is very difficult to have 20bits of output data. For this reason, we use a serial output port with the clock line SCLK. The output (16 or 20bits) can be given with 4 different modes. In Fig. 10.69, there are temporal diagrams for the filter outputs. Finally, a possible circuit is shown in Fig. 10.70. Note the two independent modulator channels (4bits left and right).

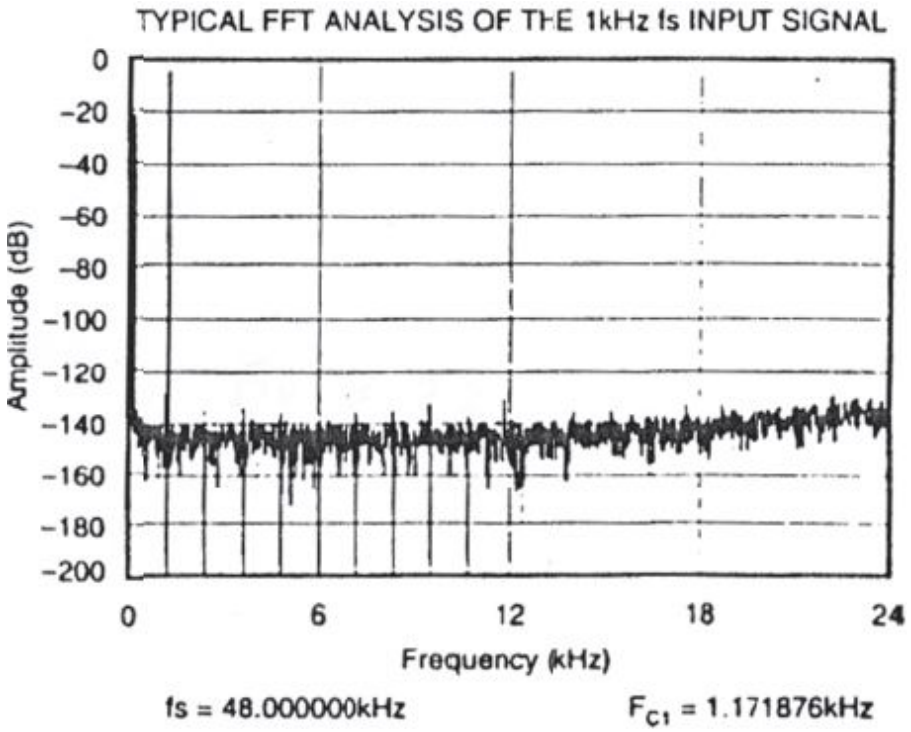


Fig. 10.68: Noise-shaping effect.

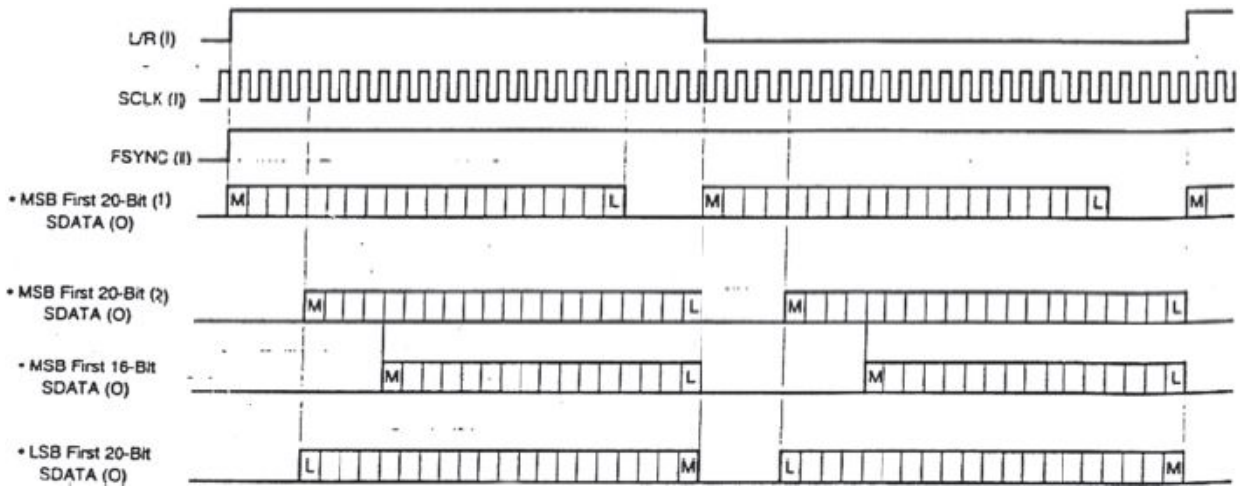


Fig. 10.69: Output filter timing diagram.

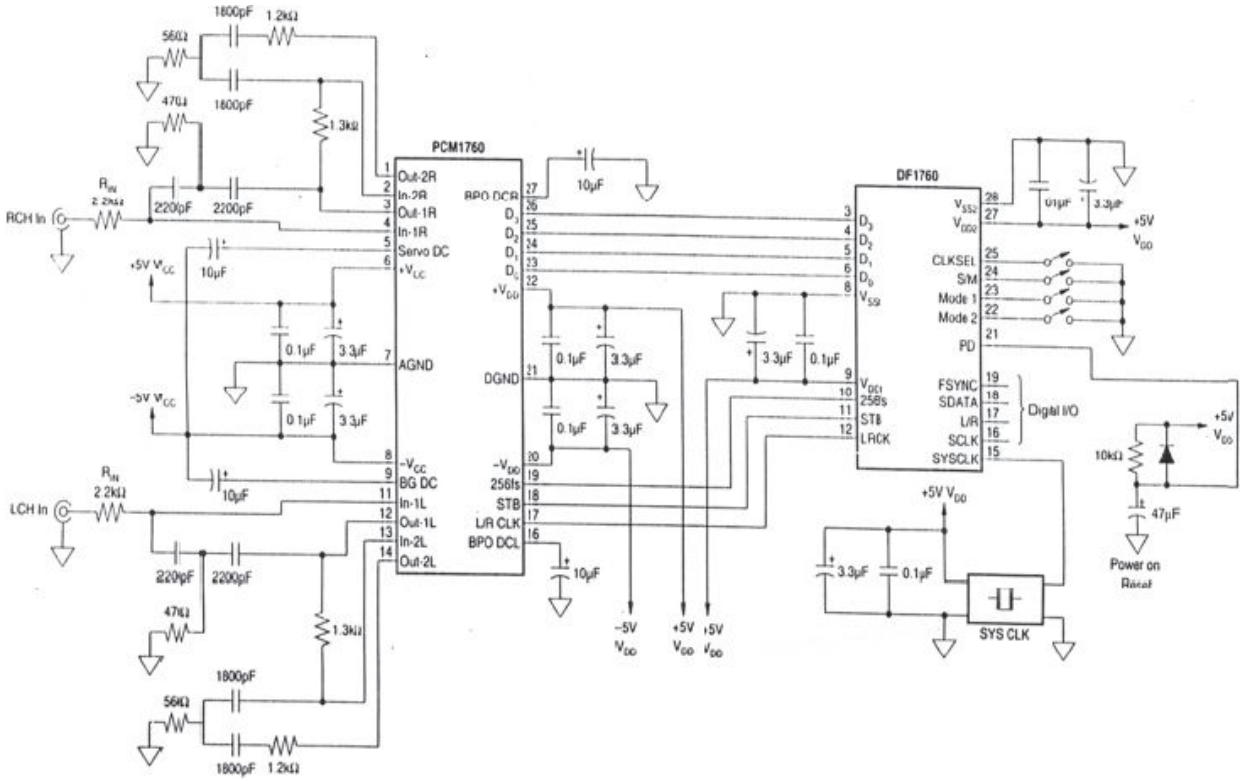


Fig. 10.70: Connections between a PCM1763 and a DF1760.

10.7.2 CS5322 and CS5323

Consider now the couple of Crystal ΣΔ CS5322 and CS5323 (Fig. 10.71). Thanks to these devices, we have a variable over-sampling with a factor 64-

4096, the modulator is a III order, and we have a digital filter with a cascade of two FIR stages. Other characteristics are high resolution (24bits) and modulator's low power consumption (100mW).

The data transfer between two blocks is ensured by only 4 pins; two are used for the low resolution high frequency digital code and the clock CK, while the other two for the synchronism. The static and dynamic characteristics are provided in summary (Fig. 10.72). Note the specifications of the offsets before and after the calibration and the variability range of the calibration itself. The low distortion of this couple is reported in Fig. 10.73.

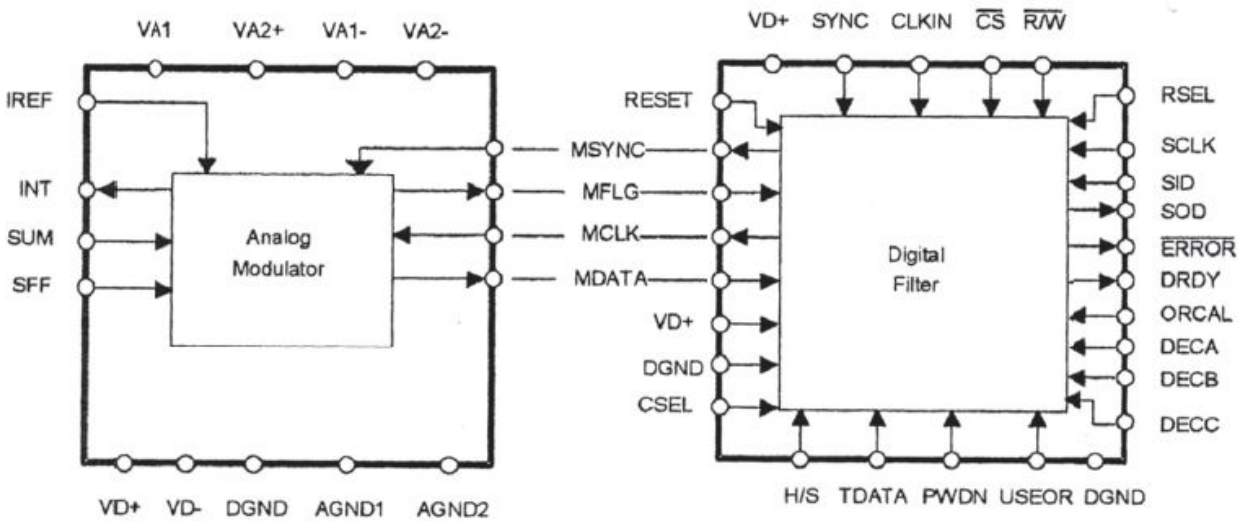


Fig. 10.71: (left) CS5322 (modulator) and (right) CS5323 (digital filter).

Parameter		Min	Typ	Max	Units
Specified Temperature Range		0	-	+70	°C
Dynamic Performance					
Dynamic Range	(Note1)				
CLKIN =1.024 MHz	f_0=4000 Hz	-	103	-	dB
	f_0=2000 Hz	-	118	-	dB
	f_0=1000 Hz	116	121	-	dB
	f_0=500 Hz	-	124	-	dB
	f_0=250 Hz	-	127	-	dB
	f_0=125 Hz	-	129	-	dB
	f_0=62.5 Hz	-	130	-	dB
CLKIN = 512 MHz	f_0=4000 Hz	-	99	-	dB
	f_0=2000 Hz	-	121	-	dB
	f_0=1000 Hz	-	125	-	dB
	f_0=500 Hz	-	127	-	dB
	f_0=250 Hz	-	130	-	dB
	f_0=125 Hz	-	132	-	dB
	f_0=62.5 Hz	-	133	-	dB
SDR	(Note 2)				
	CLKIN=1.024MH	100	110	-	dB
	CLKIN=512MHz	-	120	-	dB
IMD	(Note 3)	-	110	-	dB

Parameter	Min	Typ	Max	Units
Specified Temperature Range	0	-	+70	°C
dc Accuracy				
Full Scale Error	-	-	4	%
Full Scale Drift	-	0.005	-	%/°C
Offset	-	-	250	mV
Offset after Calibration	-	±100	-	µV
Offset Calibration Range	-	100	-	%F.S.
Offset Drift	-	500	-	µV/°C

Fig. 10.72: Static (on the left) and dynamic (on the right) characteristics for the Crystal ΣΔ couple.

10.7.3 CS5317

A ΣΔ converter which collects in a unique IC both a digital filter and a decimator is the Crystal CS5317 (Fig. 10.74). The modulator is a II order, and the FIR has a *sinc* response and a decimation factor equal to 128. Because of the lower order than the Burn Brown PCM1760 one and the lower oversampling compared to the Crystal CS5322, this converter has lower performance levels. For example, it has resolution of 16bits with dissipation of 220mW. However, it is space-saving. For this converter, the DNL and the 'Positive Full Scale Range' are also provided, which is a figure of merit of the committed error when a signal with amplitude close to the maximum allowed is converted. Note that 150mV corresponds to resolution reduced to $\log_2(5V/150mV)=5bit$, instead of the theoretical 16. In Fig. 10.75, there are the static and dynamic characteristics.

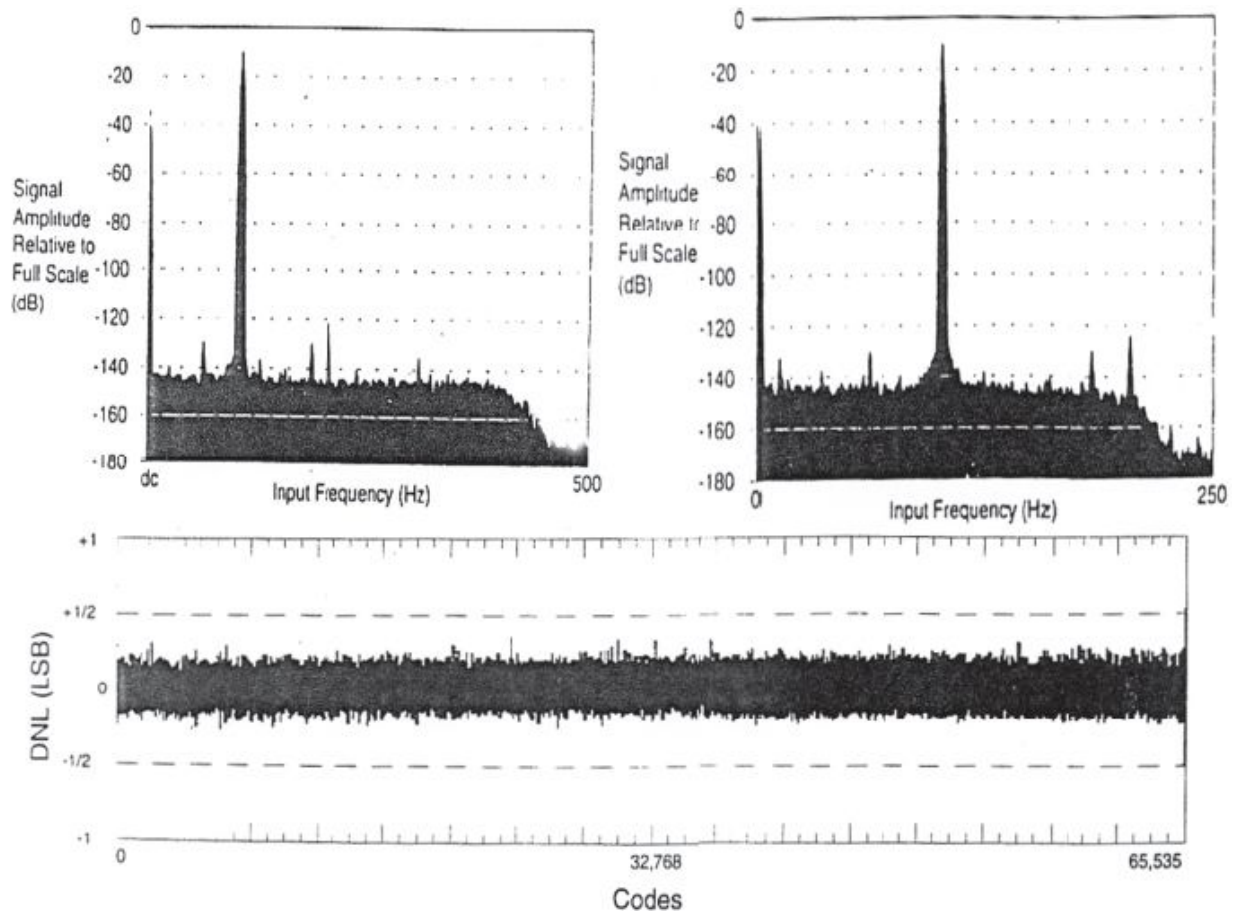


Fig. 10.73: Distortion and DNL for the couple CS5323 and CS5322.

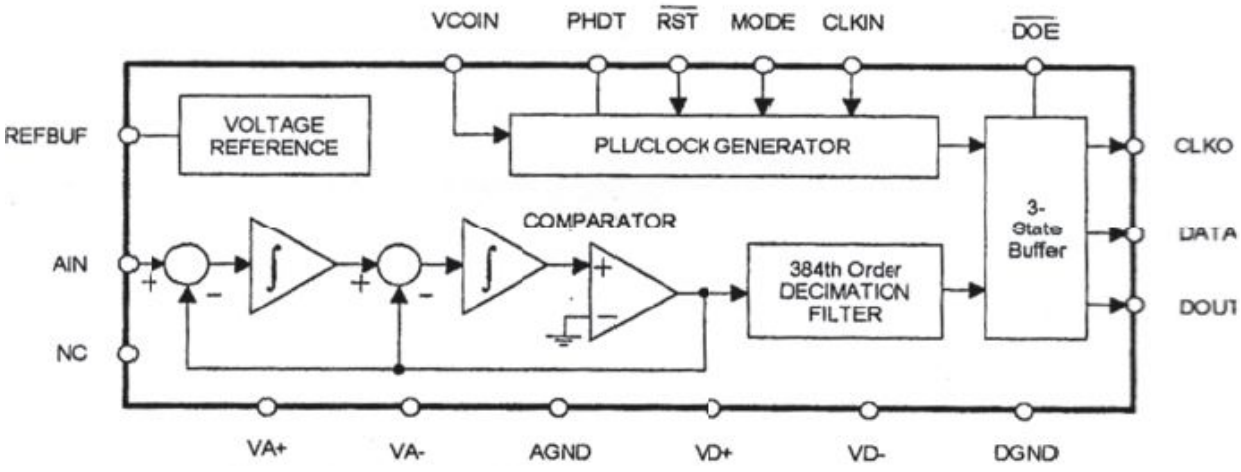


Fig. 10.74: Integrate CS5317.

Parameter	Min	Typ	Max	Units
Specified Temperature Range	0	-	+70	°C
Dynamic Performance				
Dynamic Range (Note1)	78	84	-	dB
Total Harmonic Distorsion	72	80	-	dB
Signal to Intermodulation Distorsion	-	84	-	dB

Parameter	Min	Typ	Max	Units
Specfied Temperature Range	0	-	+70	°C
dc Accuracy				
Differential Nonlinearity (Note2)	-	±0.4	-	LSB
Positve Full Scale Error	-	±150	-	mV
Positve Full Scale Drift	-	±500	-	µV/°C
Bipoar Offset error	-	±10	-	mV
Bipoar Offset Drift	-	±50	-	µV/°C

Fig. 10.75: Static and dynamic characteristics.

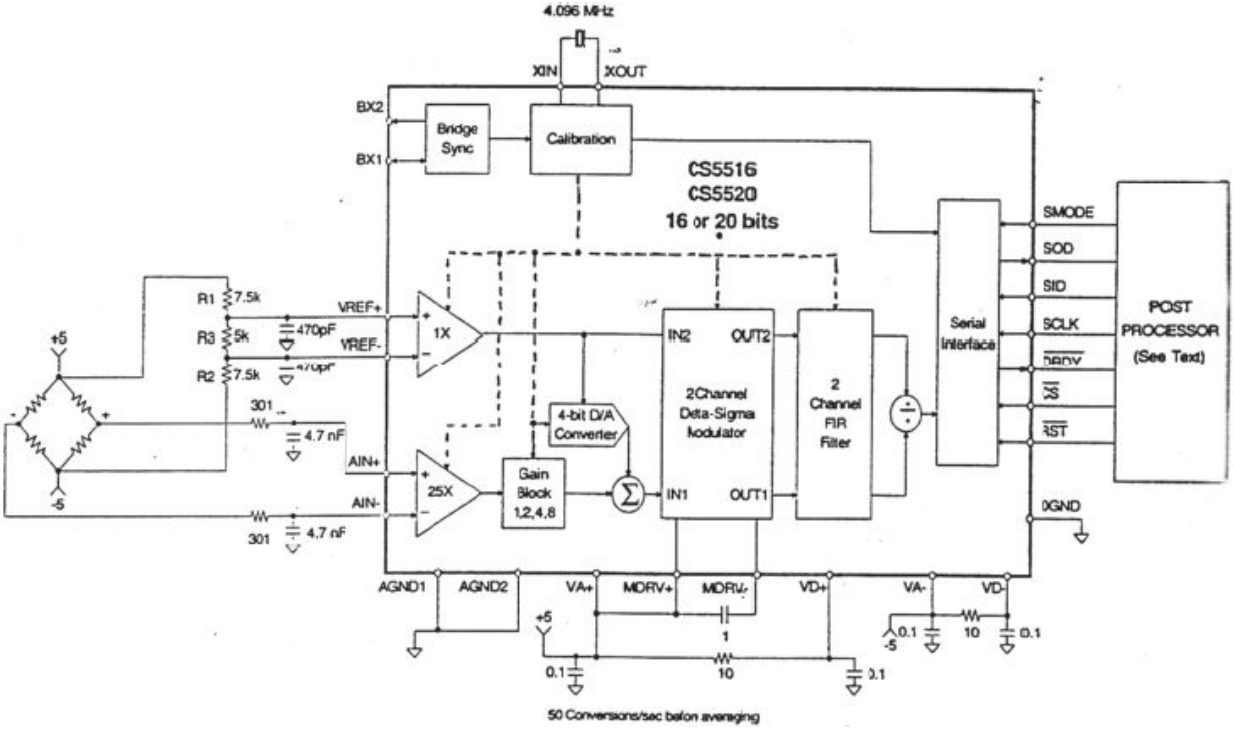


Fig. 10.76: Wheastone bridge application.

10.7.4 ΣΔ application example

A typical ΣΔ application is the front-end for sensors (Fig. 10.76), in which we desire great resolution (16-20bit) and a conversion speed not very high (therefore, it is simpler to increase the oversampling). For example, we can use the ΣΔ Crystal CS55xx. The CS5516 converts at 16 bits with a 60Samples per second while the CS5520 has precision of 20bit. Note that the CS55xx is not only a modulator and a digital filter, but it also has an INA (instrumentation amplifier) with a gain 25x, a PGA (Programmable Gain

Amplifier), and a 4bits DAC to add or subtract an offset to or from the input analog signal (Fig. 10.77). There is also an internal reference voltage source. The logic circuitry allows managing the output data transmission, the power-down function, and the calibration for 8bits registry which manage all the functions of the ΣΔ.

Also the input noise is limited: in the worst case, the minimum gain of the PGA=1, and we have an RMS noise of 250nV$_{rms}$. The noise rejection on the power supply lines (PSRR=100dB) and on the common mode (CMRR=165dB) are optimal. The dynamic characteristics (SNR, THD, and SINAD) are not provided because the application field of this ADC is for DC signals (or, by the way, very slow signals). For this reason, the static characteristics are very important.

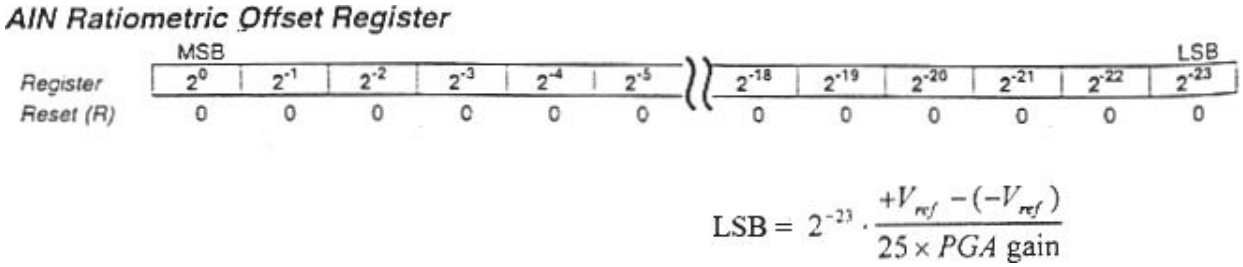


Fig. 10.77: AIN Ratiometric Offset Register; calibration registry. The LSB value is equal to 2^{-23}·[(+Vref)-(-Vref)]/(25·PGA gain).

11

Exercises on converters

“I tuoi corti capelli come sono cambiati
no, non mi dire chi li ha accarezzati.
Fossi un pittore brucerei il tuo ritratto
ma sono solo un amante distratto.”

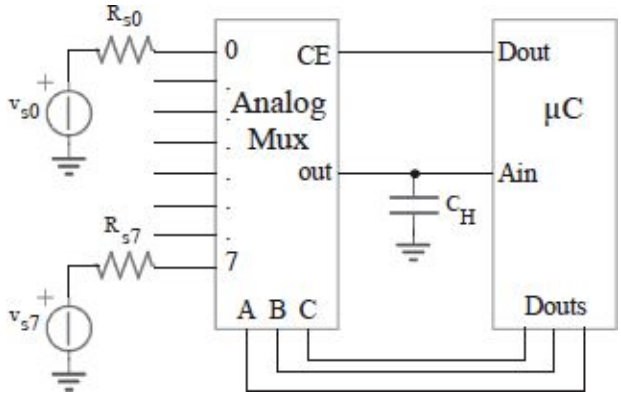
“Più ci penso”, Gianni Bella

11.1

The analog mux (R_{on}<*100Ω*, I_{leak}<*50nA*) with *8* input is controlled by µC with an internal *8 bit* ADC (V_{ref}=*5V*, T_{AD}=*5µs*, $I_{leakAin}$<*50nA*, R_{Ain}<*8MΩ*) without S&H.

a) Compute the maximum value of the R_{Sx} in order to have a *0.5·LSB* precision.

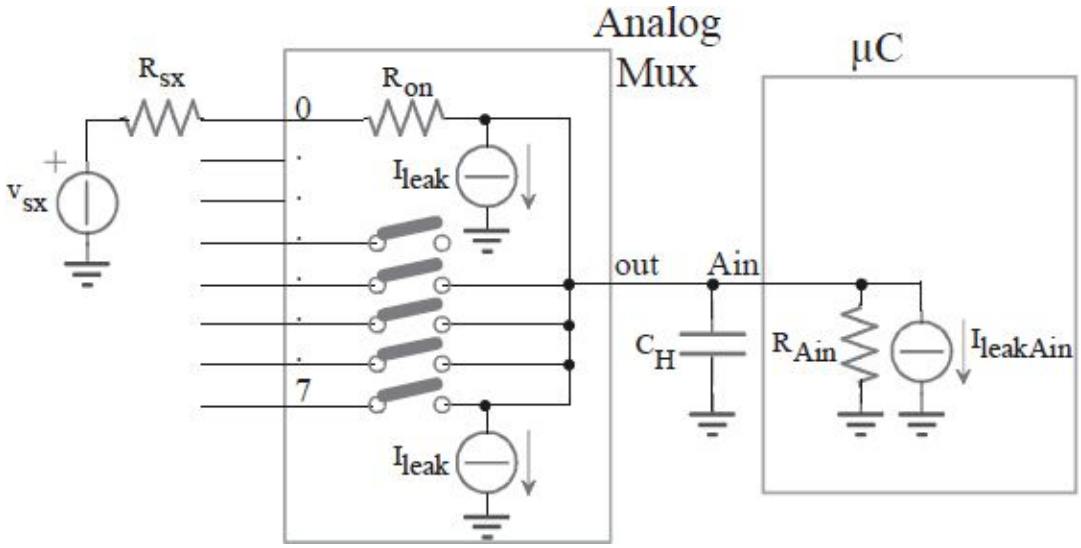
b) Given R_{sx}=*10kΩ* , find the right value of C_H in order to keep a precision of *0.5·LSB* when the system is reading in sequences all the input with a frequency of *1024sample/sec/input*.



11.1.a

It goes without saying that both the µC and the reference voltage of the internal ADC are powered at *5V*. That leads to a LSB of $5/2^8$=*19.5mV*.

The internal structure of the analog multiplexer can be schematized with the resistors R_{on} (switch on) and R_{off} (switch off) of the analog switches given the right leakage equivalent sources. Both the MUX and the µC have to be studied with their own not-ideality. The equivalent circuit is the one in the following picture.



Even if the current sources in the previous circuit are written with a direction, this is not known a priori and so, in order to choose which is the correct one, it's required to check it in the data-book. The same apply to the leakages current source at the input of the µC internal ADC, and it has to be checked in the component data sheet. In the worst case scenario, that is when both sources have the same direction, the maximum voltage gap at the input of the converter is:

$$\Delta V_{leakage} = (8 \cdot I_{leak} + I_{leakAin}) \cdot [(R_{sx} + R_{on}) \parallel R_{Ain}] \approx (8 \cdot I_{leak} + I_{leakAin}) \cdot R_{sx}$$

where we used as hypothesis $R_{on} << R_{sx}$. In order to guarantee a voltage drop less than $\frac{1}{2} \cdot LSB = 9.75mV$, it's required to use a resistor $R_{sx} < 22k\Omega$. This confirm the hypothesis $R_{on} << R_{sx}$ true.

There is one other voltage drop on the signal that is caused by the input resistor of the µC ADC and the other resistors in series:

$$\Delta V_{RAin} = v_{sx} \cdot \frac{R_{sx} + R_{on}}{R_{sx} + R_{on} + R_{Ain}} \approx v_{sx} \cdot \frac{R_{sx}}{R_{Ain}}$$

given R_{Ain} bigger than the other resistors. In order to keep this voltage drop negligible compared to the signal it's mandatory to have $R_{sx}/R_{Ain} < 1/2 \cdot 2^8$, that is $R_{sx} < 16k\Omega$.

The latter of the two statements is the more cogent one. It's important anyway to note that the two effects aren't present one or the other, but both at

the same time, so it's better to keep the source resistance quite under the limit just computed, in the level of $R_{sx}<10k\Omega$.

11.1.b

If we want to take *1024* samples per second the time from two samples of the same input has to be equal or inferior to *1/1024=976μs*. In this period we have to read *8* different inputs, so the overall time per input has to be equal or inferior to *976μ/8=122μs*.

The μC ADC is a successive approximation ADC (SAR), with a conversion time per bit equal $T_{AD}=5\mu s$. In order to fully manage the 8 bit conversion with the reset and restore phases a time of nearby $9.5\cdot T_{AD}\approx 50\mu s$ is needed. This is the conversion time that the μC has to provide for the "proper" conversion. The remaining *122-50=72μs* are used by the circuit for the acquisition of the input that is approximately the charging of the Hold capacitor.

The capacitor CH charges itself throw the equivalent resistor $(R_{sx}+R_{on})//R_{Ain}\approx 10'100\Omega$. The voltage upon its pins rises with an exponential law to the asymptotis value. The error from the actual value and the asymptotic one, the opposite, decrease with an exponential slope:

$$\varepsilon(t) = FSR \cdot e^{-t/\tau}$$

As stated in the previous calculations we can afford a maximum error of *½·LSB*, it's mandatory to wait a minimum charge time before to start the conversions in the μC. The time limit is given by:

$$\tau \le \frac{t}{\ln\left(2\cdot 2^{8}\right)} = 11.5\mu s$$

This leads to a pretty small hold capacitor $C_H<\tau/(R_{sx}+R_{on})=1.1nF$; that is a commercial value of *1nF*.

It's a major interest now to check if the value we chose is low enough to lead to an unexpectedly high amount of droop in the hold phase. This condition has to be checked when all the MUX switches are open and the leakage currents charges or discharges the hold capacitor, changing the voltage at its pins. The voltage *droop ΔV* is:

$$I_{TOTleak} = 8\cdot I_{leak} + I_{leakAin} = C_H \cdot \frac{\Delta V}{\Delta t}$$

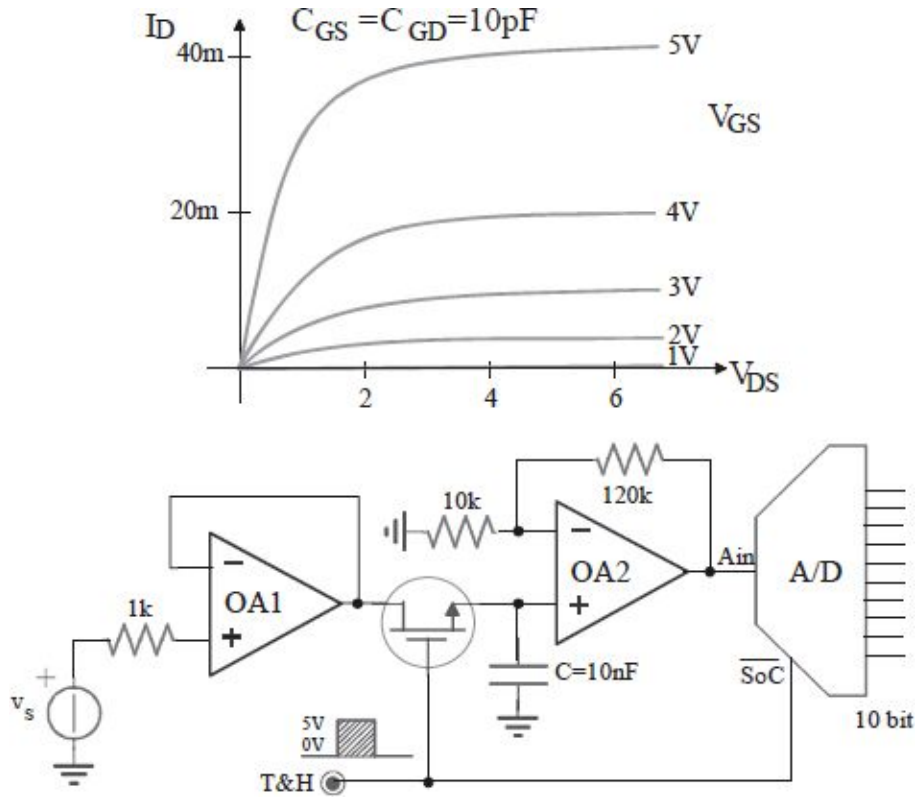
where Δt is the hold time, equal to the minimum conversion time previously computed of $50\mu s$. With those values the voltage droop is $\Delta V=22.5mV$, greater than $1 \cdot LSB$. Due to this error the converter accuracy will be less than the required 8 bit. The ways to solve the problem are lowering the leakages currents, lowering the conversion time (that is the hold time), or rising C_H in pair with a reduction of the input resistance R_{sx}.

11.2

Both the Op Amps with *Ioutmax=10mA, VOS=4mV, IB=50nA, SR=5V/μs* and the MOS with the characteristics in the picture are given.

a) Find out the maximum input frequency *f vS=200mV·[1+sin2πft]*, in order to guarantee a *1LSB* output resolution in the tracking phase ($V_{T\&H}$*=+5V*).

b) Compute the errors due to the not ideality of both the MOS and the Op Amps.



11.2.a

In the tracking phase the signal at the input of the A/D converter has to differ from the ideal one less than *1·LSB*. The ADC reference voltage is supposed to be 5V. This leads at the input of the converter to a LSB value $5/2^{10}$*=4.9mV*.

The second Op Amp has a gain of *1+120k/10k=13*, that leads to a precision in the voltage at the pins of the hold capacitor C_H of *4.9m/13=376μV*.

The real components bring are many causes of errors and not ideality in the circuit. The *slew rate* alone in the Op Amps, that's equal to *SR=5V/μs*, causes the first stage to operate at a maximum frequency of:

$$\left.\frac{dV_{out,OA1}}{dt}\right|_{\max} = \omega_{\max} \cdot V_{p,\max} = 2\pi \cdot f_{\max} \cdot 200m = SR_{OA1}$$

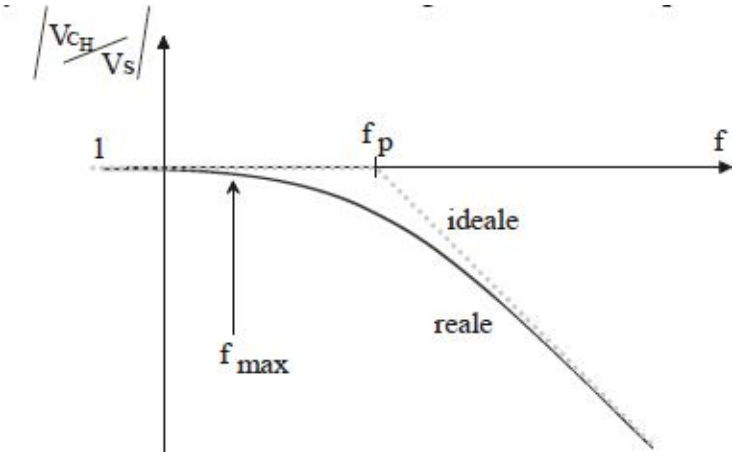
that leads to an f_{max}*=5/1μ/200m/2π=4MHz*. The second Op Amp suffers even more from this effect because of his voltage gain. Due to that gain the peek to peek voltage excursion at the output of the second stage is *±200m·13=±2.6V* with a mean value of *200m·13=+2.6V*. The first thing to note is that the maximum voltage value goes over the dynamic of the A/D converter (*FSR=5V*) and so it will cause a harmonic distortion in the conversion. For this exercise we'll skip this problem and we proceed computing the limitation due to the SR of the second OpAmp:

$$\left.\frac{dV_{out,OA2}}{dt}\right|_{\max} = \omega_{\max} \cdot V_{pOA2,\max} = 2\pi \cdot f_{\max} \cdot 200m \cdot 13 = SR_{OA2}$$

that leads to f_{max}*=5/1μ/2.6/2π=306kHz*, more cogent than the previous one.

The switch is made by a MOS transistor. The tracking voltage input at the gate of $V_{T\&H}$*=+5V* gives an effective voltage V_{GS} that changes depending to the signal v_s. It's easy to see that this signal is always under the value of *+400mV*, that permits to approximate V_{GS}*≈5V*. From the characteristic of the MOS we can see that when operating as triode it's the same as if we're using a resistor R_{on}*≈1/40m=25Ω*. This resistance fixes a low pass filter with time constant R_{on}*·*C_H*=25·10n=250ns* that is a pole at f_p*=637kHz*.

This filter forces a maximum frequency on the signal f_{max} that guarantee the *±1·LSB* accuracy. This is easily found also in the Bode diagram of the low pass filter below.



The inverting gain can be written with the following transfer function:

$$\left|\frac{V_{CH}}{V_s}(f)\right| = \frac{1}{\sqrt{1+\left(f/f_p\right)^2}}$$

The application of an input signal v_s with a peek to peek amplitude of *200mV* and frequency f_{max} causes at the output of the filtering stage a sinusoid with amplitude lower than the one at DC. This difference in amplitude can be estimated as:

$$\Delta v_{CH}(f_{max}) = v_{CH}(0) - v_{CH}(f_{max}) = v_s - v_s \cdot \frac{1}{\sqrt{1+\left(f_{max}/f_p\right)^2}} \approx v_s \cdot \left[\frac{1}{2}\left(f_{max}/f_p\right)^2 + \ldots\right]$$

in where we expanded the root square. To keep this difference under the $\pm 1/2 LSB$ value of *188μV* it's required to keep:

$$\frac{f_{max}}{f_p} \le \sqrt{\frac{2}{2^n}} = \frac{1}{22.6}$$

that is, in order to guarantee a 10 bit accuracy it's required to have an input signal with $f_{max}<28kHz$.

This last limitation is caused by the maximum current that the OA1 can supply to charge the hold capacitor C_H. The worst case scenario is the one with

R_{on}≈*0*, that requires a maximum current of:

$$I_{out\,\max} = C_H \cdot \left.\frac{dv_{CH}}{dt}\right|_{\max} = C_H \cdot 2\pi \cdot f_{\max} \cdot v_s$$

this case leads to an f_{max}=*796kHz*.

In conclusion the maximum frequency that can be applied at the overall stage is limited by the analog low pass filter caused by the resistance of the MOS switch. This forces a maximum frequency of *28kHz* at the input in order to guarantee the *1·LSB* resolution at the whole signal.

11.2.b

The offset of the first OpAmp causes an error equal to *±4m/376µ=10.6·LSB*. The offset of the second OpAmp causes the same amount of error, because of it has no voltage gain (it's in fact connected as a buffer). This can lead to an error of *2·10.6=21·LSB* in the worst case scenario, that is equal to say that nearby *5* bit of the *10* available are incorrect.

The bias current flowing in OA1 forces a voltage drop at the input of *1k·50n=50µV*, that is less than *1·LSB*. The inverting pin of the OA2 forces his current in the 120k resistor (the other input is grounded for the superposition effect, and so no current can flow in the 10k resistor) that leads to a voltage drop of *120k·50n=6mV* at the input of the converter, slightly more than a *1.2·LSB*. The effect of the bias current at the not inverting pin is futile, because of it flows through the R_{on}≈*25Ω* of the MOS and then have a gain of *1+120k/10k=13* (that is equal to flow in an effective resistor of *25·13=325Ω*, much lower than the *120kΩ* seen at the inverting pin).

In the Hold phase, when the MOS is interdict, the not inverting pin bias current causes a voltage droop at the pins of C_H. In order to prevent a charge higher than *1·LSB* it's required to choose a hold phase time T_H less than the once computed below:

$$I_B = C_H \cdot \left.\frac{dv_{CH}}{T_H}\right|_{\max} = C_H \cdot \frac{2 \cdot v_s / 2^{10}}{T_H}$$

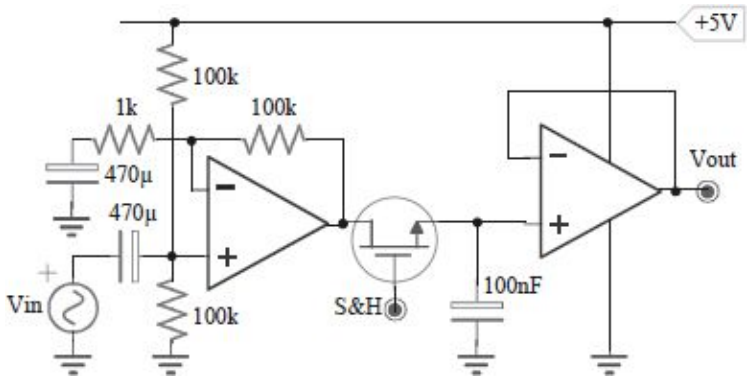
that with current values T_H<*78µs*.

11.3

The S&H is followed by a *12bit* ADC described with *FSR=5V*. The OpAmp are described with *SR=1V/µs* and I_B*=100nA*, while the MOS has R_{on}*<10*Ω.

a) Find out the maximum sample time and hold time.

b) Compute the maximum frequency of the input signal given $|v_{in}|$*<20mV*.



11.3.a

The first OpAmp has a medium frequency gain given by *1+100k/1k=101*. If the output is an ideal rail to rail, the signal can afford an output voltage swing between *0* and *+5V*. In DC the exit is maintained at midrange, that is *+2.5V*. The MOS, the capacitor and the second OpAmp (that is in a buffer configuration) are a simple Sample&Hold.

When the MOS switch is closed the gathering of the signal start (sampling). In this stage the voltage drop at the capacitor has to reach the correct value in the fewer time possible. There are two effects that slow down the capacitor charge: the first OpAmp slew rate and the exponential charge of the capacitor through the MOS resistor R_{on}.

The worst case for the transient is obviously when we have to charge the capacitor from the minimum value (*0V*) to the maximum one (*+5V*) allowed by the circuit. In the case of a perfect voltage step at the input of the MOS, the charge of the capacitor would follow the exponential relationship:

$$v_c(t) = \Delta V \cdot \left(1 - e^{-t/Ron \cdot C}\right)$$

in where the max slope is reached at the start, and is equal to:

$$\left.\frac{dv_c}{dt}\right|_{\max} = \frac{\Delta V}{R_{on} \cdot C}$$

That in our case can be computed in *5V/μs*.

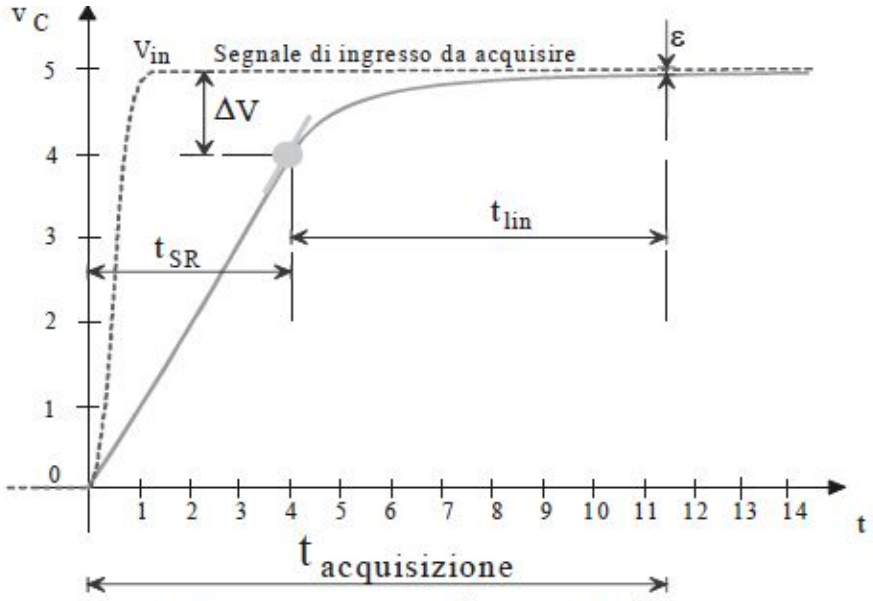
The OpAmp slew rate forces a maximum charge slope equal to *1V/μs*. This will change only after some time, or to be more exact when the voltage drop to charge the Hold capacitor lower enough to the level of:

$$\left.\frac{dv_c}{dt}\right|_{\max} = \frac{\Delta V}{R_{on} \cdot C} = SR$$

This can be computed in a voltage drop of $\Delta V = SR \cdot R_{on} \cdot C = 1V/\mu s \cdot 10\Omega \cdot 100nF = 1V$ from the full charge value *5V*.

As we can see from the following picture the first part of the charging is limited by the slew rate, and thus linear, while the second part follows the usual exponential law. This first phase will take $t_{SR} = (5 - \Delta V)/SR = (5-1)/1V/\mu s = 4\mu s$ to end. The second one instead can be choosed by the accuracy of the voltage we need.

The accuracy of the signal we're getting is forced by the accuracy of the *12bit* ADC converter. In order to keep the error to less than *1·LSB* (that is $5/2^{12} = 1.2mV$) when the *FSR* is *5V* we have to reach an asymptotic value of the signal that differ from the ideal value no more than $\varepsilon = \frac{1}{2} \cdot LSB$. The sample time of this second phase can be computed as:



$$t_{linear} \geq \tau \cdot \ln \frac{\Delta V}{1/2 \cdot LSB} = R_{on} \cdot C \cdot \ln\left(2^{12+1} \cdot \frac{\Delta V}{FSR}\right) = 1\mu s \cdot 7.4 = 7.4\mu s$$

It is important to note that the voltage drop in the formula is not the full dynamic of the system *FSR=5V*, but just the last *ΔV=1V*. It's easy to see that we can't wait only the usual *3÷4* time constant, but we have to wait for at least *7.4* time constant in order to guarantee the required accuracy. The overall sample time is given by the sum of the two terms computed:

$$t_{sample} \geq t_{SR} + t_{lineare} = 4\mu s + 7.4\mu s = 11.4\mu s$$

In the Hold phase the MOS acts as an open circuit and the capacitor voltage is transferred to the output by the second OpAmp in buffer configuration. In this phase it's primary that the voltage at the capacitors don't gets any droop. There are two primary reasons for this droop to be: leakage current in the capacitor and bias current of the buffer OpAmp $I_B=100nA$.

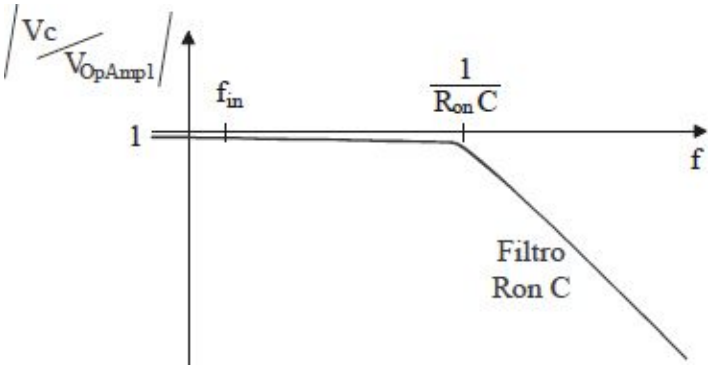
This constant current causes a constant droop of the capacitor voltage. In order to keep this droop lower than the usual *½·LSB* we have to keep the Hold phase short in the order of:

$$t_{Hold} \leq C \cdot \frac{1/2 \cdot LSB}{I_B} = 100nF \cdot \frac{1/2 \cdot 1.2mV}{100nA} = 600\mu s$$

11.3.b

In the previous exercise we studied how the Sample&Hold circuit works. We have focused on the sampling of the signal from the sampling switch to close to the full charge of the capacitor. In the case of an input signal that changes its own value during the sampling time, it sees the R_{on}·C network as a filter, thus we will find at the capacitor pins a lowered copy of the signal itself. This concept leads to the definition of the analog bandwidth of the S&H circuit.

This filtering network has a time constant equal to *10·100n=1μs*, that leads to a pole at *159kHz*. Given an input signal at the same frequency of the pole, the output of the S&H circuit would follow the signal with a *-3dB* attenuation: instead of a peek to peek sinusoid of *20mV·101=2.02V* we'd have at the output a peek to peek excursion of only *2.02/√2=1.43VI*. The loss in voltage is equal to *2.02-1.43=0.6V* that in LSB is *0.6/1.2m=500·LSB*! This is lowering the accuracy of the converter to *3bit* when we're supposed to use *12*.



In order to guarantee the usual *±½·LSB* accuracy at *12bit* it's mandatory to position the pole at a ways higher frequency than the signal, or vice versa to use a signal ways lower in frequency than the pole. This condition is drawn via the Bode diagram in the picture below, in where we have taken also the first OpAmp open loop not inverting bandwidth, with a gain of *101*. The flow of the gain function can be written as:

$$\left|\frac{v_c}{v_{OpAmp1}}(f)\right| = 1 \cdot \frac{1}{\sqrt{1+\left(f_{in}/f_p\right)^2}}$$

Given a signal with amplitude v_{OpAmp1} and frequency f_{in} at the input of the MOS the resulting signal at output is a sinusoid with lower amplitude than the

one registered with a lower frequency input. The amplitude gap can be estimated as:

$$\Delta v_c(f_{in}) = v_{OpAmp1}\cdot 1 - v_{OpAmp1}\cdot 1\cdot\frac{1}{\sqrt{1+\left(f_{in}/f_p\right)^2}} \approx v_{Opamp1}\cdot 1\cdot\left[\frac{1}{2}\left(f_{in}/f_p\right)^2+\ldots\right]$$

In which the square root has been expanded. In order to keep the difference below a $\pm x\cdot LSB$ given maximum amplitude *FSR*, it's mandatory to guarantee

$$\Delta v_c(f_{in}) \le \pm x\cdot\frac{FSR}{2^n}$$

With a little math:

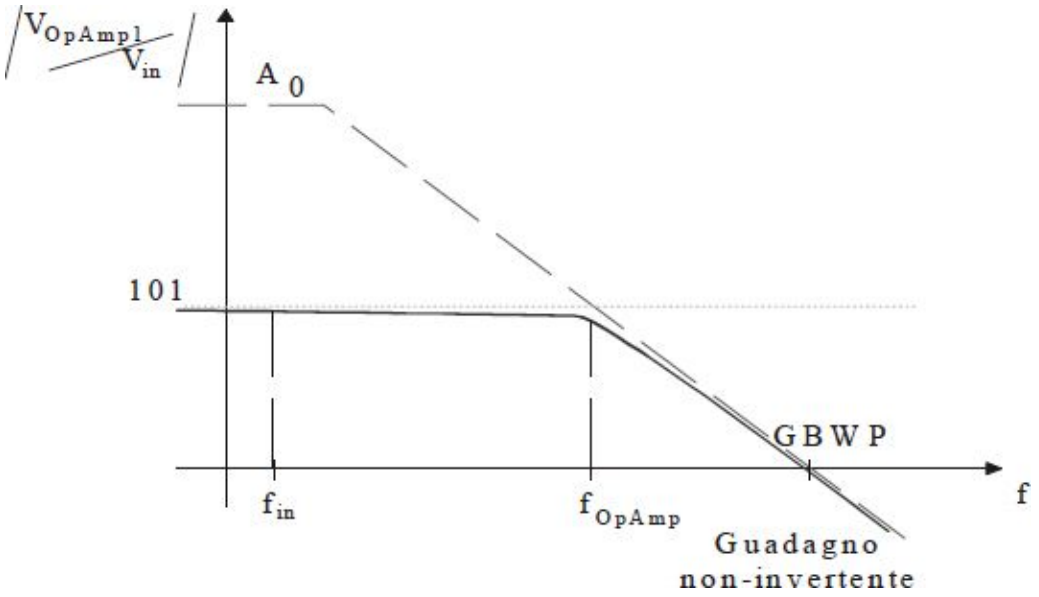
$$\frac{f_p}{f_{in}} \ge \sqrt{2^n\cdot\frac{1}{2}\cdot\frac{1}{x}\cdot\frac{v_{OpAmp1}\cdot 1}{FSR}} = \sqrt{2^n\cdot\frac{1}{4}\cdot\frac{1}{x}\cdot\frac{v_{OpAmp1,peek}}{v_{OpAmp1,peek,\max}}} \ge \sqrt{2^n\cdot\frac{1}{4}\cdot\frac{1}{x}}$$

This last equation is correct only in the case the peek amplitude of the signal $v_{OpAmp1,peek}$ at the output of the first OpAmp is equal to the maximum amplitude allowed by system $v_{OpAmp1,peek,max}=FSR/2$. In every other case the condition is less hard.

In our case, with a maximum error of $\pm\frac{1}{2}\cdot LSB$ ($x=\frac{1}{2}$) on a *12bit* ($n=12$) scale, we will have to choose an input frequency of $f_{in}<f_p/45$ that is *3.9kHz*.

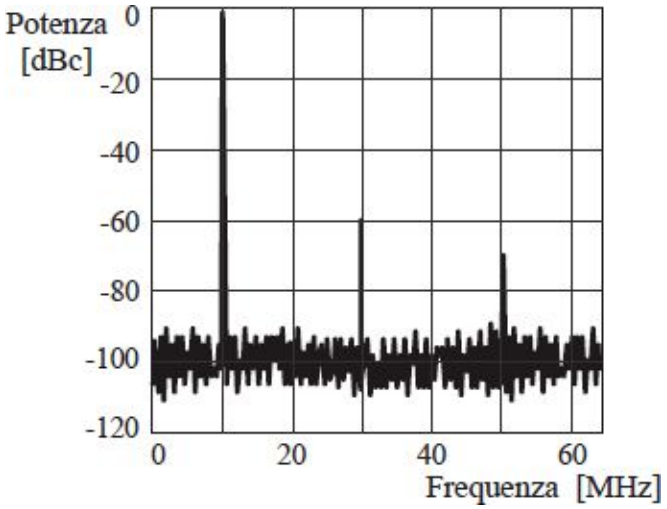
The analog bandwidth of the stage is also limited by the OpAmp itself. The OpAmp, as illustrated by the Bode diagram in the picture below, is limited by itself the pole given by the value of his real *gain bandwidth product*, *GBWP*. If we want to overlap the pole of the closed loop OpAmp f_{OpAmp} and the one we computed for the S&H circuit $f_p=1/R_{on}\cdot C=159kHz$, we need to use a OpAmp with a $GBWP>101\cdot f_{OpAmp}=16MHz$!

This last request is very expensive and thus has not to be taken carelessly. In fact, given a common cheap OpAmp with *GBWP=0.5MHz*, it would have been this the most cogent request, thus the limit in the analog band of the S&H would have been lowered to *0.5M/101=4950Hz* and the one of the input signal to $f_{in}< 4950/45=110Hz$.


V_{OpAmpl}
V_{in}
A_0
101
GBWP
f
f_{in}
f_{OpAmp}
Guadagno
non-invertente

11.4

The S&H SHC605 works at frequency f_s=*128MHz* with a sinusoidal input peek to peek ±*1V*. The spectrum at the output has a *bin-width=1kHz*.
a) Compute both *SiNAD* and *SFDR*.
b) Compute the overall harmonic distortion.



11.4.a

The Sample & Hold we're studying has been described with a sampling frequency of f_s=*128MHz*. For this reason a spectrum analysis has been taken at the output and the graph shows it at the base band between *0Hz* e $f_s/2$=*64MHz*. As seen in the picture, the S&H has been tested with a sinusoid at the input with frequency *10MHz*, and mean value null (the input goes from *–1V* to *+1V*).

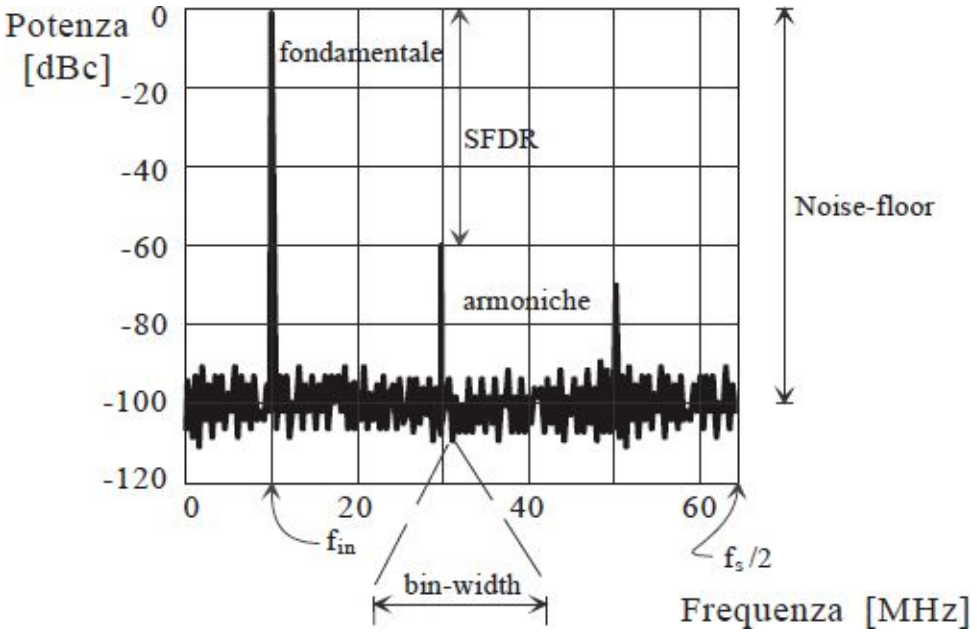
Sadly due to the not ideality of the S&H in the spectrum shown there's not only the peek at the input frequency, but it shows also both 3rd and 5th harmonic at *30MHz* and *50MHz*. For that reason at the output we'll have not only the 1V peek voltage, but also some noise, the highest of which is the third harmonic with a *–60dBc* power. This is the *Spurious-Free Dynamic Range*, *SFDR*, of the circuit.

If we want to quantify the power of the noise, it's mandatory to note that the y axis of the graph is referred to the carrier signal, and so it's a relative value, and not a strict one in W o mW. This is in fact the means of the char "c" after the "dB", thus in order to emphasize that the value is referred to the "carrier" signal. It's obvious this way that the carrier signal has 0dBc amplitude, irrespective of the peek value of the signal itself. From all this we can compute the carrier power as:

$$P_{0dBc} = \left(\frac{V_p}{\sqrt{2}}\right)^2 = \left(\frac{1V}{\sqrt{2}}\right)^2 = (707mV)^2$$

in where the voltage hasn't been raised at the second due to the loss of physical meaning of the V^2 (it refers in fact to a power at a 1Ω resistor). The distortion at the 3rd harmonic correspond to a sinusoid at *30MHz* with power that is $10^{-60/10}=10^{-6}$ times less than the carrier, that leads to peek value $10^{-60/20}=10^{-3}=1/1'000$ times under the carrier itself. In conclusion, the distortion caused by the third harmonic has a peek value of $707m/1'000=707\mu V$ and the one due to the 5th is at $707mV \cdot 10^{-70/20}=224\mu V$.

Every histogram in the spectrum has a width called *bin-width*, that in this case is *1kHz*. Thus in the overall we have $f_s/bin\text{-}width=128M/1k=128'000$ histograms. All of those has been computed through Fast Fourier Transform (FFT) with the same number $N=128'000$ of samples taken at the output of the S&H.



At this point it's easy to find out the *Noise-floor* level, that is the noise power in each histogram, and that is equal to *–100dBc*, thus equal to $(707mV)^2 \cdot 10^{-100/10} = (7.07\mu V)^2$, that is an effective *7.07µVrms* for each *1kHz* of bandwidth. In order to get the value of the noise at the whole bandwidth we have just to compute the integral of the mean of each histogram from *0Hz* to $f_s/2$, that are *N/2=64'000*. The value that it's obtained in our case is $(7.07\mu V)^2 \cdot N/2 = (1.8mV)^2$.

If we want just to compute the value of the noise, without the harmonic distortion, directly in *dBc*, we would have been able to find it out as *–100dBc·N/2*. It's mandatory not to be fooled from this equation, in fact the mean is not *–100dBc·64'000=6.4MdBc* because in order to compute the product we have to use two values with the same mean. That is, we have to transform the factor *64'000* in dB (power), and the result is $10 \cdot log_{10} 64'000 = +48dB$.

The overall noise power computed this way in *dBc* is equal to *–100dBc+48dB=-52dBc*. In order to check the result we can simply compute the effective value of the noise power, that is $(707mV)^2 \cdot 10^{-52/10} = (1.8nV)^2$, as expected.

The value of *–52dBc* we just found is the real *Signal to Noise* ratio, SNR_{real}, of the Sample & Hold, thus it's already normalized to the power of the carrier

signal. This SNR_{real} will differ from the ideal one that consider only the quantization error at *n bit*, that would lead to $SNR_{ideal}=6.02 \cdot n+1.76\ dB$.

At the end it's possible to say that if we keep only the white noise we have:

$$White\ noise = (Noise - floor)_{dBc} \cdot 10 \cdot \log_{10} \frac{N}{2} = -52 dBc$$

We just find out that there's not only the white noise to worsen the signal, but also some harmonic distortion in the output bandwidth. The power of those 2 harmonic components is focused in 2 histograms, with respectively *–60dBc* and *–70dBc* amplitude. In order to sum the power it's mandatory NOT to sum the two *dBc* values, but to convert them in effective power and then sum them; after this we can convert the sum to a *dBc* value. It's simply conceptually wrong to sum the two as harmonic components as *–60dBc-70dBc=-130dBc*, and so the real result is:

$$Hrmonic = 10 \cdot \log_{10} \left[\sum_{i=2}^{\infty} 10^{-dBc,i-esima/10} \right] = 10 \cdot \log_{10} \left[10^{-60/10} + 10^{-70/10} \right] \approx -59.6 dBc$$

We may notice that if there is an input with heavier weight among the other, the total power will be slightly higher (the value in dBc, in module, thus slightly lower) than the heavier harmonic.

The very same computation should be done for every other noise or distortion in the bandwidth, no matter of the causes. Those can be pretty varies, as the *feed-through* of the clock in the S&H, the *ripple* of the power supply, crosstalks through the wires, and so on.

The value of the distortion should be written as:

$$Distorsion = 10 \cdot \log_{10} \left[\sum_{k=1}^{toni} 10^{-dBc,k-esimo/10} \right] = 0 = -\infty dBc$$

It's mandatory to sum all those parts in the computation of the *SiNAD*:

$$SiNAD = \frac{Signal}{White\ Noise\ +\ Harmonic\ +\ Disturb} =$$
$$= \left[0dBc - \left(White\ Noise\ +\ Harmonic\ +\ Disturb\right)_{dBc}\right]_{dB} =$$
$$= 10 \cdot \log_{10}\left[10^{White\ Noise,dBc/10} + 10^{Harmonic,dBc/10} + 10^{Disturb,dBc/10}\right] =$$
$$= 10 \cdot \log_{10}\left[10^{52/10} + 10^{59.6/10} + 0\right] = 51.3dB$$

11.4.b

The total harmonic distortion is defined as the ratio between the power sum of the effective values of the first nine harmonic copies and the effective value of the main signal. In the current case we have only two harmonic copies, the 3rd and the 5th, and the sum of the power is *–59.6dBc*. The correct formula is:

$$THD = 20 \cdot \log_{10}\left\{\frac{\sqrt{\sum_{i=2}^{9} 10^{-dBc,i-esima/10}}}{V_1}\right\} = 20 \cdot \log_{10}\left\{\frac{\sqrt{(V_1)^2 \cdot 10^{-60/10} + (V_1)^2 \cdot 10^{-70/10}}}{V_1}\right\} =$$
$$= 20 \cdot \log_{10}\sqrt{10^{-60/10} + 10^{-70/10}} = 10 \cdot \log_{10}\left(10^{-60/10} + 10^{-70/10}\right) = -59.6dBc$$

In our case the calculation is simplified because all powers in *dBc* are already referred to the main signal. It’s easy to note also that it has been used a factor of 20 for the logarithm because of we’re working with voltage ratioes and not with powers (that would need the factor 10, instead). In both cases, the final result would have been the same.

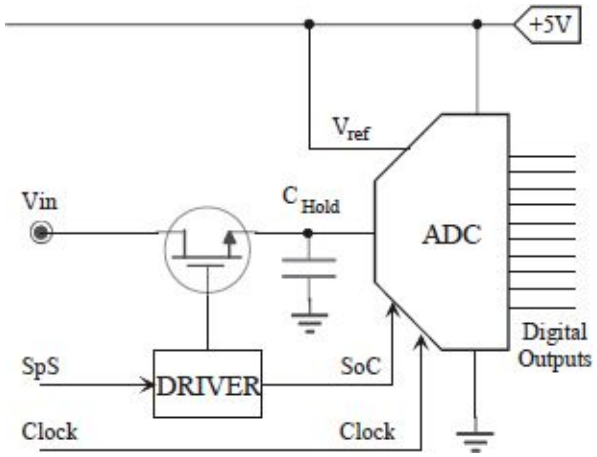
11.5

It's given the circuit of the Ex.5.2 in Chapter 5. Connect to the output of that circuit the double ramp *10bit* ADC in the picture.

a) Choose the value of *CF* in the filter of the Ex.5.2 in Chapter 5 in order to guarantee a *1LSB* accuracy.

b) Choose the *clock* frequency in order to guarantee 5*12 SpS* and then the value of *CHold* and *Ron* of the mosfet, given an ADC leakage current of 50nA.

c) Draw the FFT qualitative spectrum at the exit of the ADC. Bear in mind the power noise with and without C*bypass*.



11.5.a

Capacitor C_F causes a low pass filtering of the input signal. Because of this, the gain of the stage will decrease from the mid frequency value of *22* down to nil. The pose has a frequency of $1/2\pi \cdot C_F \cdot 220k$.

Given an input signal with the same frequency of the pole, it would get reduced by a mean of *-3dB*: thus instead of a sinusoid with peek value of $100m \cdot 22 = 2.2V$ it would show as $2.2/\sqrt{2} = 1.6V$. The loss of $2.2-1.6=0.6V$ is equal to $0.6/4.9m = 122 \cdot LSB$! This would downgrade the accuracy of the ADC to just *3bit*, even if the real ADC accuracy is *10bit*.

In order to guarantee a *±1·LSB* accuracy at *10bit* it's mandatory to position the pole of the low pass filter ways higher than the frequency of the signal, which has been given as *20Hz*. This condition is shown in the following Bode diagram.

The flow of the inverting gain and be expressed as:

$$\left|\frac{v_{out}}{v_s}(f)\right| = G_0 \cdot \frac{1}{\sqrt{1+\left(f_{in}/f_p\right)^2}}$$

If we apply at the input a sinusoid with amplitude v_s and frequency f_{in} we'll see at the output a sinusoid with amplitude lowered than a DC signal (if we assume the *470µF* as a short circuit) or, at least, with a ways lower frequency. The ratio can be expressed as:

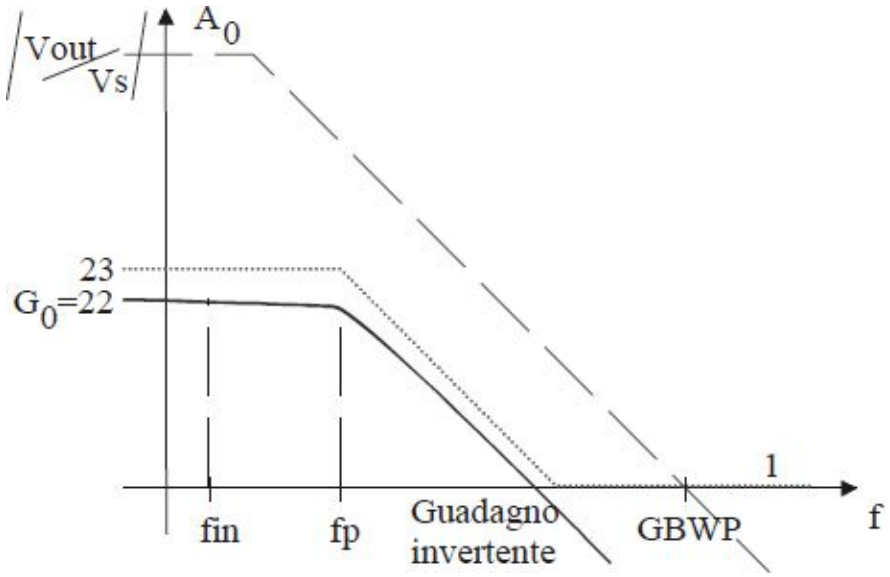
$$\Delta v_{out}(f_{in}) = v_{out}(0) - v_{out}(f_{in}) = v_s \cdot G_0 - v_s \cdot G_0 \cdot \frac{1}{\sqrt{1+\left(f_{in}/f_p\right)^2}} \approx v_s \cdot G_0 \cdot \left[\frac{1}{2}\left(f_{in}/f_p\right)^2 + \ldots\right]$$

in where we expanded as power series the square root. In order to get an error less than *±x·LSB*, with a peek value of *FSR/2*, it's required to guarantee:

$$\Delta v_{out}(f_{in}) \leq \pm x \cdot \frac{FSR/2}{2^n}$$

From which we can get:

$$\frac{f_p}{f_{in}} \geq \sqrt{\frac{2^n}{2} \cdot \frac{1}{x} \cdot \frac{v_s \cdot G_0}{FSR/2}} = \sqrt{\frac{2^n}{2} \cdot \frac{1}{x} \cdot \frac{v_{out,picco}}{v_{out,max}}} \geq \sqrt{\frac{2^n}{2} \cdot \frac{1}{x}}$$



This last disequation is correct only in the case of peek value $v_{out,picco}$ equal to the maximum allowed by the system $v_{out,max}=FSR/2$. In every other case the condition will be less cogent.

In our case, in order to get an error less than $\pm 1 \cdot LSB$ ($x=1$) with a *10bit* ($n=10$) accuracy we'll have to choose a pole frequency $f_p > 22.6 \cdot f_{in}$ that leads to a pole at *452Hz*. This condition fixes the value of the capacitor $C_F > 1/2\pi \cdot 452 \cdot 220k = 1.6nF$; we'll get the *1nF* commercial capacitor.

It's useful to note that the oversizing of the filter can be extremely hard to find, mostly if there are strictly conditions of low errors on a high number of bits. The following table shows off some of the most meaningful values for the most common or important cases. As example, if we want to read a 1kHz input signal with 16bit accuracy and a fourth of LSB as error with a gain of 20, it's required to position the pole at *724kHz* and to use an OpAmp with *GBWP=15MHz*.

	f_p/f_{in}		
n	**¼·LSB**	**½·LSB**	**1·LSB**
8	46	**32**	23
10	91	**64**	46
12	181	**128**	91
14	362	**256**	181
16	724	**512**	362
18	1450	**1024**	724
20	2897	**2048**	1450

11.5.b

In order to have *512* conversion per second it's required to have a conversion start at every *1/512=1.95ms*. If we can suppose the acquisition time as null, that is an instant charge of the hold capacitor, the time just computed would be the time given to the converter for the charge and discharge ramps of the integrated capacitor in the ADC. With the given *10bit* accuracy, the clock frequency required can be computed as $2 \cdot 2^{10} \cdot 512 = 1'048'576Hz$. To be honest we have to give a reasonable acquisition time for the Hold capacitor to charge in the order of *50μs*, thus the time given to the ADC for the conversion is *1.9ms*. The minimum clock frequency results to be *1.1MHz*.

While the conversion takes time we have to check that the ADC leakage current at the input do not discharge too much the Hold capacitor. If we keep as limit a droop of ± *x·LSB* it's required to use:

$$C_{Hold} \geq I_{leakage} \cdot \frac{T_{Hold}}{x \cdot FSR/2^n} = I_{leakage} \cdot \frac{T_{Hold}}{FSR} \cdot \frac{1}{x} \cdot 2^n$$

That for the requested accuracy of *±1·LSB* (*x=1*) with *10bit* leads to $C_{Hold} > 19nF$; we can choose a standard *22nF* commercial capacitor.

In order to guarantee the *50μs* sampling phase from the MOS switch to close to the full charge of the capacitor, we have to check that the time constant of the capacitor charge is fast enough. As usual we'll check the worst case scenario, that is then the capacitor is at the start fully discharged and has to charge up to the maximum value of *FSR*. The flow of the charge will be an exponential if there are no other factors at play such as the OpAmp slew-rate, or its $I_{out,max}$, or the saturation of JFET switch, etc...

$$v_{CHold}(t) = FSR \cdot \left(1 - e^{-t/\tau}\right)$$

the error to the asymptotic value, equal to *FSR*, has a flow $\varepsilon(t) = FSR \cdot e^{-t/\tau}$. With a maximum time of *t=50μs* and a maximum error of $\varepsilon < x \cdot LSB$, we can compute:

$$\tau \leq \frac{t}{\ln\left(\frac{2^n}{x}\right)}$$

in our case $\tau<t/6.9=7.2\mu s$. The maximum resistor allowed for the MOS switch is $7.2\mu/22n=327\Omega$, that is a reasonable value for a common MOS.

	t/τ		
n	¼-LSB	½-LSB	1-LSB
8	7	6	6
10	8.3	8	7
12	10	9	8.3
14	11	10.4	10
16	12.5	12	11
18	14	13	12.5
20	15.2	14.5	14

As we did in the previous case we're reporting a table in where we show the oversizing ratio of the time constant τ of the circuit in order to keep a given accuracy.

It's easy to see that often the usual $4\div5$ time constant are not enough and it's required to wait e time in the order of $10\div15$ time constant. In the real case we would have to consider not only the MOS resistor R_{ON} but also the OpAmp output resistor. If we can estimate the OpAmp output resistor with a maximum value of 27Ω we have left still 300Ω for the MOS resistor R_{ON}. Last we have to consider the limiting factors of the OpAmp that feed the MOS. In the worst case we just analyzed we have to check that the OpAmp can provide the initial output current at the start of the charging exponential $I_{out,max}>FSR/R_{ON}=17mA$. The second thing to check is the *Slew-Rate* that has to be able to grant the maximum slope of the signal, that is $SR>2\pi\cdot f_{max}\cdot V_p$. Given an OpAmp with SR in the order of some $V/\mu s$, it's possible to get up to $63kHz$ as maximum frequency, which is higher than the requested one.

11.5.c

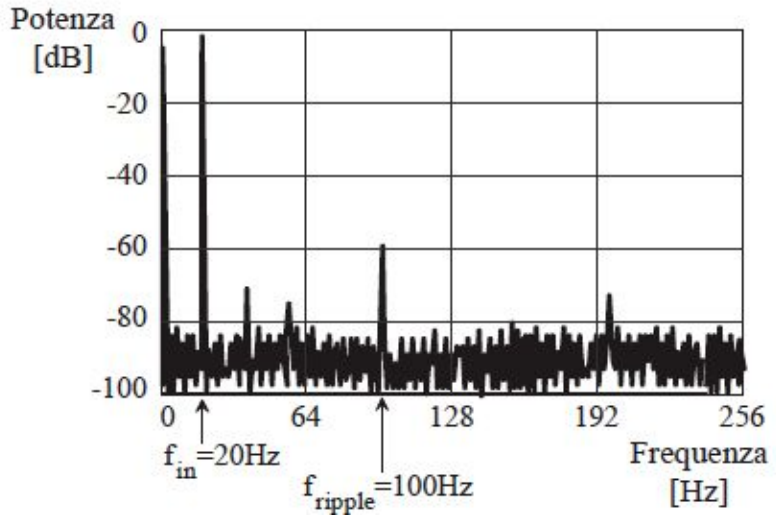
We can compute a numerical FFT with the numerical data at the output of the ADC. If With $N=4096$ samples we will get an accuracy in the frequency domain equal to the *bin-width* on f_s/N, in where f_s is the sampling frequency equal to $512SpS$, and so $f_s=512Hz$ that leads to a *bin-width*$=0.125Hz$. The use

of *4096* samples means we need to acquire the digital stream of samples at the output of the converter for a time $4096 \cdot 1/f_S = 8s$. The spectrum will be replicated at the multiples of f_S, thus we can draw just between *0* and $f_S/2$, like in the following picture. The graph consists of *N/2* histograms with width equal to the bin-width, which is *2048* channel of *125mHz*. Every histogram will have amplitude that represents the power of the signal in the corresponding frequency gap.

Just the quantization noise would lead to a *SNR* ideal of $6.02 \cdot n + 1.76 = 62dB$. Because of we're using *4096* samples for the FFT, the *noise-floor* in the picture will be lowered by a factor N/2, that is *33dB* in power. This is explained as we split the whole frequency spectrum $0 \div f_S/2$ in *N/2* histograms, each of this will get a portion *1/(N/2)* of the noise *–62dB* referred to the carrier. In the picture is easy to see that the *noise-floor* is at $-62 - 33 \approx -95dB$.

The y axis is normalized at *0dB*, equal to the power of a sinusoid with an effective power equal to the maximum dynamic, which is with a peek value of *FSR/2*:

$$P_{0dB} = \left(\frac{FSR/2}{\sqrt{2}} \right)^2 \quad \Rightarrow \quad 0 \quad [dB]$$



In our case *FSR=5V* is equal to a power of $(1.77V)^2 = 3.125V^2$. The DC signal at *FSR/2* has a power of $6.25V^2$. The sinusoid applied at the ADC gives a peek value of *2.2V* and so will give a power of $2.42V^2$, that is a factor

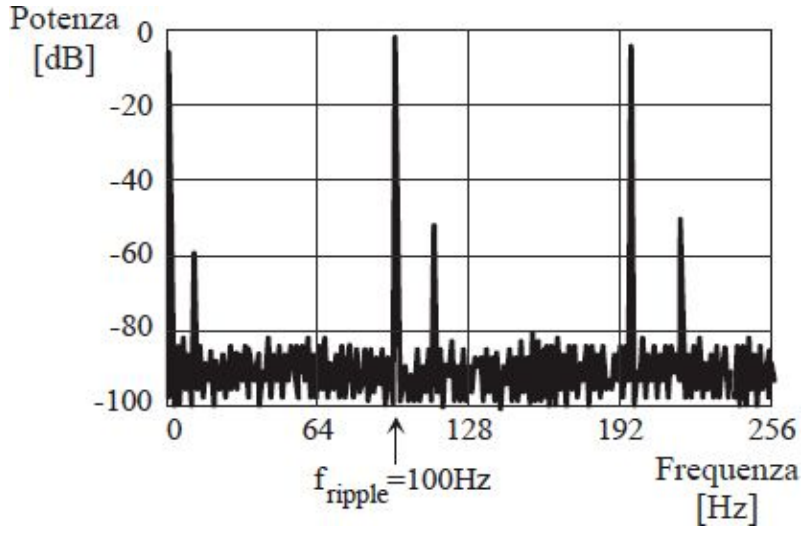
2.2/2.5=-1.1dB less than the maximum power, which is the *–1.1dB* we can see in the picture referred to the *0dB* value.

In the same graph we can find both the input signal at *20Hz* and its harmonic components, caused by the not linearity of the DAC transfer function. It's indeed present the CD component at *2.5V*. There's also the power sources ripple at *100Hz*: the spurious signal at *±1·LSB* that it creates at the output causes a ripple equal to:

$$P_{1\cdot LSB} = \left(\frac{FSR/2 \cdot 2^n}{\sqrt{2}} \right)^2 = \frac{P_{0dB}}{2^{2n}} \qquad \Rightarrow \qquad -6.02 \cdot n \qquad [dB]$$

That is the *–60dB* signal we can see in the previous picture.

If we remove the by-pass capacitor the power ripple would saturate the output of the OpAmp. The overall output would be this was more or less a square wave and the input signal would be completely hidden. In this case the spectrum would be the one in the following picture in which it's easy to see the spectrum of the square weave with the carrier at *100Hz* and all the harmonic components both at high frequency and their copy caused by the aliasing of the FFT.

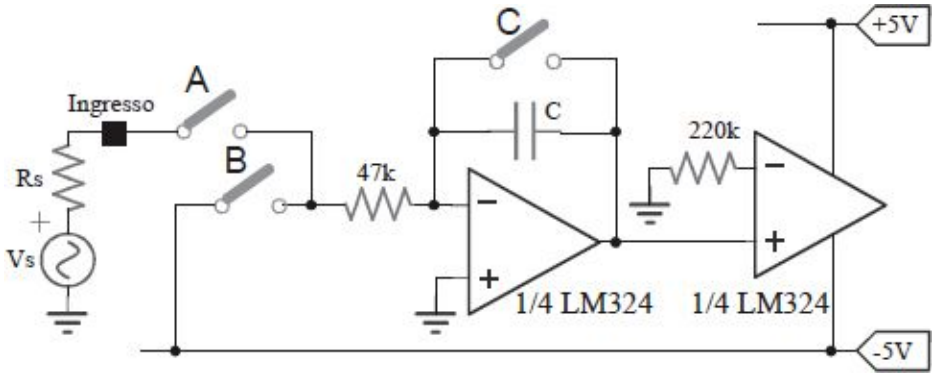


11.6

The circuit in the picture realize the analog part of a double ramp ADC with LM324 (*IB=500nA* e *VOmax=±3.5V*) that works at *100sps* and with *12bit* accuracy.

a) Compute the clock frequency and choose the value of *C*.

b) Discuss the influence of *IB* and the role of the *220k*Ω resistor.



11.6.a

The two OpAmps develops a double tramp ADS with the proper command at the switches from a control logic stage. Both the OpAmps are powered at *±5V*, but their output is limited by an *output voltage swing* from V_{Omax}=*±3.5V*. The first OpAmp works as an integrator stage and it converts at the end the voltage applied to a current that charges the feedback capacitor. The second OpAmp at the opposite works open loop, thus as a ccomparator.

At the beginning of each acquisition period the control logic has to close the C switch in order to fully discharge the capacitor, before to start the acquisition of a new sample. This way the voltage at the output of the first OpAmp gets to *0V*. After the discharge of the capacitor the C switch opens and the proper sampling operations can start.

After that the A switch close and the input signal integration phase starts. Given a source resistor negligible compared to the *47kΩ*, the charging current

results to be $v_s/47k\Omega$. The voltage at the output of the OpAmp starts to become negative with the usual expression:

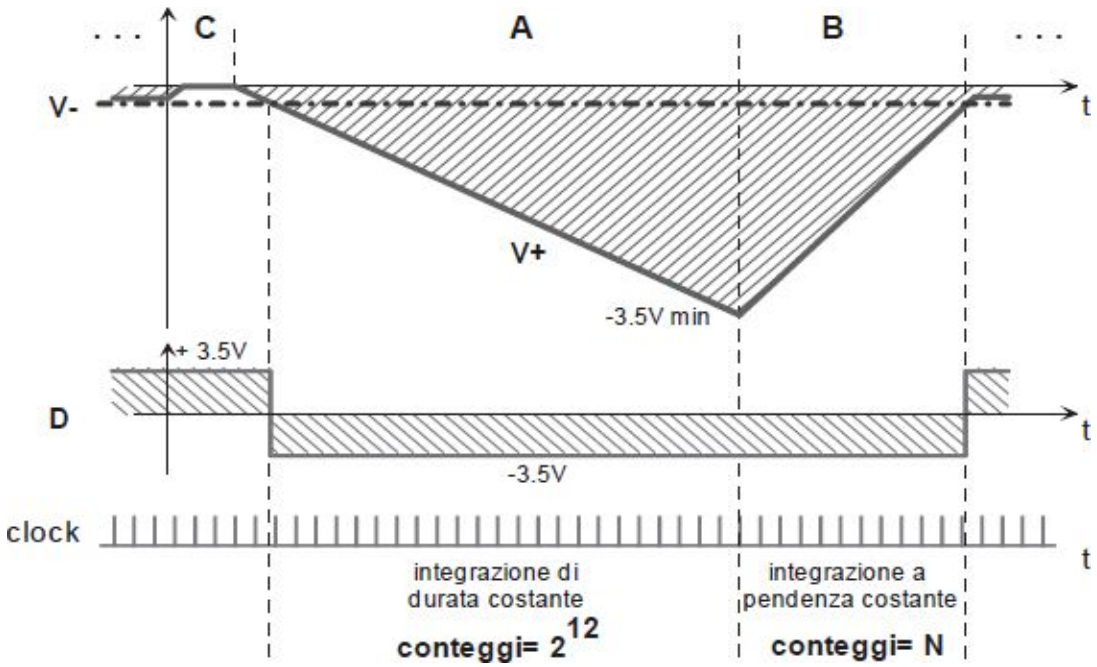
$$I_C = C \cdot \frac{dV_c}{dt} \qquad \Rightarrow \qquad V_{o1}(t) = -\frac{1}{47k \cdot C} \cdot \int_0^t v_s(\tau) d\tau$$

If v_s is constant the voltage will decrease as a negative ramp.

The control logic has to maintain this phase up to a constant. In the real case a counter has to be used that, following a clock with its own clock frequency f_{ck}, after the initial reset will keep on counting up to the overflow. The number of the bit of the counter has to be the same as the resolution of the ADC, which is *12bit*.

At the end of the integration phase, that is after a time equal to $2^{12} \cdot 1/f_{ck}$, the control logic has to open the A switch, to close the B switch and to reset the counter. The change of the switches leads to the discharge (from a negative value to *0V*) of the capacitor through the constant current $-5V/47k\Omega = -106\mu A$, and last until the voltage at the capacitor gets to *0V*.

This condition will be seen by the second OpAmp, used as a comparator, which will commute the output. During the whole first phase in which the A switch is closed the not inverting pin will be at a negative value thus the exit is saturated at *-3.5V*. At the end of the second phase, with the B switch closed, as soon as the voltage at the not inverting pin becomes a little positive the comparator commutes to the positive saturation value *+3.5V*, reporting to the control logic that the B switch has to get opened and the conversion is ended. The value in the *12bit* counter is the time, in clock cycles, between the closure and the opening of the B switches and it represent the digital value of the analog input. As examples we can see that if the input v_s has the maximum value *+5V* (equal to the V_{ref}=*-5V* at which the B switch is connected) the second phase of integration with B closed would last the same time as the first when A was closed and then the counter would be exactly at 2^{12}. At the opposite, given v_s=*0* the second phase would last the minimum time, that is only 1 clock cycle, in order to find out that the input is already at *0V*; and the result of the conversion would be exactly 2^0. Every other value of v_s will give a count between those two values.



The timing of the various operations is scheduled as in the picture below, together with the ramp succession and the commutation of the comparator D. In order to guarantee that this last snaps correctly, bear in mind that in the reset phase (C closed) the not inverting pin is at *0V*, it's required that the inverting pin (marked as *V-* in the picture below) is slightly below the 0V value in order to keep the D output at the upper saturation value *+3.5V*. As soon as the A switch close we just have to wait that the comparator to snap again before to start the counting the clock impulses. This way it's avoided the not linear start of the integration caused by the closure of the switch that isn't instant and a few other parasitisms. The maximum length of the conversion is achieved for $v_s=|V_{ref}|=5V$ and it's equal to the double of $2^{12}\cdot 1/f_{ck}$. In order to grant *100sps*, it's required a maximum conversion time of *1/100=10ms*. From those calculations it's easy to compute the minimum clock frequency, that has to be inferior to $10m/2/2^{12}=1.2\mu s$, thus $f_{ck}>820kHz$.

The capacitor has to be chosen in order to integrate the maximum signal $v_s=+5V$ all the second phase long that is *5ms*. Because of the maximum excursion at the output of the first OpAmp is the saturated value of *–3.5V*, it's required to limit the capacitor $C<I_{max}\cdot\Delta t/\Delta V=5/47K\cdot 5m/3.5=152nF$. It's important to get a real capacitor the closest possible to that value (e.g. *150nF*), because of too low values would reduce too much the output swing and, at last,

the accuracy. Of the conversion due to effects we just skipped so far like noise, parasitic currents, tolerances, offsets and so on.

11.6.b

First OpAmp Bias current has a double effect. When all the switches are open it causes a progressive discharge in the capacitor C that may causes the saturation of the first OpAmp and thus the saturation of the comparator. That is why it's always required a full reset of C before every conversion with the closure of C. If we're working free-running, the phase with all the open switches can last no more than *5ms* (if $v_s=0$); in this worst case the capacitor would charge at $dV=I_B \cdot dt/C=16.6mV$, that is not enough to saturate the OpAmps, but surely more than enough to bring some huge errors without the reset phase.

When the C switch is closed the current I_B gives no errors. At the opposite when it's A or B to be closed it forces an integration over C with a slope of $I_B/C=500n/150n=3.3V/s$. This voltage increase or decrease depending on the sign of I_B, and will show at the output of the first OpAmp, thus would be able to snap the comparator if not taken in proper consideration, or at the opposite bring it further to the switching threshold. For this reason and for the one just studied it's mandatory to place correctly the switching threshold of the second OpAmp and thus decrease the effect of I_B.

In regard of this problem it's possible to take advantage from the very same bias current at the inverting pin of the second OpAmp. If we connect a proper resistor to the inverting pin this causes a voltage drop in the reason of the connected resistor. With a $220k\Omega$ we get a threshold of $V\text{-}=I_B \cdot 220k\Omega=110mV$. Be careful with the sign! In order to have the whole working properly it's mandatory for the voltage drop to be negative (take a look at the picture). This leads to a current sink by the OpAmps. That's not the case in the LM324, which has a *pnp* differential stage at the input. This leads to just two choices: changing the OpAmps or (better) the way the negative reference voltage is achieved (and possibly more stable).

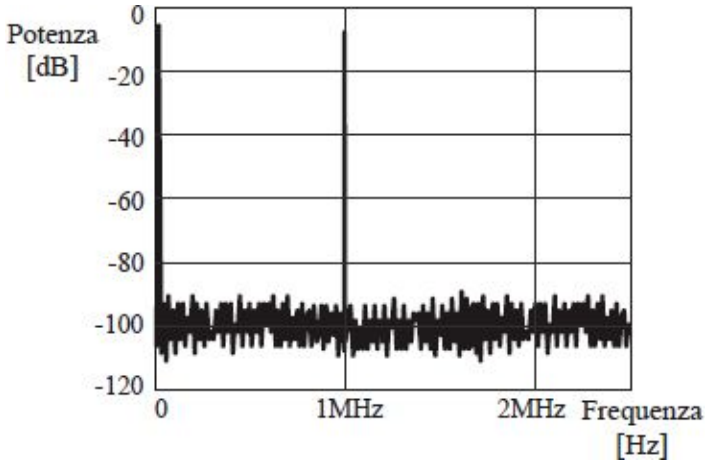
11.7

Given a *14bit* ADC with *FSR=5V, fS=5MHz*. The input is a sinusoid with *1V* peek amplitude. The FFT at the output is computed using *8192* samples (*bin-width=610.4Hz*).

a) Find out the *SNR* **ideal** (from the only quantization error), the *SFDR* and the dB value referred to the carrier.

b) Compute the **real** noise power.

c) Compute the *ENOB*, and comment the reason for which it's not *14bit*.



11.7.a

We start computing the reason of a *least significant bit* that in a *14bit* converter with *FSR=5V* is equal to $LSB=5/2^{14}=305\mu V$. The quantization error, as usual always between $\pm\frac{1}{2}\cdot LSB$, has to be taken into consideration by itself and unrelated to the input signal, thus managed as it was a noise with power equal to:

$$\sigma_q^2 = \frac{LSB^2}{12} = (88\mu V)^2$$

This power has to be compared to the carrier one and, in the case of a sinusoid with maximum amplitude *FSR/2* is equal to:

$$S_{\max} = \left(\frac{FSR/2}{\sqrt{2}} \right)^2 = (1.77V)^2$$

In the idea case in which there's no real noise but only the quantization error this last can be considered as the only disturb other than the signal and thus the signal to noise ratio can be computed as:

$$SNR_{ideal,\max} = \frac{S_{\max}}{quantization\ Noise} = \frac{S_{\max}}{\sigma_q^2} = \frac{(1.77V)^2}{(88\mu V)^2} = 86dB$$

In where we preferred the decibel value and skipped the usual "*10·log()*" formula. Anyway, because of you students never know whether to use the multiplication factor of 10 (to be used in the case of power ratio) or 20 (to be used in the case of voltages or currents ratio) it's useful to compute the result in both ways:

$$\frac{S}{Noise} = 10 \cdot \log \frac{(1.77V)^2}{(88\mu V)^2} = 20 \cdot \log \frac{1.77V}{88\mu V} = 86dB$$

What we just found out for the $SNR_{ideal,max}$ is usually written in a more general way in this formula:

$$SNR_{ideal,\max} = \frac{S}{\sigma_q^2} = \frac{\left(\frac{FSR/2}{\sqrt{2}} \right)^2}{\frac{\left(FSR/2^n \right)^2}{12}} = 6.02 \cdot n + 1.76 \qquad [\text{dB}]$$

That in the case of a *14bit* ADC gives exactly the value of *6.02·14+1.76=86dB* we just computed.

Because of the real input signal has only *1V* peek and not the whole *2.5V* allowed by the circuit, the ideal noise to signal ratio is lower than the one we just computed. In our case in fact the signal power does not hit the maximum value of $S_{max}=(1.77V)^2$, but just:

$$S = \left(\frac{1V}{\sqrt{2}}\right)^2 = (0.7V)^2$$

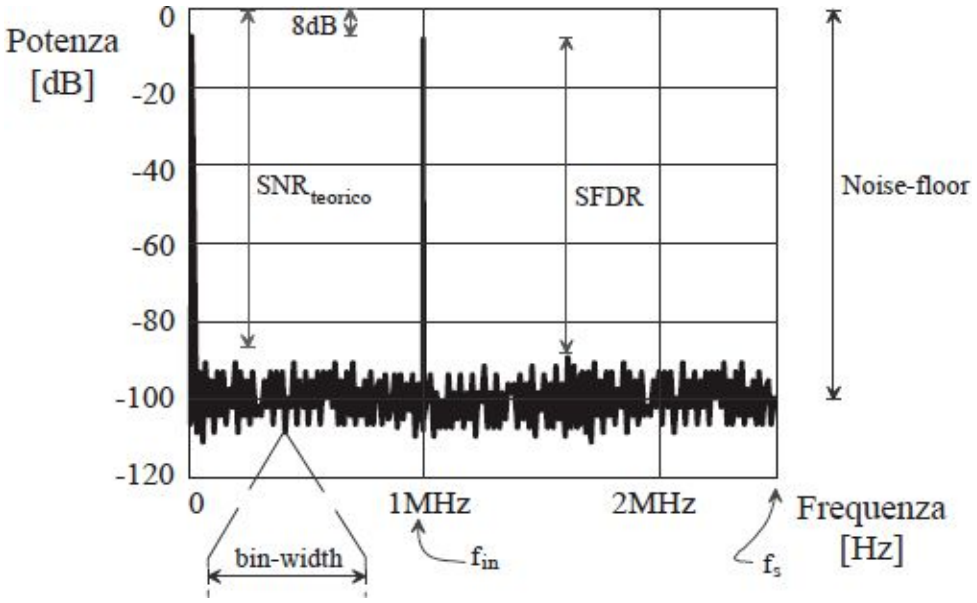
That leads in the case of the only quantization error to a *SNRideal* equal to:

$$\mathrm{SNR}_{\text{teorico}} = \frac{\mathrm{S}}{\text{quantization Noise}} = \frac{\mathrm{S}}{\sigma_q^2} = \frac{(0.7\mathrm{V})^2}{(88\mu\mathrm{V})^2} = 78\mathrm{dB}$$

In each spectral diagram is useful to draw at the y axis the power in *dB* referred at each histogram, which is the power of the entire signal in the range of frequency of the histogram itself. In the case of a sinusoid centered at a given frequency, *1MHz* in our case, the power in the histogram will be likely to be caused by the signal itself.

To help the comparison between ADC with different FSR it's a common use to normalize the value of the y axis in order to have at *0dB* the power of the sinusoid with the maximum peek value *FSR/2* that, in our case, it's equal to $S_{max}=(1.77V)^2$. In this way, the normalized output signal of the 1V peak sinusoid is under the maximum value allowed by a factor of 1/2.5 = 0.4, that is equal to 8dB (care to use the expression "*20·log()*"because it is a voltage ratio) below the origin, arbitrarily placed at 0dB. And this is why, in the following spectral diagram, the sine wave signal has peak equal to-8dB.

The component that appears at frequency *f = 0Hz* is the average value of the signal, that is the DC value overlapped to the sine wave to have it translate inside the dynamic allowed by the circuit, thus between *0V* and *5V*.



In addition to the sine wave signal and the DC component, there are no other signals, harmonics or noise. Thus, the spurious-free dynamic range, SFDR, defined as the ratio between the useful signal power and the highest spectral contribution (that from the picture can be estimated of more or less *–90dB*), is nearby *–8dB-(-90dB)=82dB*

11.7.b

The output of the ADC is represented by a digital data stream, clocked at the sampling frequency *fs = 5 MHz*, that is one every *1/5M = 200ns*. Taking a number of samples it's possible to compute an FFT on these numbers. The result of this math is exactly the spectral diagram shown above. The number of samples in the frequency domain is equal to the number of samples used in the time domain. Because of the sampling causes the spectrum to repeat periodically in the neighborhood of all integer multiples of *fs*, we can just confine the analysis in the range between *0Hz* and *fs / 2 = 2.5MHz*. In our case, we have used *N=8192* samples, the acquisition phase lasted *8192·200n≈1.6ms*, and in the range of interest there will be *N/2* samples in frequency called, simply, histograms. The width of each will be equal to *fs/N=610.4Hz*.

If we had taken a single sample in the time domain (N = 1), the FFT would have calculate only the power of that point, thus we would have had just a single histogram. Increasing the number of N, the number of histograms increases, and thus the frequency resolution. The power of the input signal,

assumed it's a sine weave at a given frequency (without jitter), is put in a single histogram, the one centered at 1MHz.

The quantization noise, however, is distributed over the frequency range. Then, as the number of histograms increases, the power of quantization error that falls into each of them will be gradually decrease. If this was a single histogram, and there was no useful signal, the power of this histogram would have a value of $SNR_{ideal,\ max}$ below the *0dB*. Instead, given a fixed number of samples, such power will be split into these histograms and the corresponding average level in each of them, called *noise-floor*, will be:

$$Noise - floor = -\left(6.02 \cdot n + 1.76 + 10 \cdot \log \frac{N}{2}\right) \qquad [\text{dB}]$$

In the theoretical case of complete absence of noise except the quantization error, the average would be *Noise-floor=-86-36=122dB*.

As a proof we can compute the overall noise power simply by multiplying the average one for the number of histograms:

$$Noise = (Noise - floor) \cdot \frac{N}{2} = -122dB \cdot 4096 = \ !!! \ = -122dB + 36dB = -86dB$$

The resulting number is equal to the level of quantization noise expected in theory, 86dB below the one of the maximum sine wave amplitude (i.e. the opposite of $SNR_{ideal,max}$). Note that while the noise-floor is a concept that has an immediate mean on the spectral plot shown in the previous page, the SNR, either theoretical or real, is not reflected directly on the graph. In fact, it does not represent any relationship between distance and height (power) of histograms, because the SNR is a concept of "integral" thus it represents the ratio between the power of a well-defined sinusoid (the signal, represented by the height of on the histogram) and the noise (which, however, is distributed among all histograms in the spectrum).

The above formula has highlighted an improper writing, but often used, that is the product of a value in *dB* (power) and a dimensionless factor: it is obviously incorrect to say *122dB·4096=500'000dB*, since each factor is measured in different ways.

In fact, measured from the spectrum (or, rather, calculated using an FFT) can be seen that the noise floor is established at around *-100dB*. This means that

there are additional contributions to the noise, leading to a degradation of *SNR*. The real power of noise is:

$$Noise = (Noise - floor) \cdot \frac{N}{2} = -100dB \cdot 4096 = !!! = -100dB + 36dB = -64dB$$

This value does not mean that the noise power is equal to $10^{-64/10}=(398nV)^2$, because the value in *dB* always refers to the *0dB* normalized power corresponding to $S_{max}=(1.77V)^2$. Thus, in the end, the total power of noise output of the ADC is equal to:

$$Noise = (Noise - floor) \cdot \frac{N}{2} \cdot S_{\max} = -64dB \cdot (1.77V)^2 = !!! = (1.1mV)^2$$

This value exceeds the level of quantization error of $(88\mu V)^2$ by a factor of *1.1m/88μ=12.5*, that is equal to *22dB*, equal to the ratio between the computed ideal *noise-floor* (*-122dB*) and the real one (*-100dB*) or, again, from the quantization noise (*-86dB*) and the total one (*-64dB*).

11.7.c

The *effective number of bit*, *ENOB* is defined as the number of bits of a converter that is equivalent to the present one in which the input signal is the same as the real case, but in where the noise is believed to be due solely to quantization theory. The accounts we have just seen, leads to a SNR_{real} equal to:

$$SNR_{real} = \frac{S}{Noise} = \frac{(0.7V)^2}{(1.1mV)^2} = 56dB$$

which could be described with a number of bits obtainable from the usual expression *SNR=2.6·n+1.76*, that is *9bit*. The general expression for the ENOB is:

$$ENOB = \frac{SNR_{reale} - 1.76dB}{6.02dB} = \frac{56dB - 1.76dB}{6.02dB} = 9$$

It’s interesting to see where those bits have been “lost”. First, it is not true that we have lost some “pieces” in the way (and neither that some outputs of

the ADC went on fire!). The bits that the ADC continues to provide are always *14*. Too bad that upon all of these, only *9bit* are useful, because some of the other *5bit* are never used (because of we failed at using the entire available output dynamic) and others are meaningless because of they “dance” with the noise superimposed to the signal. We will study in detail the causes of this loss.

Entering with a sine wave of 1V peak, instead of the maximum of 2.5V, we’re using only a small part of the available dynamic, in the mean of *1/2.5=40%*. This means that upon the $2^{14}=16'348$ codes available at the output, we’re using only *6553.6*. It’s like saying that it would take an ADC with only $log_2 6553.6=12.6bit$, instead of the 14 available; lost 1.4bit.

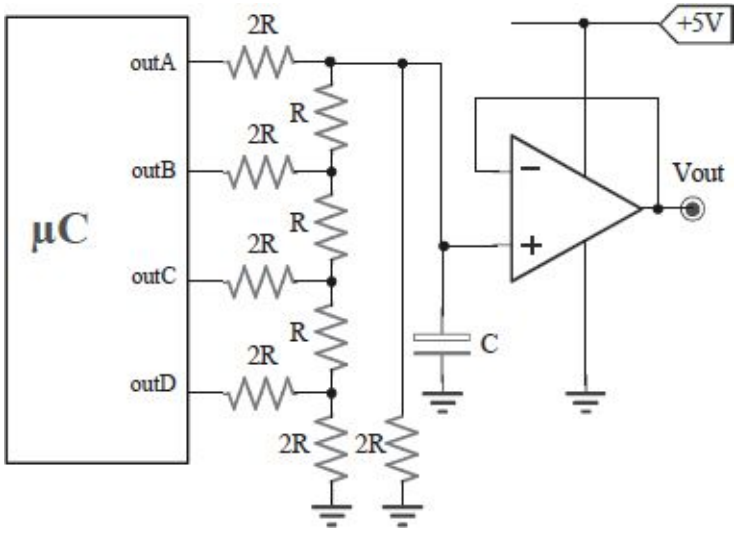
Also, instead of having just the quantization noise equal to $(88\mu V)^2$, we’re measuring a total of noise $(1.1mV)^2$. The deterioration by a factor of 22dB let us say that the *14bit* ADC is heavily noisy and, in practice, behaves as an ADC with only the theoretical quantization noise and a smaller number of bits, the mean of $LSB_{eq}^2/12=(1.1mV)^2$, thus *10.4bit* (obtained through $log_2(FSR/LSB_{eq})=log_2(5/3.8m)$). We lost pretty much this way$14-10.4=3.6bit$, because they are not useful at describing the signal, in that they follow the fluctuations of the noise. Last, as written before, we’re using wrongly $1.4+3.6=5bit$ of the *14bit* available.

11.8

Realize the R-2R *4bit* DAC, using a μC with $I_{out,max}$ = 1mA.

a) Rate the resistance and compute the tolerance to ensure an error of *±0.5LSB*.

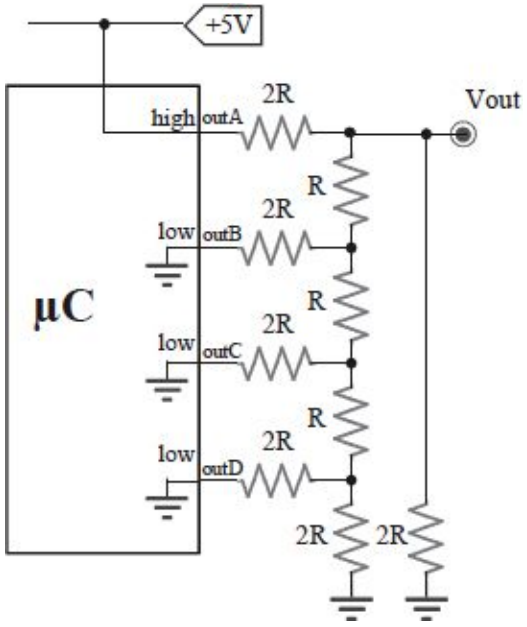
b) We want to convert *1'000 nibble/s*. Size C to smooth the output, ensuring the accuracy of 4bit.



11.8.a

First, it is of interest to verify that each output of the μC sees to the outside world, an equivalent resistance of value equal to 2R + R = 3R. Of all the releases, the most significant bit is outA while outD the least significant. In particular, if outA is at a *high* level, equal to 5V, the analog voltage at the output is *5·R/3R=5/3=1.67V*. If all the outputs were at a *high* level, the analog voltage output would be worth twice (actually slightly less in the mean of *1·LSB*) than this value, since the least significant bits, add one half of ... half of this value: the n^{th} bit of a DAC has weight *M* (where $M=2^{n-1}$), all *n*-bit encoding *2·M* levels, the value corresponding to all *n* bits at high level is *2·M-LSB*. At last, given all the outputs to *low* level, the voltage is equal to *0V*, and

when they are all at *high* voltage is 3.33V. In this case *1·LSB* is equal to $3.33/2^4 = 208mV$.



In order to deliver no more than *1mA* from each pin of the µC after the transients, we have to choose $R > 1/3 \cdot 5V/1mA = 1700\Omega$; that leads to a commercial $R = 3.3k\Omega$.

In order to do an analysis on the tolerance of the resistance, it is convenient to consider the most significant bit, outA, and assume that all resistances are unbalanced in the worst way, as shown in the figure. Because of the tolerance, *toll*, non-zero, the voltage at the output will differ from the theoretical value of *5V/3* by a value ΔV equal to:

$$\Delta V = \frac{5V}{3} - 5V \cdot \frac{R \cdot (1-toll)}{2R \cdot (1+toll) + R \cdot (1-toll)} = \frac{5V}{3} - 5V \cdot \frac{R \cdot (1-toll)}{3R + R \cdot toll} =$$

$$= \frac{5V}{3} - \frac{5V}{3} \cdot \frac{(1-toll)}{1+toll/3} \approx \frac{5V}{3} \cdot \left[1 - (1-toll) \cdot (1-toll/3)\right] \approx \frac{5V}{3} \cdot \frac{4}{3} \cdot toll$$

In which is supposed a tolerance much lower than *100%* (*toll<<1*) and has been ignored the term $toll^2$, as higher-order infinitesimal. This error will add to those of the other bits. However, having assumed the unbalanced resistance as shown above, it is impossible that the other bits give a contribution as much

high. In addition, their weight is gradually decreasing, the more we're getting closer to the least significant bit.

In conclusion, if we want to forces that the error is less than ±½·*LSB*, thus:

$$\Delta V < \frac{1}{2} \cdot LSB = \frac{1}{2} \cdot \frac{FSR}{2^4} = \frac{1}{2} \cdot 5V \cdot \frac{2}{3} \cdot \frac{1}{2^4}$$

We obtain *toll<0.046=4.6%*. This tolerance is certainly available on the market; however, it forces to resort to resistors at 1%, instead of the much cheaper resistance at 10% or 5%.

11.8.b

Convert *1'000nibble/s* means to have for each nibble (half byte of data, i.e. *4bit*) 1ms. The typical charge on the capacitor is exponential from the initial value to the asymptotic regime. In the worst case, the capacitor must charge from 0V to the maximum value *2/3·5V=3.33V*.

The error between the actual value during the charge and the regime value is exponentially decreasing, with the usual law:

$$\Delta V_c(t) = \frac{2}{3} \cdot 5V \cdot e^{-t/\tau}$$

In order to ensure at the end of the maximum time available *4bit* accuracy it's enough to ensure a charging time long enough when compared with the time constant of the charge. In particular, in order to ensure $\Delta V_c < ½ \cdot LSB = ½ \cdot 2/3 \cdot 5V/2^4$, it's required to use a constant charging time faster than:

$$\tau < \frac{t_{max}}{\ln(2 \cdot 2^4)}$$

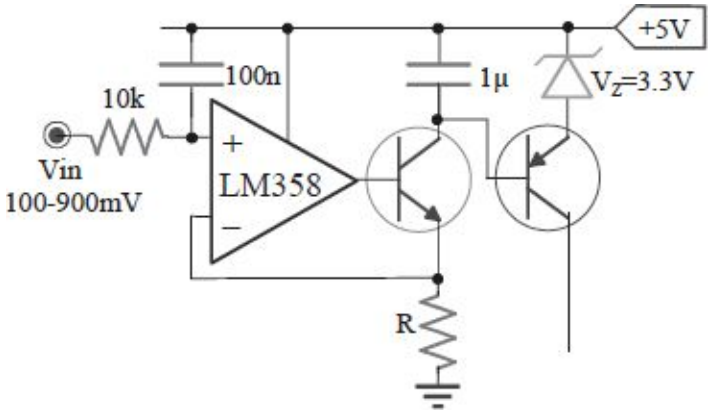
Given t_{max} equal to, it's required to size $\tau < 1ms/3.5 = 288\mu s$. Since the equivalent resistance seen from the capacitor is equal $2R//2R//2R = 2/3 \cdot R = 2.2k\Omega$, we can choose a capacitor of $288\mu/2.2k = 131nF$, we should choose a lower value, e.g. *100nF* or *47nF*, depending on the operations that the downstream electronics will be called upon to perform.

11.9

An ADC is realized using the circuit shown in the picture. The input changes from *+100mV* and *+900mV*.

a) Project the network that short-circuits the *1µF* capacitor for *1µs* when the pnp turns on.

b) Insert a $S_{TART}O_FC_{ONVERSION}$ input and complete the ADC in order to have at the output an *8bit* output.



11.9.a

The circuit achieves a kind of ramp ADC, where the slope of the ramp is determined by the input signal through the controlled current generator, while the threshold that determines the end of conversion is determined by network *BJT-ZENER*: As soon as the voltage to across the capacitor exceeds *3.3V+0.7V=4V*, the *BJT* turns on and causes the $E_{ND}O_FC_{ONVERSION}$. In order to discharge to zero (i.e. short circuit) the 1µF capacitor after the BJT turns on, we can exploit its collector current sourced to a resistor for switching on an N-MOSFET. This, in turn, turns on a PMOS placed in parallel to the capacitor by injecting current in resistor. The way it is, the circuit does not perform the intended function, however. In fact, as soon as the voltage across the capacitor drops below *4V*, the two *MOS* and *BJT* torn off, and the capacitor does not fully discharge. In order to avoid the problem it's enough to slow down the

shutdown of the *MOS* by overloading their gates with a capacitor. In fact, the charge of the capacity will be fast through the current injected by the previous, and the discharge will only happen through the load resistance.

11.9.b

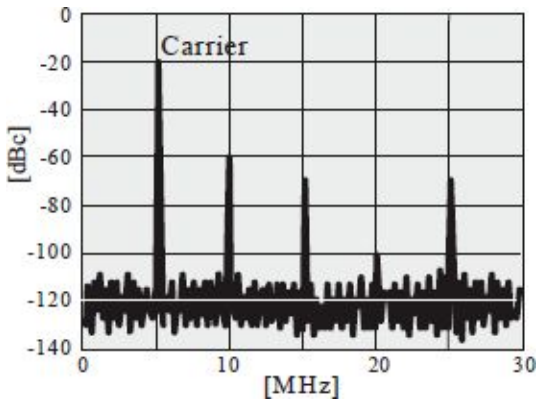
To introduce a $S_{TART}O_FC_{ONVERSION}$ signal in the previous circuit, it has to be slightly modified. In fact, the capacitor has to be kept short-circuited all the time from a signal $E_{ND}O_FC_{ONVERSION}$ to the subsequent $S_{TART}O_FC_{ONVERSION}$. It is easy to obtain this functionality by inserting a flip-flop between the gate of the *PMOS* and the two signals *EOC* and *SOC*. These two signals must also enable the clock of the counter 8-bit encoding that provides the output. Both the value of *R* and the period of the clock will depend on the time we are willing to spend for conversion.

11.10

Given the **unilateral** spectrum of a *16bit* ADC with *FSR=5V* and *bin-width* of *3kHz*.

a) Compute SNR_{ideal} (maximum amplitude and only quantization noise), theoretical noise-*floor* and rms value of the carrier (f_0=*5MHz*) and the second harmonic ($2 \cdot f_0$).

b) Compute $SNR_{theoretic}$ (real amplitude but only quantization noise), $SNR_{measured}$ (both real amplitude and noise) and the following *ENOB*.



11.10.a

The SNR_{ideal} is what you get considering only the quantization noise and the signal maximum amplitude. It is easy to derive the following useful relationship, knowing that the quantization noise has a power equal to $LSB^2/12$:

$SNR_{ideal} = 6.02n + 1.76 = 98dB$

The *noise-floor* is the theoretical base noise due only to quantization noise. As we are dealing with spectra obtained by FFT, what we see in the chart is not a power density, but rather the power present in band as wide as the *bin-width*. The spectrum also displayed only extends from frequency *0* to $f_s/2$. In fact it is periodical and for convenience only half period is reported. These considerations show that the theoretical *noise-floor* is equal to the total noise

power split between those bins that are in $f_s/2$. If you choose to calculate in *dBc* (*dB* related to the power of the carrier amplitude), the total power of noise - SNR_{ideal} is equal to:

$noisefloor_{teo} = -(6.02n + 1.76) - 10log(30MHz/3kHz) = -98 - 40 = -138dBc$

The root has a power of -20dBc, which is 20dB below the maximum amplitude of the carrier power. Since *FSR=5V*, the rms amplitude of the carrier is equal to $V_p/\sqrt{2}=2.5V/\sqrt{2}=1.76V$. The carrier therefore has an effective value of:

$$-20dBc = 20\log\left(\frac{V_1}{1.76V}\right) \rightarrow V_1 = 10^{\frac{-20}{20}} \cdot 1.76V = 176mV_{eff}$$

Similarly, the second harmonic has a power of-60dBc, which is 60dB below the maximum amplitude of the carrier power. Thus the second harmonic has an effective value of:

$$-60\text{dBc} = 20\log\left(\frac{V_2}{1.76V}\right) \rightarrow V_2 = 10^{\frac{-60}{20}} \cdot 1.76\text{V} = 1.76\mu V_{\text{eff}}$$

Note that the calculation was carried out directly on the effective value, making sure to put “20” in front of the logarithm. The result would have not changed by using the powers provided to properly calculate in dB.

11.10.b

The $SNR_{theoretic}$ is what you get considering only the quantization noise, but the real extent of the input signal. It will therefore be worse than the SNR_{ideal}, degraded by a factor $10log(P_{FSR}/P_{sig})$. In this case, the degradation is immediately readable from the spectrum, since the fundamental has a capacity of only *-20dBc*. Thus $SNR_{theoretic} = SNR_{ideal}$*-20dBc=78dB.*

The $SNR_{measured}$ is actually measured, and includes quantization noise, additional noise, and signal amplitude lower than the dynamic. It is still less than $SNR_{theoretic}$. To calculate it should make the real noise from the spectrum. It is nothing but the noise-floor visible spectrum multiplied by the number of bins. Then:

$$SNR_{misurato} = Segnale_{mis} - Rumore_{mis} = -20dBc - \left(-120dBc + 10\log\frac{30MHz}{3kHz}\right) =$$

$$= -20dBc - (-120dBc + 40dB) = 60dB$$

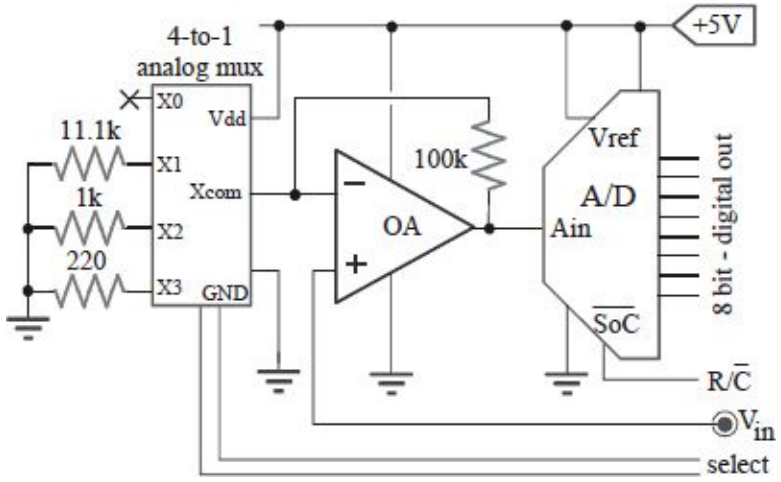
To this *SNR* corresponds an *ENOB* equal to *(60-1.76)/6.02=9.67bit*.

11.11

The circuit uses an OpAmp (A_0=*100dB, GBWP=10MHz*) and an analog mux (R_{on}≤*100Ω*, R_{off}≥*100kΩ*).

a) Compute the worst time of acquisition needed from when the gain is set until the conversion can start.

b) With an **ideal mux**, modify the circuit to have gains of 1, 2, 4, 8 and by halving the gain automatically as soon as the output of the ADC exceed half of full scale dynamic.



11.11.a

The proposed circuit allows you to vary the gain of the amplifying stage that precedes the ADC operating on the AnalogMux. At four different configurations corresponds four different gains with each own different band of the stage. We must then wait for an acquisition time longer or shorter depending on the configuration selected. In particular, the worst time of acquisition has the highest gain. Considering the ideal Mux (R_{on}=*0*, R_{off}=∞), the gains of the four configurations are:

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{00} = 1 + \frac{100k\Omega}{\infty} = 1 \qquad \left.\frac{V_{OUT}}{V_{IN}}\right|_{10} = 1 + \frac{100k\Omega}{1k\Omega} = 101$$

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{01} = 1 + \frac{100k\Omega}{11.1k\Omega} = 10 \qquad \left.\frac{V_{OUT}}{V_{IN}}\right|_{11} = 1 + \frac{100k\Omega}{220\Omega} = 455$$

To be honest the real gains will differ from those just computed due to the AnalogMux input resistances.

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{00} = 1 + \frac{100k\Omega}{(11.1k\Omega + R_{off}) // (1k\Omega + R_{off}) // (220\Omega + R_{off})} \cong 4$$

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{01} = 1 + \frac{100k\Omega}{(11.1k\Omega + R_{on}) // (1k\Omega + R_{off}) // (220\Omega + R_{off})} \cong 12$$

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{10} = 1 + \frac{100k\Omega}{(11.1k\Omega + R_{off}) // (1k\Omega + R_{on}) // (220\Omega + R_{off})} \cong 94$$

$$\left.\frac{V_{OUT}}{V_{IN}}\right|_{11} = 1 + \frac{100k\Omega}{(11.1k\Omega + R_{off}) // (1k\Omega + R_{off}) // (220\Omega + R_{on})} \cong 313$$

The lowest band will correspond to the configuration with the highest gain, which is 11. Because of in this case $1/\beta = 313$, $f_{pole} = GBWP \cdot \beta = 10MHz/313 = 32kHz$.

If you're not limited by the operational *slew-rate*, the minimum acquisition time to wait to make sure you get the correct signal is:

$$T_{acq} \geq \tau \cdot \ln \frac{FSR}{\varepsilon}$$

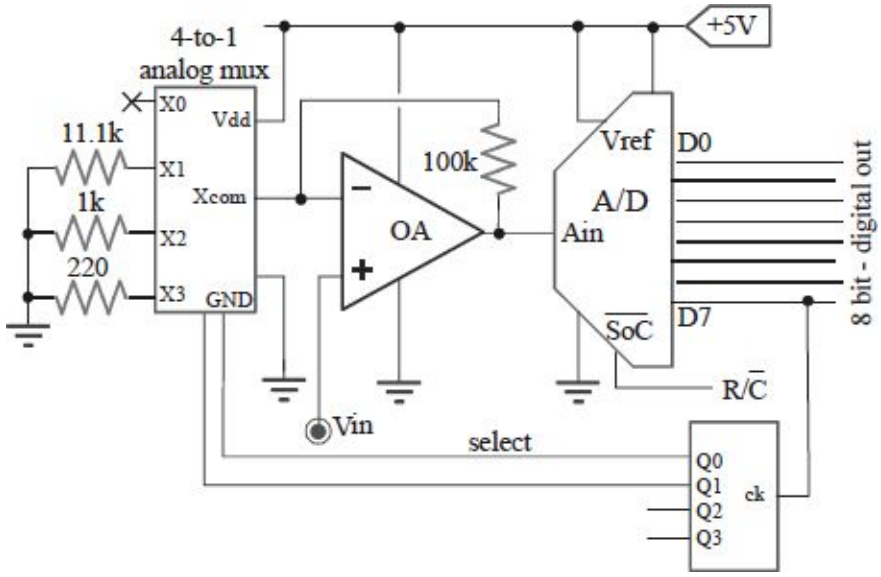
Given a maximum error equal to 1/2LSB we can compute:

$$T_{acq} \geq \tau \cdot \ln 2^{n+1} \cong 5\mu s \cdot 6.2 = 31\mu s$$

11.11.b

To ensure that every time is exceeded half the dynamic gain is halved, it is sufficient to take the last bit of the ADC cycle and use it to change the gain by a binary counter, as shown in the figure. It is clear however that this gain can only be decremented and not incremented. In addition, once you reach the minimum gain, exceeding half a subsequent dynamic causes a jump to the

maximum gain. For exercise it’s proposed to improve the circuit to solve these two problems.

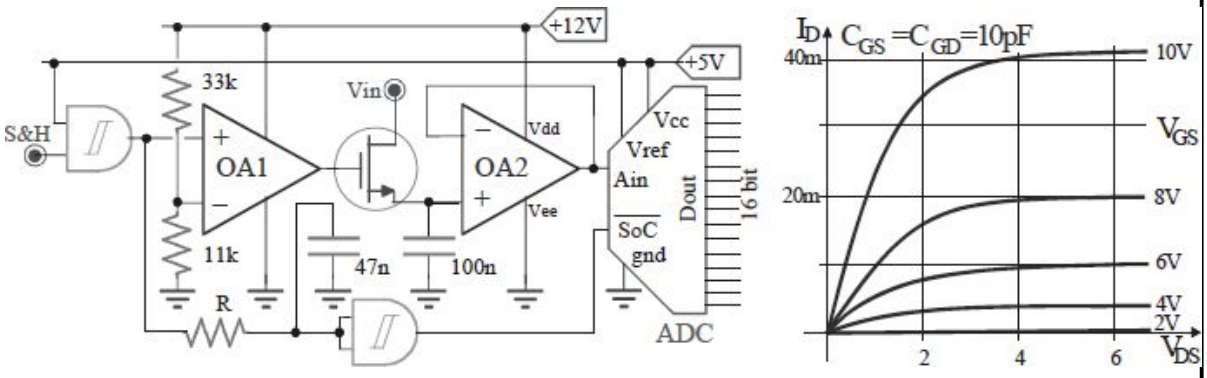


11.12

S&H command and NAND have $t_{rise}=t_{fall}=t_{delay}=20ns$. OpAmps have an *output voltage swing* between V_{ee} and $(V_{dd}-2V)$ and $SR=5V/\mu s$. MOS has $V_t=2V$. The given ADC has *16bit* accuracy. The input has $f_{max}=2kHz$.

a) Compute $V_{in,max}$ in order to guarantee an acquisition time $<200\mu s$ thus resize R.

b) Compute the highest error in *LSB* caused by *1V* ripple on the +12V power supplies.



11.12.a

The RC network formed by the hold capacitor and the MOS R_{on} leads to an exponential charge of the hold capacitor. The longer the acquisition time used, the lower mismatch of the voltage. In addition, the error is also a function of the maximum input signal applied. The mismatch, in the worst case is:

$$\Delta V = V_{in,\max} \cdot e^{-T_{acq}/\tau}$$

When in the process of sampling, *10V* are applied to the gate of the MOS. So at best $V_{gs}=10V$. In reality, the voltage V_{gs}, and thus MOS R_{on}, depends on the hold capacitor voltage. If we study the worst limit, we can freely assume that the input signal will be smaller than *5V* regardless of the maximum error due to the limitation imposed by RC, since this is the dynamic of the ADC. Therefore, the V_{gs} is still greater than $10V-5V=5V$. At this voltage corresponds

R_{on}*=2V/10mA=200Ω* , thus *τ=200Ω ·100nF=20μs.* In order for the error to be less than *1/2LSB* and the acquisition time less than 200μs it's required:

$$V_{in,max} < \frac{1}{2}\text{LSB} \cdot e^{T_{acq}/\tau} = 38\mu V \cdot e^{\frac{200\mu s}{20\mu s}} = 837mV$$

Assuming to have an acquisition time of *200μs*, the *SoC* command has to come to the ADC *200μs* after the MOS has been closed. Apparently the *SoC* command may be the inverse of the *S&H* command. In fact, because of the slew rate of the comparator, the MOS fully turns on only *2μs* after the signal went up. We have to compensate for this delay, otherwise the ADC start conversion when the acquisition is not completed yet. Thus *R* has to introduce a delay of *2μs*. Assuming that the switching threshold of the AND gate is 3V, this requires a time constant *τ=2.1μs*, i.e. *R=45Ω*

11.12.b

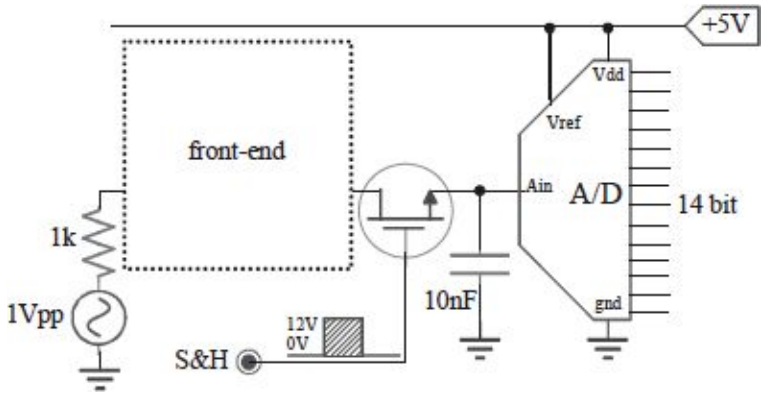
The power sources ripple causes an *aperture time jitter*, because it forces an oscillation of the comparator threshold of *1V/√2·11kΩ/(11kΩ+33kΩ)=176mV*. Since the control signal of the comparator has a speed of *5V/20ns*, the oscillation causes a jitter on the opening of the MOS equal to *176mV / 5V/20ns=0.7ns*. For a signal at *2kHz* and $5V_{peek\text{-}to\text{-}peak}$, which has a maximum slope of *2.5V·2π·2kHz=31.4mV/μs*, this corresponds to an error of *31.4mV/ μs·0.7ns=22μV*, equal to 0.3LSB.

11.13

MOS has $R_{on}<100\Omega$, $C_{GD}=C_{GS}=C_{DS}=2pF$ and $I_{leak}=10nA$. Input range between $\pm 500mV$ with phone bandwidth ($4kHz$).

a) Size the front-end stage in order to use the whole *FSR*.

b) Find out the error in *LSB caused by* charge-injection, leakage and R_{on} after the MOS has been kept open over $40\mu s$ and a sampling phase that has last maximum time allowed by the sampling theorem.



11.13.a

To exploit the whole dynamic of the ADC is necessary to amplify and translate the *1Vpp* input signal so that it has a range $0 \div 5V$. One way is to use an adder, in which we sum the signal amplified in a not-inverting stage for $1+R_2/R_1=1+470k\Omega/120k\Omega \cong 5$ (which then produces an output signal of amplitude $\pm 2.5V$), with a constant component of 500mV, which is also amplified by 5.

In order to avoid loading the divider with the DC component responsible for the continuous output of 2.5V, we have to decouple the DC input.

11.13.b

The LSB for the configuration under consideration is $LSB=5V/2^{14}=300\mu V$. The ADC input leakage during the hold time of 40µs results in a voltage droop

on the capacity of $\Delta V=I\cdot t/C=10nA\cdot 40\mu s/10nF=40\mu V$, much less than 1 LSB.

The presence of the MOS R_{on} leads to a pole with time constant $\tau=100\Omega\cdot 10nF=1\mu s$. The presence of this pole has a dual effect. It creates a low pass filter that even with infinite time of acquisition may result in an attenuation of the input signal. To compute this attenuation, we have to assess the transfer function of the filter to the signal frequency. That is: is:

$$\left|\frac{V_{in}}{V_{out}}\right| = \frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}}$$

thus:

$$\Delta V_{in} = V_{in} - V_{in}\cdot\frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}} \approx V_{in}\cdot\left[\frac{1}{2}\left(f_{in}/f_p\right)^2+\ldots\right] = 5V\cdot\left[\frac{1}{2}\left(4kHz/160kHz\right)^2+\ldots\right] = 1.5mV$$

Equal to 5 LSB.

In addition, the RC also determines an exponential charge of the hold capacitor. The mismatch made is lower the longer the acquisition time. In this case we are sampling at the Nyquist, i.e. every 11/8kHz=125μs. Since the hold phase lasts $40\mu s$, $80\mu s$ remain available for sampling. Therefore the mistake made is, in the worst case:

$$\Delta V = FSR\cdot e^{-T_{acq}/\tau} = 5V\cdot e^{-80} \approx 0$$

In the S&H MOS switching, due to C_{gs} capacitor, a small charge is injected in the Hold capacitor. In the worst case this contribution is equal to:

$$\Delta V = \Delta V_{gs}\cdot\frac{C_{gs}}{C_{gs}+C_H} = \left(V_{in\,max}+V_T\right)\cdot\frac{2pF}{2pF+10nF} = 600\mu V$$

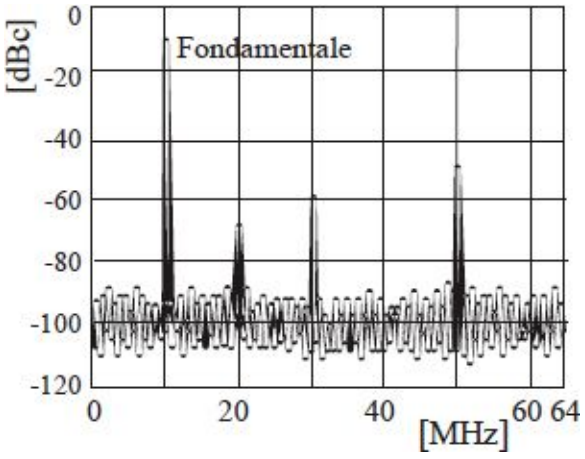
Equal to 8 LSB, given a threshold voltage equal to V_T=2V.

11.14

This unilateral spectrum (*16kHz bin-width*) correspond to a*16bit* ADC with *FSR=10V*.

a) Compute the theoretical *noise-floor*, SNR_{ideal} (maximum amplitude and only quantization noise) and the rms value of the 5^{th} harmonic.

b) Compute $SNR_{theoretic}$ (real amplitude only quantization noise) and $SNR_{measured}$ (real amplitude and noise) with the subsequent *ENOB*.



11.14.a

The SNR_{ideal} is what you get considering only the quantization noise and the signal maximum amplitude. It is easy to get the following relationship, knowing that the quantization noise has a power equal to $LSB^2/12$:

$SNR_{ideale} = 6.02n + 1.76 = 98dB$

The *noise-floor* is the theoretical base noise due only to quantization noise. As we are dealing with spectra obtained by FFT, what we see in the chart is not a power density, but rather the power that lies on a band as wide as the bin-width. The spectrum also displayed only extends from 0 to $f_s/2$. In fact it is periodical and for convenience we can study only half period. These considerations show that the theoretical *noise-floor* is equal to the total noise

power split between all the *bins* in $f_s/2$. If we choose to compute the whole in *dBc* (*dB* related to the power of the maximum amplitude carrier), the total noise power is equal to SNR_{ideal} changed sign. Then:

$noisefloor_{teo} = -(6.02n + 1.76) - 10log(64MHz/16kHz) = -98 - 36 = -134dBc$

The fifth harmonic has a power of-50dBc, which is 50dB below the power of the maximum amplitude carrier. *Since FSR=10V*, the rms of the fundamental is equal to $V_p/\sqrt{2}=5V/\sqrt{2}=3.53V$. The fifth harmonic has thus an effective value of:

$$-50dBc = 20\log\left(\frac{V_5}{3.53V}\right) \rightarrow V_5 = 10^{\frac{-50}{20}} \cdot 3.53V = 11mV_{eff}$$

Note that the calculation was carried out directly on the effective value, making sure to put "20" in front of the logarithm. The result would have not changed by using the powers provided to properly calculate in *dB*.

11.14.b

The $SNR_{theoretic}$ is what you get considering only the quantization noise, but the real extent of the input signal. It will therefore be worse than the SNR_{ideal}, degraded by a factor $10log(P_{FSR}/P_{sig})$. In this case, the degradation is immediately readable from the spectrum, since the fundamental has a power of only *-10dBc* (its power isn't the computed *3.53V*, but just $10^{-10/20}\cdot 3.53=1.11V$). Thus $SNR_{theoretic} = SNR_{ideale}-10dBc=88dB$.

The $SNR_{measured}$ is actually measured, and includes quantization noise, additional noise, and signal amplitude lower than the dynamic. It is still less than $SNR_{theoretic}$. To calculate it should make the real noise from the spectrum. It is nothing but the noise-floor visible spectrum multiplied by the number of bins. Then:

$$SNR_{misurato} = Segnale_{mis} - Rumore_{mis} = -10dBc - \left(-100dBc + 10\log\frac{64MHz}{16kHz}\right) =$$

$$= -10dBc - (-100dBc + 36dB) = 54dB$$

To this *SNR* correspond an *ENOB* equal to only *(54-1.76)/6.02=8.67bit*

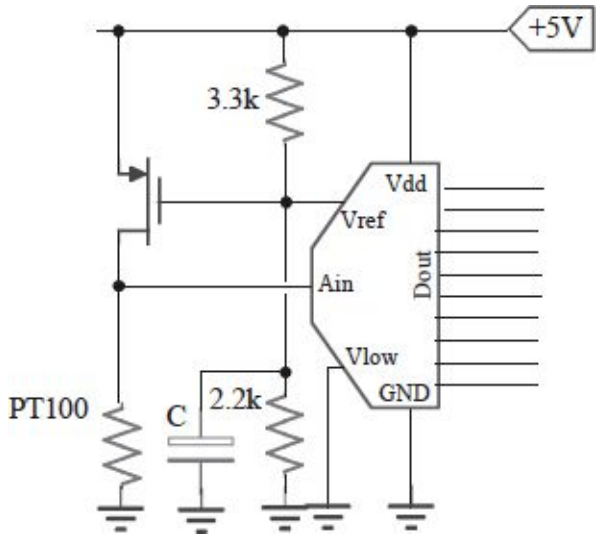
11.15

MOS has to provide *10mA*, given $|V_t|=1.2V$ and $½\mu C_{ox}=300\mu A/V^2$. PT100 (that is *100Ω* at *0°C* and increase by *0.5Ω/°C*) has to measure temperatures between *0°÷300°C*. ADC has *10bit*.

a) Size transistor *W/L* and the C capacitor in order for a *500mV* ripple at *50Hz* in the power supply not to change V_{ref} more than *0.5mV*).

b) Find out the hex code at the output provided at *0°C* and *100°C*; compute the *ENOB* for an input range *0°C÷100°C*.

c) Modify the circuit in order to avoid distortions and use all the *10bit* range, given $V_{ref}=5V$ e $V_{low}=0V$.



11.15.a

To size the MOS *W/L* we can simply use the drain current formula of MOS while in saturation zone. Because of $|V_{gs}|=5V\cdot 3.3k\Omega/(3.3\ k\Omega+2.2\ k\Omega)=3V$ we have that:

$$\frac{W}{L}=\frac{I_{D,sat}}{\frac{1}{2}\mu Cox(V_{gs}-V_t)^2}=\frac{10mA}{300\mu A/V^2\cdot(3V-1.2V)^2}=10$$

To choose the value of C we have to size the R-RC filter. Instead of computing the transfer of the filter in the frequency domain, we can exploit the concept of impedance of a capacitor, because we're interested to attenuate a single frequency (50 Hz). The ripple of *500mV* has not to vary V_{ref} more than *0.5mV*, the filter has to attenuate by a factor of 1000 the grid frequency. That is to say that, computing as in a resistive divider:

$$\frac{1}{sC} \parallel 2.2k\Omega \cong \frac{1}{1000} 3.3k\Omega = 3.3\Omega$$

Since it is obvious that the parallel will be dominated by C, we can write:

$$C \geq \frac{1}{2\pi f \cdot 3.3\Omega} \cong 1000\mu F$$

11.15.b

The PT100 changes from *100Ω* at *0°C* with a slope of *0.5Ω/°C*. Thus, at *0°C*, it is a *100Ω* resistor and the voltage drop at his pins is equal to *100Ω·10mA=1V,* while at *100°C* it is *150Ω* and the voltage drop is *150Ω·10mA=1.5V*. Because of V_{ref}*=2V* we have:

$$0°C \rightarrow Dout = 2^{10} \cdot \frac{1V}{2V} = 512 = 200Hex$$

$$100°C \rightarrow Dout = 2^{10} \cdot \frac{1.5V}{2V} = 768 = 300Hex$$

Note that temperature above 200 ° C are out the dynamics of the ADC, producing a voltage> 2V. They will therefore not be recognized, producing in all cases the output code *1023*.Trascrizione fonetica

To compute the ENOB, in case of transition 0 to 100 ° C is sufficient to note that it uses only a quarter of the total dynamic (0.5V to 2V, 256 encodings over 1024), so the actual number of bits is reduced to *10-*log_2*2=8bit*.

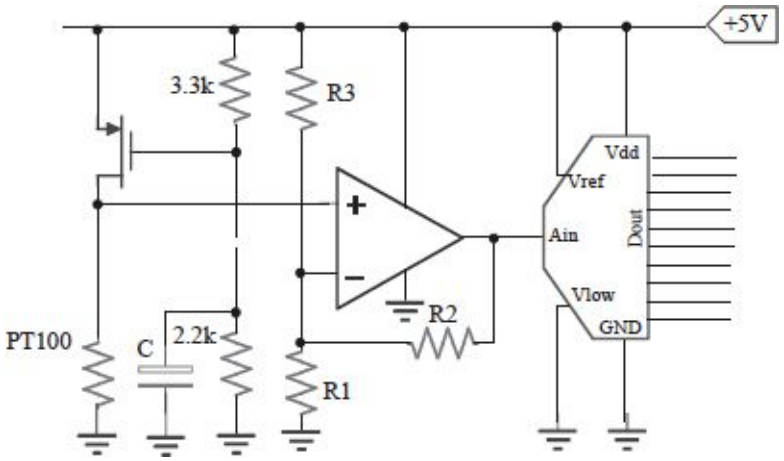
11.15.c

To avoid distortion and structure the whole dynamic of the ADC, with V_{ref}*=5V* and V_{low}*=0V*, it's required that the input temperature 0°C input to the ADC corresponds to 0V and the temperature 300°C corresponds to 5V.

It should then enter a stage amplifying between ADC and the heat resistor leading the voltage range for the 300°C from *1.5V* to *5V*, and at the same time remove the component due to polarization of heat resistance, equal to 1V.

To do this we can use the operational stage in the figure. It acts as an adder of two signals, the voltage signal from the RTD which is amplified by a +3.3

factor, and a current injected into the virtual ground in charge of a component at the output equal to *-1V*. To size the stage should be noted that the resistor R_3 is in parallel with R_1 in the computation of the non-inverting transfer, then the value of the latter must be adjusted with respect to the classic non-inverting.

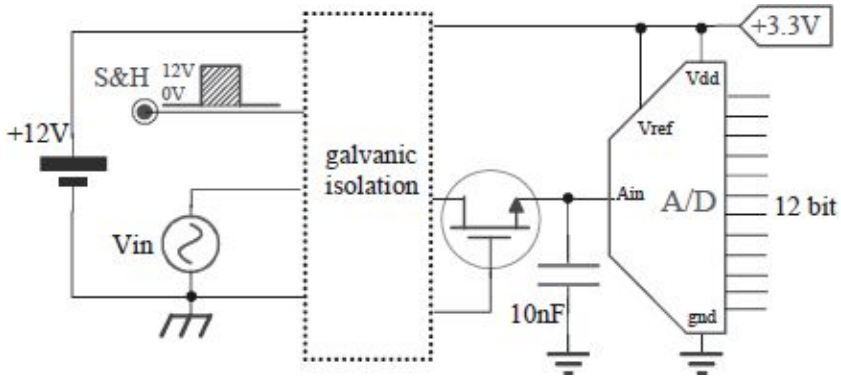


11.16

A 12V world has to be measured using a 3.3V electronic, remaining electrically isolated from it. The input is a *±500mV* sinusoid at audio band.

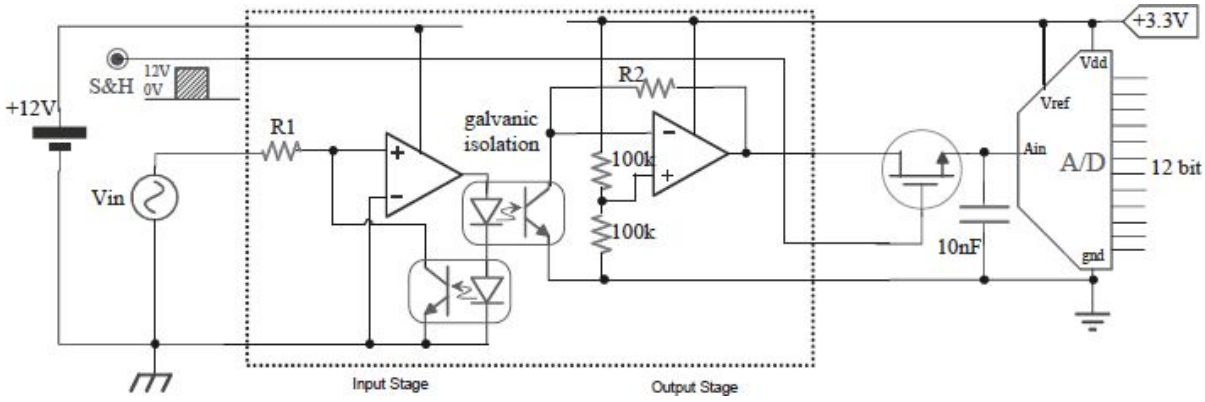
a) **Without** using integrated ISO, just normal OpAmp and **optocouplers**, design the front-end block that forces the input signal to tap the full FSR.

b) Replace the 12-bit ADC with a pair of ADC (one *8bit* and the other by *4bit*) and the further electronics if required.



11.16.a

If we want to ensure isolation between input and output, we have to create an amplifying stage that uses optocouplers as in the following picture:



In essence, the coupled pair of optocouplers, driven by the same LED, produces a transfer of power from the input stage to the output. The feedback is negative, then it tends to keep the non-inverting terminal of the first OpAmp in the potential of the second, i.e. ground. Therefore, the current entering the input node is equal to Vin/R_1 and can only flow in photodiode. That is because of the reaction, which requires the proper current in the LED. Since the photodiodes are coupled, the same photocurrent flow in R_2, resulting in an output voltage equal to $V_{in} \cdot R_2/R_1$. To ensure that the output has a $0 \div 3.3V$ instead of $-1.65V \div +1.65V$ is enough to put the positive pin of the output OpAmp at $+1.65V$.

Unfortunately, it is to note that this stage does not work for negative input signals. The Reader will have to remedy this problem.

In addition, the S&H input command is not galvanically isolated from the MOS transistor. To do so would be enough to put an optocoupler in series with the LED driven by the S&H command and the phototransistor which brings the MOS gate to power up the right side, i.e. 3.3V. Sadly this power is not enough to "close" well the MOS switch. Even here, therefore, the Reader will have to find the appropriate solution.

11.16.b

In order to obtain a *12bit* accuracy resolution from two converters with only *8bit* and *4bit*, we have to make a *"raw"* conversion of the signal, such as *8bit*, and then a *"thin"* conversion on the mismatch made during the raw

conversion, using the other converter. The process is similar to the principle of operation of a SAR converter.

The circuit must then:

- Convert the signal to *8bit*;
- Convert the captured data in analog form;
- Make the difference between the original signal and reconstructed signal;
- Convert the difference in *4bit*, using *1·LSB* of the 8-bit converter as Full-Scale Range.
- *12bit*-scrambling: the less significant bits will be those supplied by the *4bit* “thin” converter.

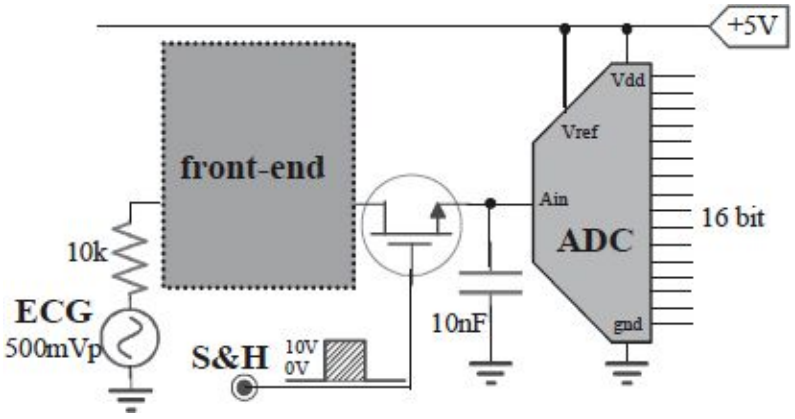
The following figure, if there were, would show a basic diagram of the circuit. Since there is not, we leave the Reader eager and artistically ready to plot it.

11.17

MOS has $R_{off}>10M\Omega$, $C_{GD}=C_{GS}=C_{DS}=3pF$ and $I_{leakage}=1nA$. Input is an ECG signal with $\pm 500mV$ amplitude and $0.2Hz$-$200Hz$ bandwidth.

a) Project the front-end stage in order to use the whole *FSR*.

b) Find out the error in LSB due to charge-injection, leakage and R_{off} after the MOS switch to be kept open over $2ms$ WITHOUT the front-end stage and with the signal short-circuited.



11.17.a

To exploit the whole dynamic of the ADC is necessary to amplify and translate the $1Vpp$ input signal so that it has a range $0 \div 5V$. One way is to use an adder, in which we sum the signal amplified in a not-inverting stage for $1+R_2/R_1=1+470k\Omega/120k\Omega \cong 5$ (which then produces an output signal of amplitude $\pm 2.5V$), with a constant component of 500mV, which is also amplified by 5.

We have to introduce a decouple capacitor in order to avoid loading the divider with the DC component responsible for the continuous output of 2.5V.

11.17.b

The LSB for the configuration under consideration is $LSB=5V/2^{16}=75\mu V$.

The leakage of the ADC during a hold time of *2ms* results in a voltage droop on the capacitor of $\Delta V = I \cdot t / C = 1nA \cdot 2ms / 10nF = 200\mu V$, i.e. *2.7LSB*.

The presence of R_{off} causes during the hold phase the discharge of the capacitor through the series of R_S and R_{off}. In the worst case this leads to an error of:

$$\Delta V = V_P \cdot (1 - e^{-T_{off} / C(R_{off} + R_S)}) = 500mV \cdot 0.02 = 10mV$$

That is 133LSB (!!).

During the S&H MOS switching, through to capacity C_{gs} a small charge is injected in the hold capacitor. The complete network should consider all three capacities C_{GD}, C_{GS} and C_{DS}. To simplify we can consider only the C_{GD} to the C_H. The transition at the gate, which causes the charge injection [the one that happens at MOSFET open, since then $V_G = V_{inmax} + V_T$, where V_T is the MOSFET threshold voltage (e.g., $V_T = 2V$) to the minimum value of V_G (in this case equal to *0V*). In conclusion, we have:

$$\Delta V = (V_{in\,max} + V_T) \cdot \frac{C_{gs}}{C_{gs} + C_H} = 2V \cdot \frac{3pF}{3pF + 10nF} = 600\mu V$$

Equal to 40 LSB.

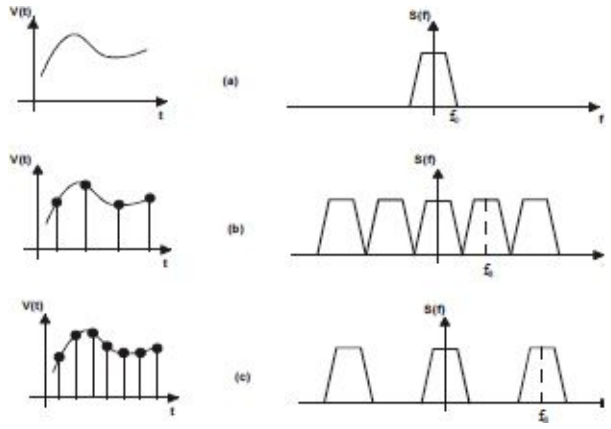
11.18

Oversample with an ADC to process an ECG signal (*0.2Hz÷250Hz*).

a) Find out the maximum OS reached by a SAR (T_{bit}=*2μs* e *10bit*).

b) Compute the maximum bandwidth of the input signal if we want to reach a *16bit* equivalent accuracy with that SAR.

c) Compute the clock frequency to provide a single bit ∑Δ converter in order to get up to *16bit* accuracy.



11.18.a

The SAR converter uses *2μs* for each bit of resolution. The total conversion time is then *20μs*, i.e. the time of sampling, and the maximum sampling frequency is 50kHz. Since the input signal has a bandwidth of 250Hz, the maximum oversampling is equal to *OS=50kHz/500Hz=100.*

11.18.b

Using oversampling, we can demonstrate that there is an improvement of SNR due only to quantization noise equal to $SNR_{out}=SNR_{teo}\cdot\sqrt{OS}$. The SNR_{teo} for a *10bit* converter is equal to *62dB*, while a *16bit* converter is equal to *98dB*, so the oversampling provides *36dB*. This corresponds to apply for a minimum of oversampling OS = 4096 (in fact every quadrupling of the sampling rate corresponds to an increase in the resolution of 1 equivalent bit). Since the drive

works at a maximum of *50kHz*, the maximum Nyquist frequency will be *50kHz/4096=12Hz*, then the maximum bandwidth of the input signal will be *6Hz*.

11.18.c

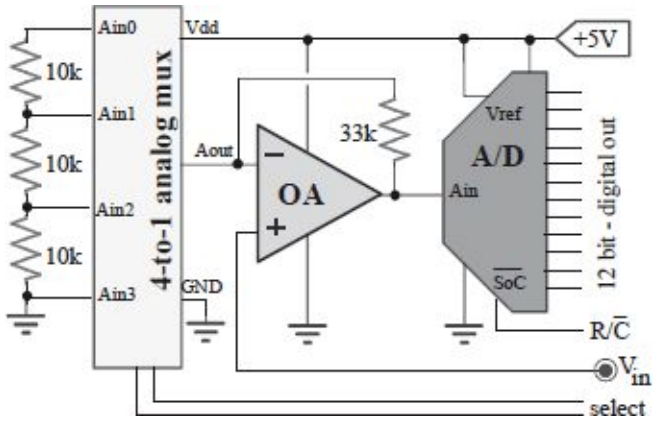
In the case of ∑Δ converter we can write the following relation: $SNR_{dB} = 6.02(c + 1.5L) - 3.41$, where c is the number of bits of the converter and 2^L the oversampling factor. In essence, we gain 1.5bit at each doubling of sampling frequency. Since we have to increase by 15bit, we have to perform 10 doublings, i.e. $OS=2^{10}=1024$. On the other way we can simply impose $c+1.5L=16$. Thus, the ECG signal has to be sampled at $f_{\Sigma\Delta}=250Hz \cdot 2 \cdot 1024=512kHz$.

11.19

The conversion circuit uses an OA (*rail-to-rail*, A_0=*100dB*, *GBWP=5MHz*, I_B=*100nA*, V_{OS}=*2mV*) and an analog mux (R_{on}≤*10Ω*, R_{off} almost infinite, I_{leak}<*300nA*).

a) Find out the function of the stage for each *select* and the value of $V_{in,max}$ not to saturate the OA.

b) Compute the maximum input frequency allowed for the *select*=10 to have *12bit* accuracy.



11.19.a

The circuit is a front-end to amplify the signal, the gain can be choose through the analog mux. It's easy to check that the 4 gains are:

$$00 \rightarrow G = 1 + \frac{33k\Omega}{30k\Omega} = 2.1 \qquad 10 \rightarrow G = 1 + \frac{33k\Omega}{10k\Omega} = 4.3$$

$$01 \rightarrow G = 1 + \frac{33k\Omega}{20k\Omega} = 2.65 \qquad 11 \rightarrow G = \infty \text{ quindi agisce da comparatore}$$

Note that in 11 case the gain is null, since the negative terminal of the operational is fixed to ground, and then the feedback cease to work.

Since the OpAmp is powered at *+5V* and is *rail-to-rail*, the maximum input signal that does not saturate the output is *5V/2.1=2.3V* 2.3V in the case *00*, *5V/2.65=1.88V* in the case *01* and *5V/4.1=1.2V* in the case *10*.

11.19.b

The gain of *10* configuration is *4.1*. In addition, the configuration introduces a pole at frequency f_p=*GBWP/4.1=1.2MHz*. It creates a low pass filter that even with infinite time of acquisition may result in an attenuation of the input signal. To compute this attenuation, we have to assess the transfer function of the filter to the signal frequency. We get:

$$\left|\frac{V_{in}}{V_{out}}\right| = \frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}}$$ that leads to: $$\Delta V_{in} = V_{in} - V_{in} \cdot \frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}} \approx V_{in} \cdot \left[\frac{1}{2}\left(f_{in}/f_p\right)^2 + \ldots\right]$$

Therefore to obtain an error less than *1/2LSB*, signal frequency has to be:

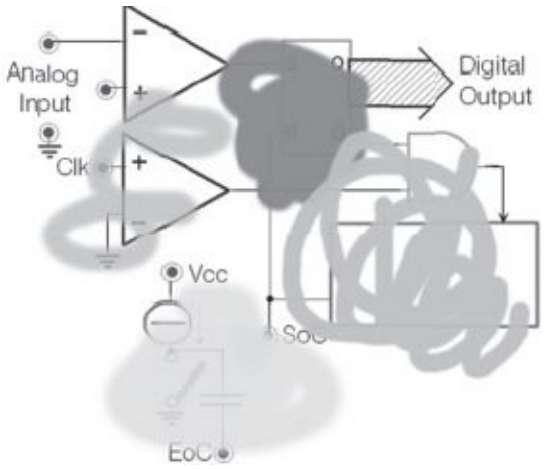
$$f_{\max} < f_p \cdot \sqrt{\frac{LSB}{V_{PP}}} = 1.2MHz \cdot \sqrt{\frac{1mV}{5V}} = 16.97kHz$$

11.20

The stage is a single ramp ADC reviewed by Vavà.

a) Size the circuit in order to have *12bit* accuracy, *FSR=5V*, powered at *5V* and $t_{conv,max}=1ms$.

b) Find out the error computed in *LSB* caused by the *1%* tolerance of *C* and MOS stray capacitors $C_{gd}=C_{ds}=C_{gs}=100pF$.



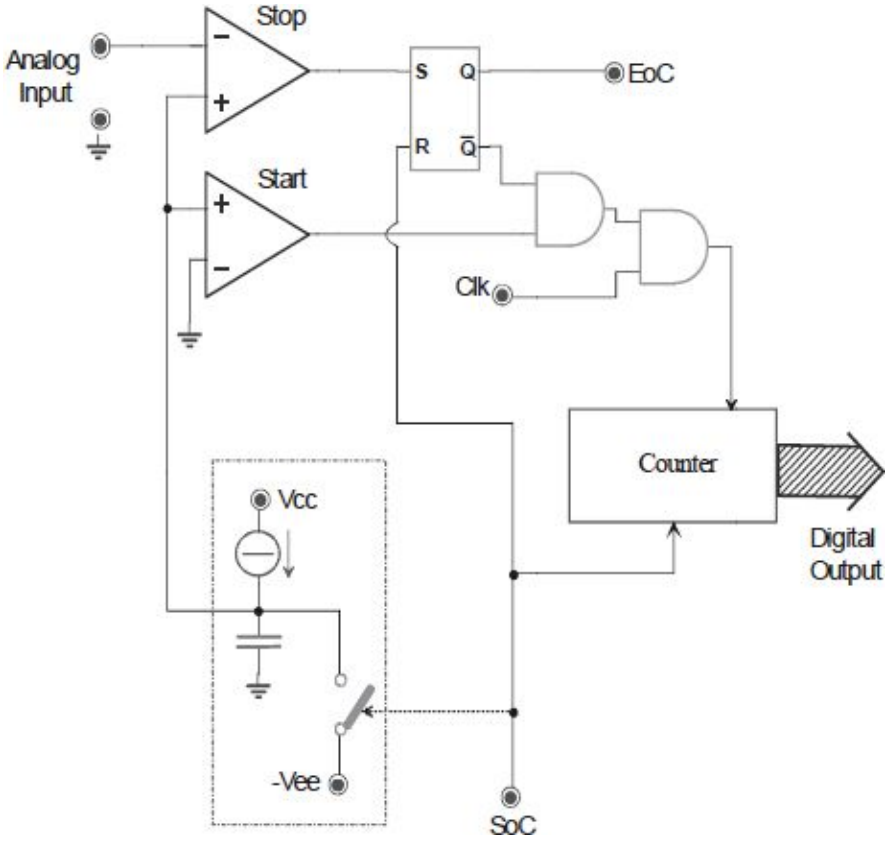
11.20.a

The correct circuit is shown in the picture below. To ensure 12bit accuracy with $t_{conv,max}=1ms$ and *FSR=5V*, over the maximum conversion time, the counter has to counts up to $2^{12}=4096$, so the clock must be *4096/1ms=4.096MHz*. Furthermore, the capacitor has to be charged to *5V* over *1ms*. For example, we can choose *I=1mA* and *C=200nF*.

11.20.b

The capacitor value is in direct proportion to the duration of the voltage ramp. A variation of 1% results in a change of 1% of the ramp duration. Since the coding output is proportional to the length of the ramp, also the output changes by 1%. The worst case is when the ramp raise to the maximum, i.e. $V_{in}=FSR=5V$. The error is equal to *5V/100=50mV*, thus 40LSB.

Also the stray capacitance of the MOS results in errors. C_{ds} and C_{gd} are in parallel to the timing capacitor, slowing down the charge. Because of in our cases this is increment is equivalent to having a capacity 0.1% larger, the error is equal to $5V \cdot 0.1\% = 5mV$, thus 4LSB.



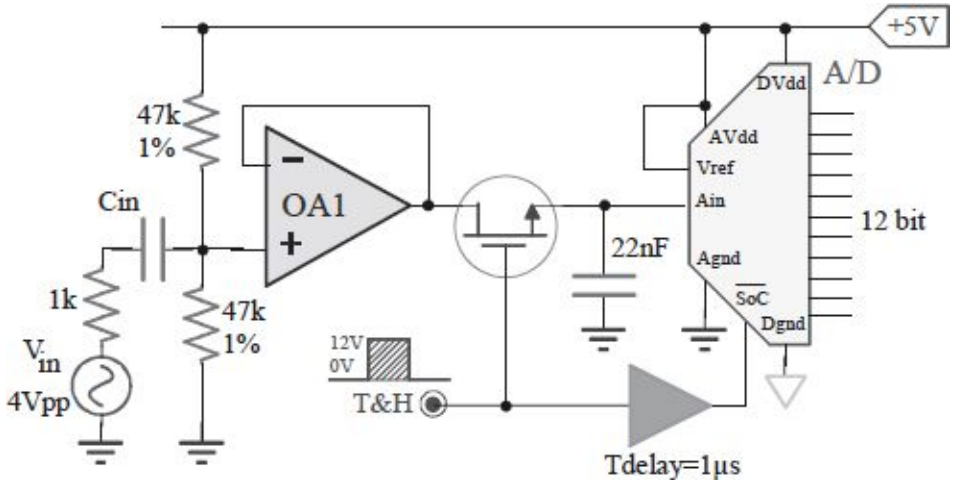
11.21

Given OpAmp with $I_B=100nA$, $V_{OS}=2mV$, $A_0=100dB$, $I_{OUT,max}=10mA$, $SR=10V/\mu s$, MOS with $R_{on}\leq 100\Omega$, $C_{gd}=C_{gs}=C_{ds}=5pF$, ADC with $T_{conv}=5\mu s$ and $I_{in}\leq 200nA$.

a) Compute the maximum input frequency in order to reach the OpAmp limit.

b) Find out the maximum sampling frequency f_S which ensure an accuracy of ± 1/2LSB.

c) Find out total error (static and dynamic) at each conversion.



11.21.a

The OpAmp works to its limits if the input signal is close to GBWP, if it causes the amplifier to reach its *slew-rate* or its maximum output current. In this case we have $SR=10V/\mu s$, while the limit on the output current imposes a maximum slope of the output signal equal to $I_{OUT,max}/C_{hold}=10mA/22nF=0.45V/\mu s$. The latter is the more cogent. Since the amplifier is in buffer configuration, this limits the input frequency to $f_{max}=0.45V/\mu s\ /\ (2\pi\cdot 2V_P)=35.8kHz$.

11.21.b

The presence of MOS R_{on} results in a pole with time constant $\tau=100\Omega\cdot22nF=2.2\mu s$. It creates a low pass filter that even with infinite time of acquisition may result in an attenuation of the input signal. To compute the attenuation, we have to assess the transfer function of the filter at the signal frequency. We get: $\left|\frac{V_{in}}{V_{out}}\right| = \frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}}$

thus:

$$\Delta V_{in} = V_{in} - V_{in}\cdot\frac{1}{\sqrt{1+\left(\frac{f_{in}}{f_p}\right)^2}} \approx V_{in}\cdot\left[\frac{1}{2}\left(f_{in}/f_p\right)^2+\ldots\right]$$

In order to have an error lower than 1/2LSB, signal frequency has to be lower than:

$$f_{\max} < f_p\cdot\sqrt{\frac{LSB}{V_{PP}}} = 72kHz\cdot\sqrt{\frac{1mV}{4V}} = 1.14kHz$$

And so $f_{S,max}=2.28kHz$.

RC network forces also an exponential charge of the hold capacitor. The smaller the error the longer the sampling time, that is:

$$\Delta V = V_{PP}\cdot e^{-T_{acq}/\tau}$$

In order to have an error lower than 1/2LSB, sampling time has to be:

$$T_{acq} > \tau\cdot\ln\frac{V_{PP}}{1/2LSB} = 20\mu s$$

Given a conversion time equal to $5\mu s$, this limits the maximum sampling frequency to $1/25\mu s=40KHz$. This limit is heavier than the one given previously.

11.21.c

In addition to the error just found the stage introduces several others.

OpAmp bias currents causes a static error on the hold capacitor equal to $I_B\cdot47k\Omega/2=2.35mV$, thus almost 3LSB.

The offset voltage causes a static error on hold capacitor equal to a $V_{OS}\cdot1=2mV$, thus 2LSB.

The real OpAmp gain causes an error equal to $V_H/A_0=V_{in}/A_0$, thus dependant on the input. In the worst case ($V_P=2$) it is $(2V+2.5V)/A_0=45\mu V$, that is negligible.

In the S&H MOS switching, due to C_{gs} capacitor, a small charge is injected in the Hold capacitor. In the worst case this contribution is equal to:

$$\Delta V = \Delta V_{gs} \cdot \frac{C_{gs}}{C_{gs} + C_H} = 12V \cdot \frac{5pF}{5pF + 22nF} = 2.7mV$$

almost 3 LSB.

In addition C_{ds} causes in the hold phase a signal feed-through, that in the worst case leads to:

$$\Delta V = \Delta V_{in} \cdot \frac{C_{ds}}{C_{ds} + C_H} = 4V \cdot \frac{5pF}{5pF + 22nF} = 0.9mV$$

So more or less 1 LSB.

ADC leakage causes a droop in the hold phase equal to $I_{in} \cdot T_{hold}/C=200nA \cdot 5\mu s/22nF = 45\mu V$, thus 1/22 LSB.

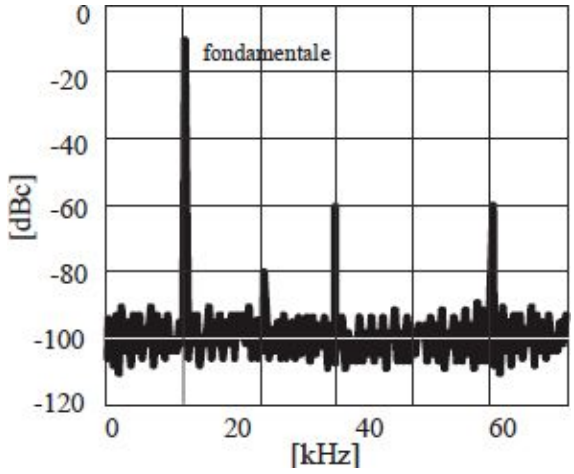
11.22

It's given the spectrum at the output of a *16bit* ADC with *FSR*=5V. La FFT has been computed on a *42ms* data stream.

a) Compute *bin-width*, *noise-floor* theoretic and SNR_{ideal} (maximum amplitude and only quantization noise).

b) Compute $SNR_{theoretic}$ (real amplitude but only quantization noise), $SNR_{measured}$ (both real amplitude and noise) and the following *ENOB*.

c) Compute *THD* and *SiNAD*.



11.22.a

The bin-width is a ratio of the FFT frequency accuracy. Since the sampling frequency is equal to *60kHz·2=120kHz*, and the stream has lasted *42ms*, 5040 samples were acquired. So in frequency the frequency domain every bin will be *120kHz/5040=24Hz* wide.

The SNR_{ideal} is what you get considering only the quantization noise and the signal maximum amplitude. It is easy to get the following relationship, knowing that the quantization noise has a power equal to $LSB^2/12$:

$SNR_{ideale} = 6.02n + 1.76 = 98dB$

The *noise-floor* is the theoretical base noise due only to quantization noise. As we are dealing with spectra obtained by FFT, what we see in the chart is not a power density, but rather the power that lies on a band as wide as the bin-width. The spectrum also displayed only extends from 0 to $f_s/2$. In fact it is periodical and for convenience we can study only half period. These considerations show that the theoretical *noise-floor* is equal to the total noise power split between all the *bins* in $f_s/2$. If we choose to compute the whole in *dBc* (*dB* related to the power of the maximum amplitude carrier), the total noise power is equal to SNR_{ideal} changed sign. Then:

$noisefloor_{teo} = -(6.02n + 1.76) - 10log(60kHz/24Hz) = -98 - 34 = -132dBc$

11.22.b

The $SNR_{theoretic}$ is what you get considering only the quantization noise, but the real extent of the input signal. It will therefore be worse than the SNR_{ideal}, degraded by a factor $10log(P_{FSR}/P_{sig})$. In this case, the degradation is immediately readable from the spectrum, since the fundamental has a power of only *-10dBc*. Thus $SNR_{theoretic} = SNR_{ideale}$*-10dBc=88dB.*

The $SNR_{measured}$ is actually measured, and includes quantization noise, additional noise, and signal amplitude lower than the dynamic. It is still less than $SNR_{theoretic}$. To calculate it should make the real noise from the spectrum. It is nothing but the noise-floor visible spectrum multiplied by the number of bins. Then:

$$SNR_{misurato} = Segnale_{mis} - Rumore_{mis} = -10dBc - \left(-100dBc + 10\log\frac{60kHz}{24Hz}\right) =$$

$$= -10dBc - (-100dBc + 34dB) = 56dB$$

To this *SNR* correspond an *ENOB* equal to only *(56-1.76)/6.02=9bit*.

11.22.c

THD is the total harmonic distortion. It's defined as the ratio between the sum of the power of all the harmonics and the power of the carrier:

$$THD = 10\log\frac{\sum_{i=2}^{n} v_i^2}{v_1^2}$$

Because of we're using powers expressed in *dBc*, we get:

$$THD = 10\log\left(10^{\frac{-80+10}{10}} + 10^{\frac{-60+10}{10}} + 10^{\frac{-60+10}{10}}\right) = -46.9dB$$

SiNAD is the ratio between the carrier power and that of the sum of the power of all the harmonics and the noise:

$$SiNAD = \frac{signal}{noise + harmonic} = -10dBc - 10\log\left(10^{\frac{-66}{10}} + 10^{\frac{-80}{10}} + 10^{\frac{-60}{10}} + 10^{\frac{-60}{10}}\right) = 46.5dB$$

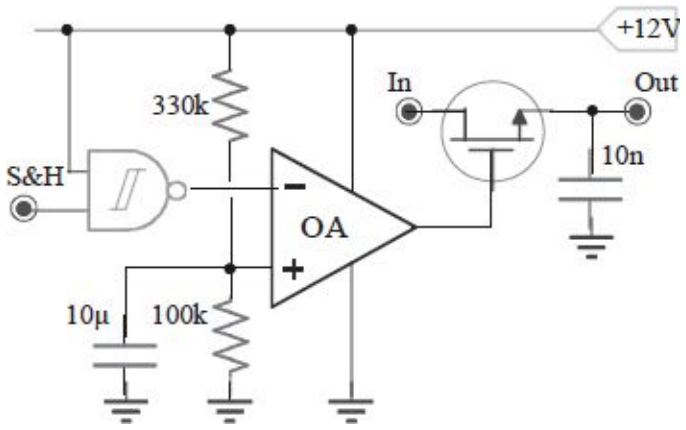
As expected the *SiNAD* is dominated by the power of harmonics, in fact, it is only *10dB* worse than the *SNR.*

11.23

The comparator has *SR=15V/μs*, the *5V* input CMOS command has t_{rise}*=20ns*. The input changes between *0÷5V*. with *1kHz* as maximum frequency. MOS has V_T*= 2V*.

a) Compute both minimum and maximum opening time (different because of the change of the input signal) and find out the subsequent *opening-time error* at the output.

b) Compute the *aperture time-jitter error* using resistor noise and the *12V* power supply ripple by *100mV* at *100Hz*.



11.23.a

The opening time is the time between the front of the S&H signal is the actual opening of the MOS. In this case it depends on the delay introduced by logic port to bring its output to the threshold of the comparator, and the time required to the comparator output to bring the MOS sub-threshold.

The first contribution is independent of the input signal. The MOS opens when its output switches from high to low level, i.e. when its inverting input goes from low to high. The switching threshold of the comparator is *2.8V*. The NAND output takes *20ns* to switch from *0V* to *12V*, then it takes *4.6ns* to cross the threshold. At this point the output of the comparator begins to move at a speed imposed by its slew rate of *15V/μs*. If the signal was the at lowest value

(*0V*), the gate has to reach *2V* for the MOS to switch off, thus the time required is *(12V-2V)/ 15V/μs=0.6μs*. However, if the signal was at the maximum value (*5V*), the gate has to reach *7V* for the MOS to switch off, and so the time required is *(12V-7V)/ 15V/μs=0.3μs*. In both cases, the opening delay is forced by the comparator *slew-rate*.

In the case we're interested in a synchronous sampling of the signal, the delay in the opening results in an error on the sampled value, the *aperture-time error*. This error is greater the higher the slope of the input signal. In this case the maximum slope of the input signal is $V_P \cdot \omega = 2.5V \cdot 2\pi \cdot 1kHz = 0.017V/\mu s$, then in the two extreme the error is equal to $0.017V/\mu s \cdot 0.6\mu s = 10mV$ and $0.017V/\mu s \cdot 0.3\mu s = 5mV$.

11.23.b

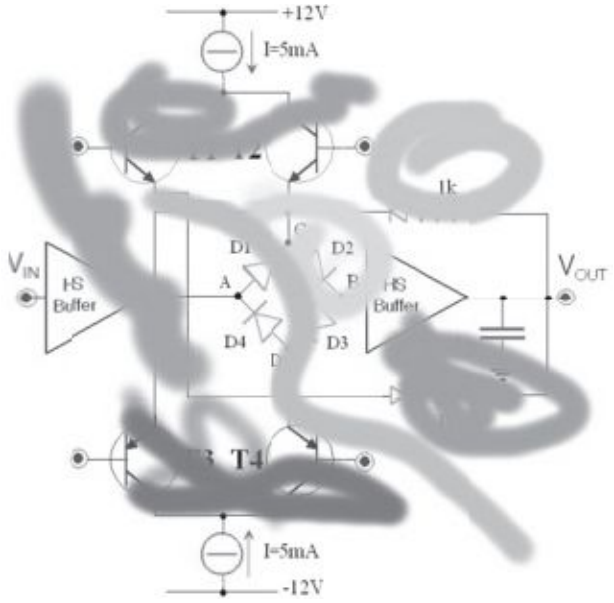
Both resistors noise and power supply ripple ensure that the reference node of the comparator is floating referred to its ideal value. This causes jitter on the threshold crossing, which is reflected as a jitter on the opening time and then the error. The noises of the two resistors causes a noise power density at the positive pin equal to $(4kT/R_1 + 4kT/R_2) \cdot (R_1 // R_2)^2 = (34nV/\sqrt{Hz})^2$. This noise has a pole at frequency $1/(2\pi C \cdot R_1 // R_2) \cdot \pi/2 = 0.3Hz$, thus the effective noise is equal to 18nV. Since the control signal of the comparator has a speed of *12V/20ns*, the error introduced is surely negligible. The power supply ripple with *100mV* amplitude at *100Hz* causes an effective ripple on the reference pin equal to $100mV/\sqrt{2} \cdot Z(100Hz)/(Z(100Hz) + 330k\Omega)$, where *Z(100Hz)* is the impedance of the parallel between the capacitor and the *100kΩ* resistor at *100Hz*. We can compute $Z(100Hz) \cong 160\Omega$, then the effective ripple is equal to *35μV*. Since the comparator control signal has *12V/20ns* speed, jitter opening appears to be negligible in our case.

11.24

Valentina and Marco have bungled the S&H, which employs diode described as $V_{on}=0.5V$ and $R_{on}=10\Omega$.

a) Draw the proper circuit and the level translation network required to drive the transistor through a single CMOS command.

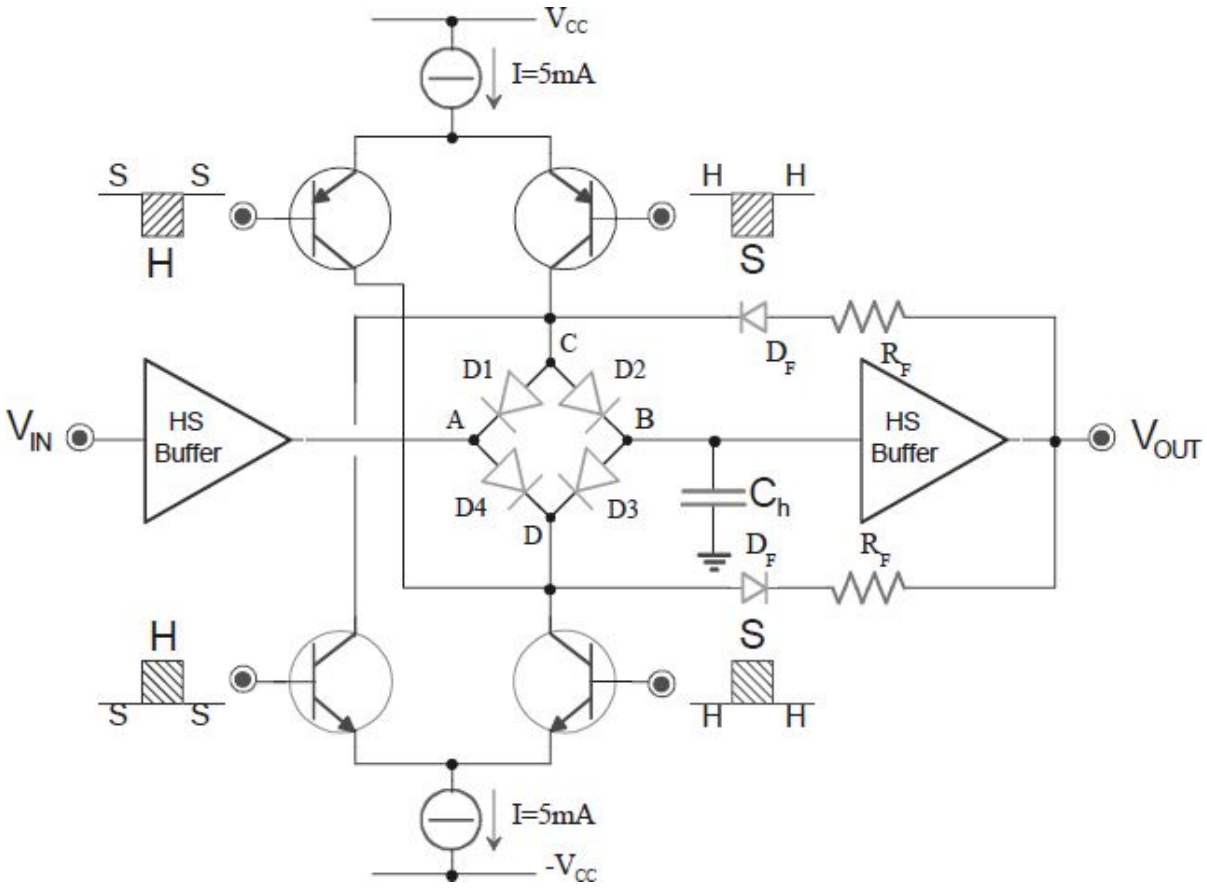
b) Given a mismatch of *1%* in the current generators and a mismatch in timing of 1ns transistors, find out the error on V_{out}.



11.24.a

The proper circuit is drawn in the following picture.

To drive the four transistors with a single CMOS command we have to use two differential stages, one for the upper pair and one for the bottom, where a transistors of the pair is connected to *+2.5V* and the other to the CMOS control. Neglecting *BJT* base currents, the upper differential stage will ensure a swing from *+12V* to $+12V-R_L \cdot I_{coda}$ while the lower stage will grant swing from *-12V* a $-12V+R_L \cdot I_{coda}$. The tail current and the load resistors will be chosen to ensure the switch on of the S&H transistor.



11.24.b

In the case of a transistors timing mismatch of 1ns of transistors, for example the case in which the transistor in the upper right is already switched on while those below are both off, we get an error equal to $1ns \cdot 5mA/C_{hold}$. To be honest, the error will be smaller, because during the switching process the current is gradually switched from a transistor to another of the lower pair.

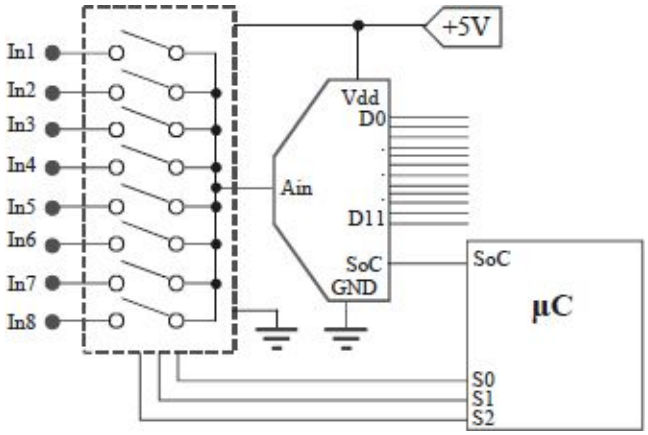
Given a mismatch of 1% in the current generators, node A and node B are no longer the same potential. We can study for instances the case when the upper source carries a current greater than the lower one, and that the potential of A and B are initially equal. The excess current flow equally in D1 and D2, then, unable to flow through in D3 and D4, go through C_{hold} and ends in the buffer. The voltage of node A does not change, as being imposed by the buffer, while the node B rises. In this situation, the current in D2 will begin to decrease, while the one in D1 increases. At regime state all the excess current will flow through the buffer and D2 will be turned off. The exact value of the voltage at

node B is not easy to compute, since it requires the study of the translinear network composed of four diodes. The error will still be in the hundreds of mV.

11.25

A *12bit* ADC that has $T_{conv}=10\mu s$, has to get 8 analog input cycling (each with the same bandwidth and sampled as Nyquist). The inputs have $R_S=200\Omega \div 10k\Omega$, mux has R_{on} negligible and S&H in the ADC has $C_H=50pF$.

a) Compute the minimum $T_{acquisition}$, maximum throughput at the output (in samples/sec, thus ADC conversions/second) and the maximum bandwidth of the input signals inputs.



11.25.a

When you want to capture a signal via an ADC we have to give the Sample & Hold time enough to ensure that the voltage level reached by the hold capacitor and the regime value mismatches for less than *1/2LSB*. This ensures that the errors caused by the Sample&Hold aren't higher than the quantization noises.

In this case, the pole of the exponential charge of the hold capacitor is forced by the product of the capacitance itself and the source resistance, because of R_{on} is negligible. When considering the worst case, we should use as R_S the highest value, i.e. *10kΩ*.

Inverting the expression of the exponential charging of a capacitor through a resistor, we get immediately the lowest acquisition time ensuring an error *<1/2LSB* when compared with the regime:

$$T_{acq} > \tau_{\max} \ln\frac{FSR}{\varepsilon} = 10k\Omega \cdot 50\,pF \cdot \ln\frac{FSR}{1/2LSB} = 500ns \cdot \ln 2^{n+1} = 4.5\,\mu s$$

To compute the maximum throughput, i.e. the maximum number of conversions per second that the ADC is able to do, we must consider that each acquisition requires a sample time (*4.5μs*), a conversion time (*10μs*) and the time for the mux to switch (we can suppose in the order of *500ns*). Overall, the conversion requires 15μs, hence the maximum throughput is equal to *1/15μs=66ksps*.

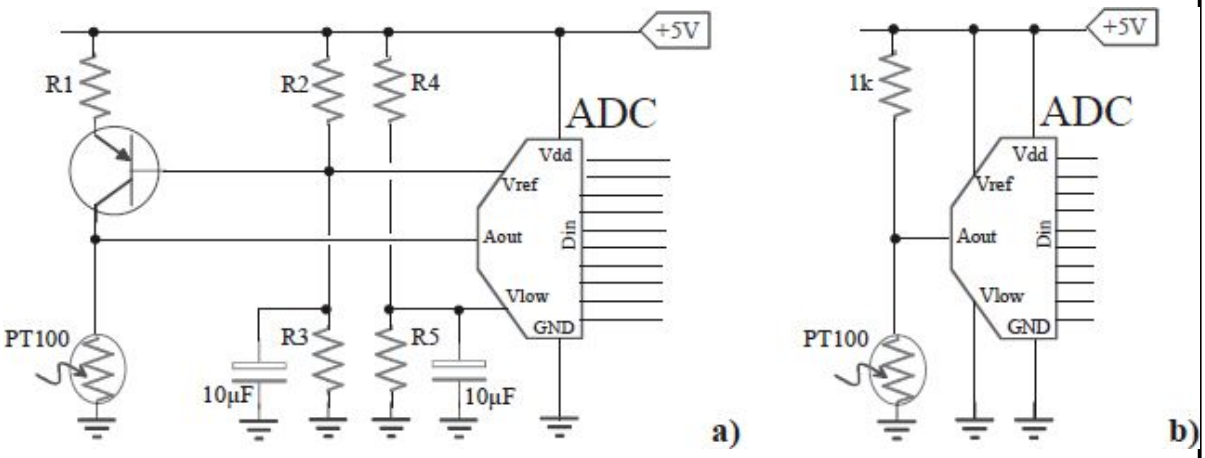
The channels are 8, thus each channel will be scanned every *15μs·8=120μs*, i.e. with a frequency of *8333sps*. Since the Nyquist sampling is done, the maximum frequency of the input signal to prevent aliasing is f_{max}*=8333/2=4.15kHz*.

11.26

The temperature is measured between *0÷500°C* with *10bit* accuracy. The PT100 shows *100Ω* at *0°C* and increases *0.4Ω/°C*.

a) Given the ADC with *10bit* in the picture, compute the values of the circuit in (a) in order to use at best the allowed dynamic.

b) Given instead the simplified circuit in (b): analize the not linearity introduced and choose the correct ADC in order to maintain the same accuracy as before.

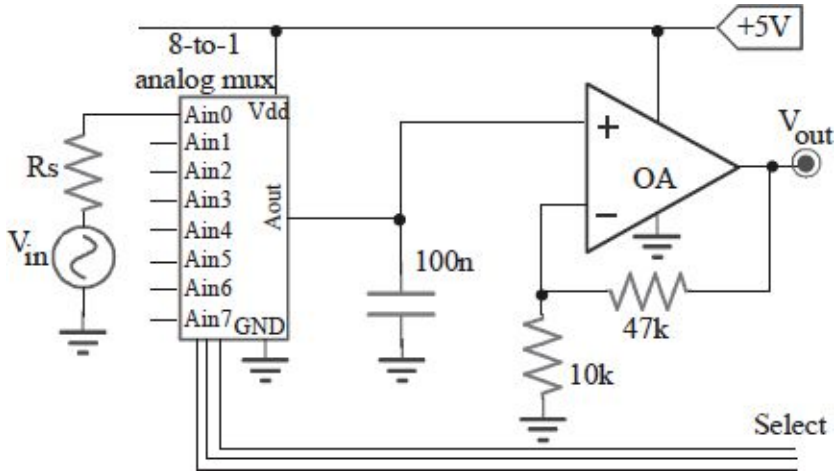


11.27

There are 8 sensors with *Rs* from *1kΩ* to *10kΩ* and signal *Vin* from *0V* to *400mV*.

a) Compute the maximum static error (both in Volt and bit), given a MUX with R_{on}*=100Ω*, R_{off}*=10MΩ*, $I_{leakage}$*=10nA* and the OpAmp with I_{bias}*=50nA* e R_{in}*=100MΩ*.

b) Compute the maximum sample time in order to guarantee *10bit* accuracy.

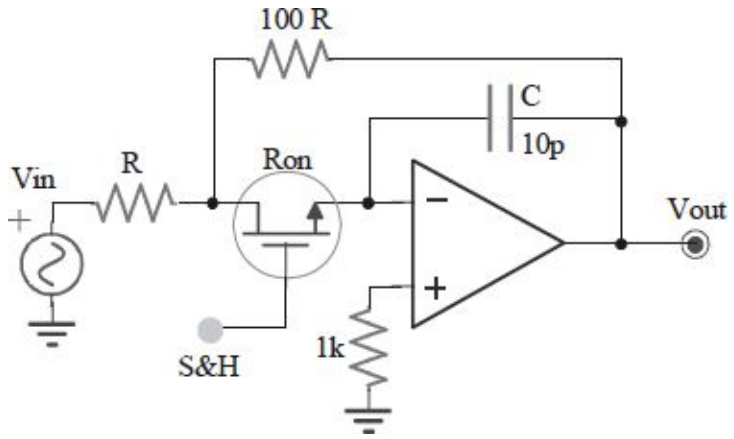


11.28

The OpAmp has a *GBWP=17MHz*.

a) Compute the pole in the sampling phase given the resistor of the closed MOS switch *Ron=70Ω*; choose the correct value of *R*.

b) What happens to stability if *R=400Ω*.

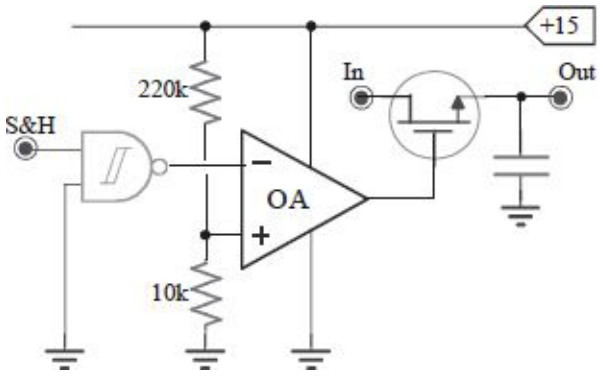


11.29

This circuit pilots the gate of a Sample&Hold, through a comparator (*4nV/√Hz, 1pA/√Hz, $C_{in,diff}$=1pF* e *GBWP=30MHz*). The command is given with a CMOS, t_{rise}≈*10ns*.

a) Find out the effective noise overlapped to the comparator threshold and the subsequent time jitter.

b) Find out what is limiting the jitter in the opening of the MOS, and provide some numerical example.

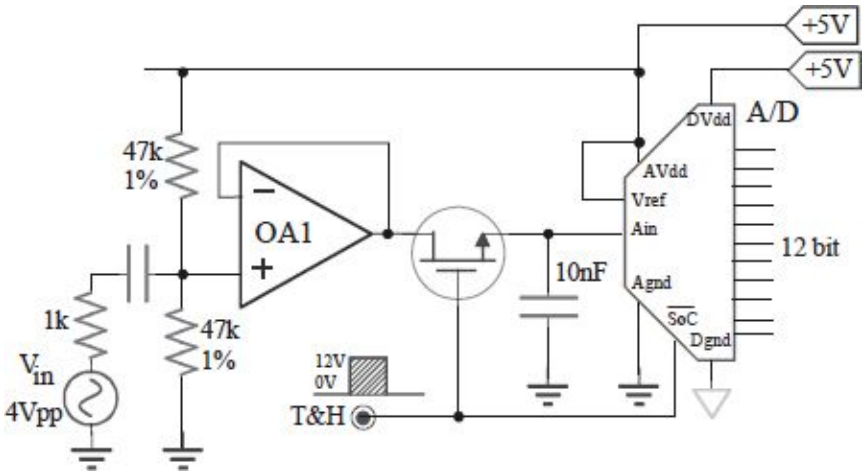


11.30

OpAmp with IB= -500nA, VOS=3mV, A0=100dB, CMRR=90dB. *12bit* ADC with Vref=5V.

a) Compute the static error in *LSB*, caused by *IB*, *VOS* e tolerances in the resistors. Put forward how to modify the circuit in order to wipe out those errors after the MOS.

b) Compute the errors caused by *A0* and *CMRR*.

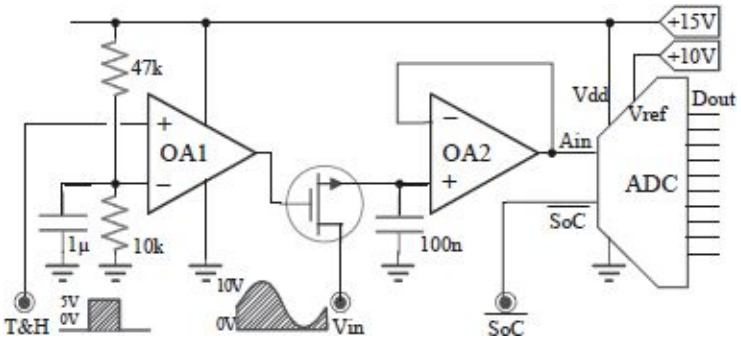


11.31

The command of the T&H ha $t_{rise}=t_{fall}=10ns$. Gli OpAmp rail-to-rail hanno $PSRR=-80dB$, $SR=10V/\mu s$, $I_{outmax}=\pm 10mA$. Il MOS ha $V_t=2V$. L'ingresso è $0 \div +10V$ ad $f_{inmax}=20kHz$.

a) Compute the maximum error in LSB due to aperture delay.

b) Compute the maximum error due to power supply ripple of $1V$.



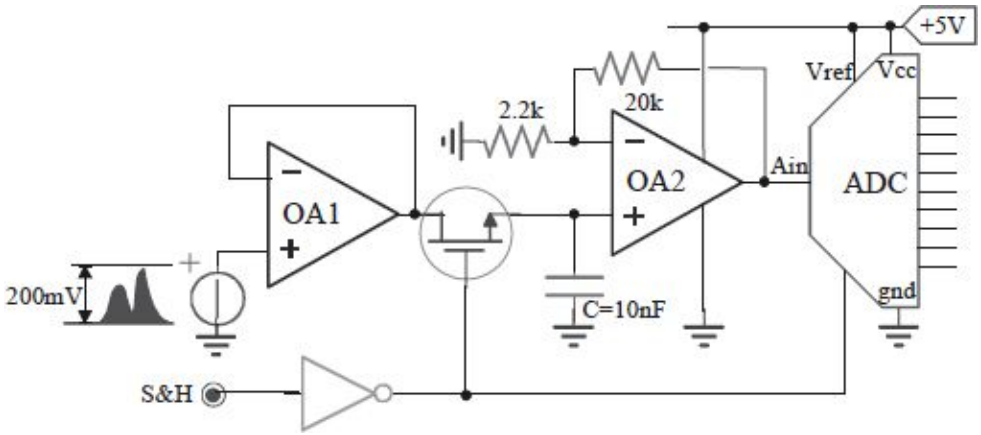
11.32

Given an input signal between *0*V and *200mV*. Find out:

a) the numbers of bit in order to obtain an accuracy of 1/3000 at the input signal.

b) The maximum *charge injection* computed in *LSB*, caused by the capacitors of the MOS (C_{stray}*=1pF*).

c) the static error overlapped to *Ain* during the sampling phase (R_{DSon}*=50Ω*) caused by the bias currents pushed from OA2 and equal to *50nA*.

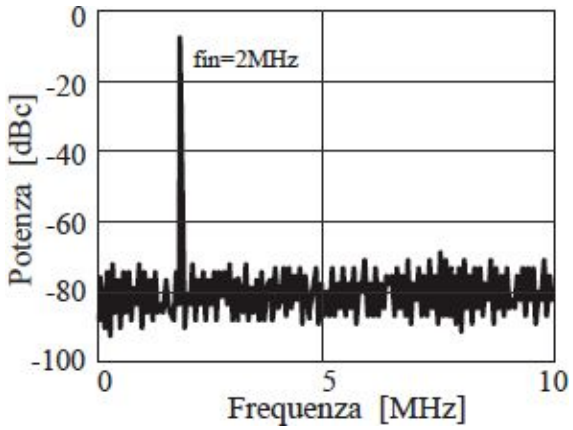


11.33

It's given an ADC with *12bit* accuracy, *Vref=10V* and the FFT drawn in the picture (it has been used *512* samples of the input sinusoid with *7Vpp*).

a) Forecast the theoretic *noise-floor* and compute the real *SNR*.

b) Estimate the number of samples *N* needed to see a harmonic distortion with power *76dB*.

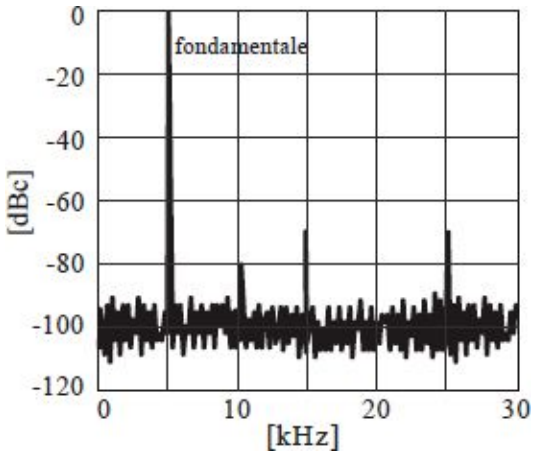


11.34

It’s given in the picture the FFT of a *12bit* ADC with *FSR=5V*. The FFT has been computed with a data stream of *170ms*.

a) Compute both the theoretic SNR and the measured one. Compute the following ENOB.

b) Compute the SiNAD and the THD.

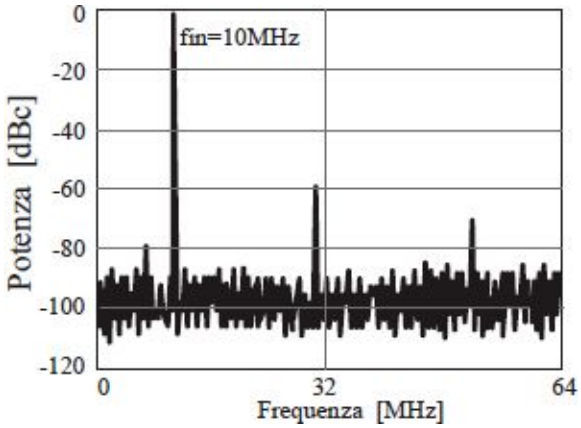


11.35

The spectrum in the picture has a *bin-width* of *1kHz* and represent the output of a Sample&Hold SHC605 given an input signal between *±1V*.

a) Compute the real SNR_{real} and the *SFDR*.

b) Compute the *THD* and the *SiNAD*.

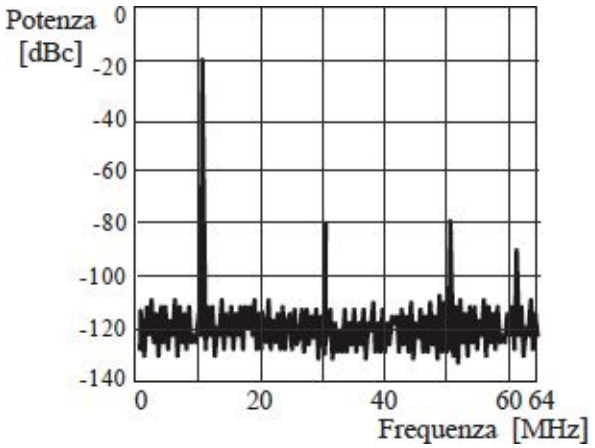


11.36

The ADC has FSR from *–2.5V* to *+2.5V*. The spectrum at the output in the picture has a *bin-width=2kHz*.

a) Compute the real *SNR*, estimate the number of bits needed by the ADC and compute the *ENOB*.

b) Compute the *THD*, the *SiNAD* and the *SFDR*.

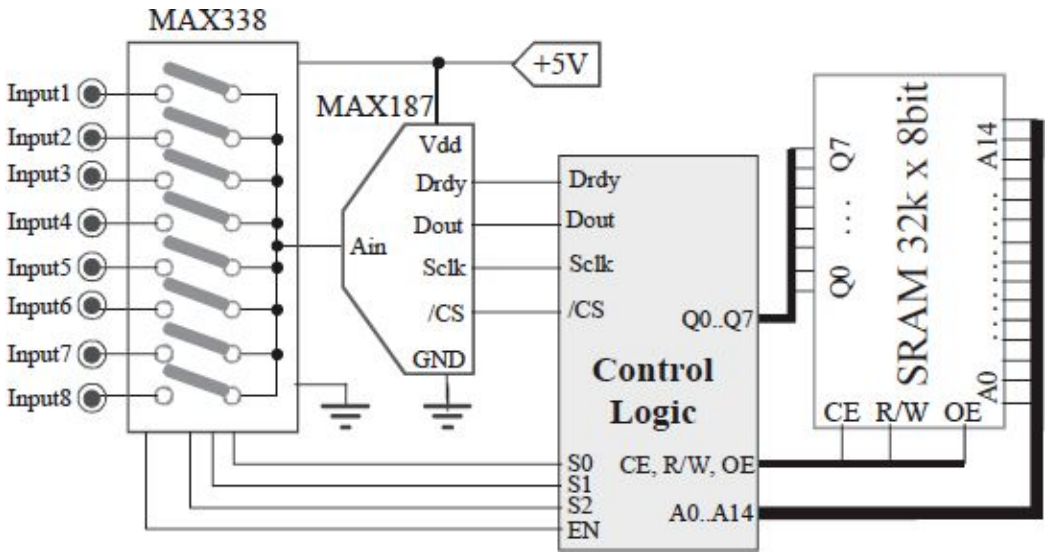


11.37

It's given a serial *12 bit* ADC that acquires 8 analog inputs and store them in a SRAM 32k x 8bit the 8 MSB. The signals are read with a cycle at every 1ms.

a) Project the Control Logic stage in order to control both the mux and the ADC and that store the values in 8 latches.

b) Project the control logic stage that writes the values of the 8 in the memory.

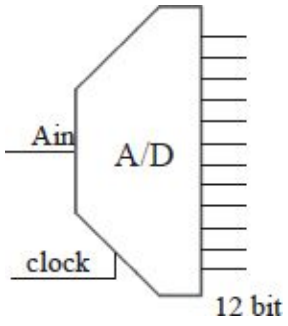


11.38

The IC ADS801 is a *12bit* pipeline ADC that works at *25MHz*, with *GainError=±1.5%*, *GainTempCo=±85ppm/°C* and *Offset error=±2.5%*. The power required is *340mW* and it has ϑ_{ja}*=75°C/W*.

a) Compute the number of codes lost in the worst case due to gain and offset errors.

b) It's given a *data latency=6.5convert cycles*. Draw the conversion timing (if *clock=high* were in tracking phase; the sampling is in time with the *falling edge*).

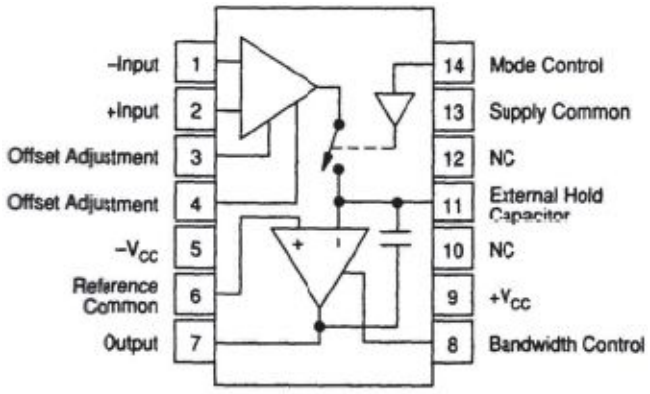

Ain
A/D
clock
12 bit

11.39

Project an inverting S&H with a gain of *10*, using the IC in the picture.

a) Draw both the schematic and the layout, adding a hold capacitor of *100nF*.

b) We want to be able to modify digitally the gain of the S&H between *1*, *10*, *100* and *1000*. Project all the selection electronics.



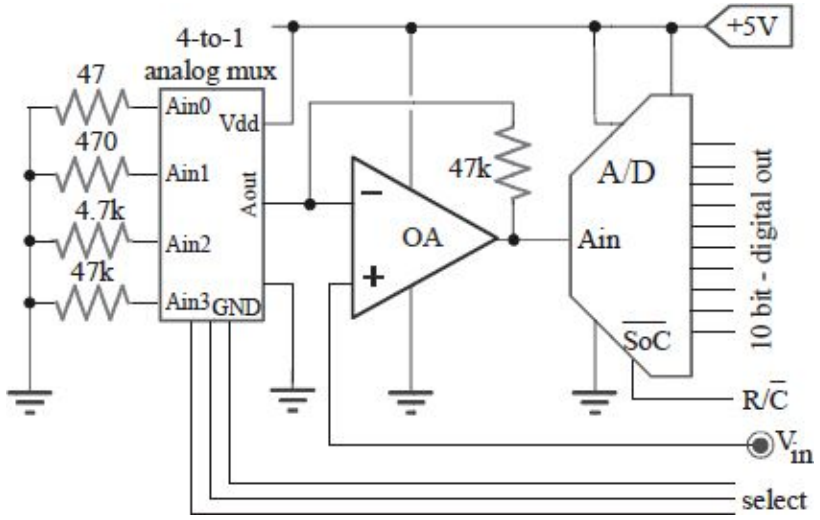
SHC5320KH, KP

11.40

The conversion circuit uses an OpAmp (A_0=*100dB*, *GBWP=50MHz*, I_B=*100nA*, V_{OS}=*2mV*) and an analog mux (R_{on}≤*10Ω*, R_{off}≥*1MΩ*, I_{leak}<*300nA*).

a) Find out the maximum **static** error (both in Volt e number of bit) and the worst case scenario for the *select* lines.

b) Compute the maximum input frequency that guarantees a *10bit* accuracy and the following worst case scenario for the *select* lines.

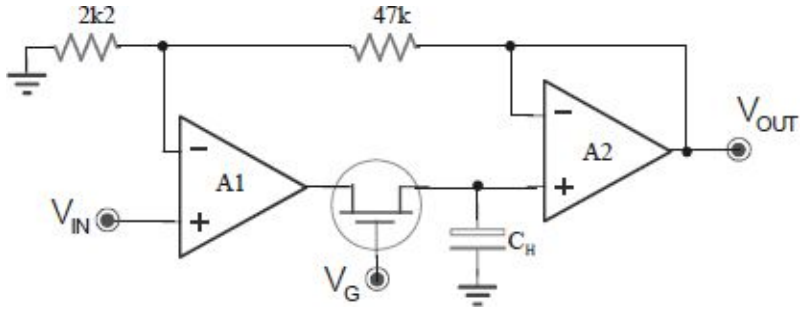


11.41

The S&H uses two compensated OpAmp, both with *IB=500nA*, *VOS=3mV*, *A0=100dB*, *CMRR=80dB* and *GBWP=5MHz*. The input has a bandwidth of *100kHz*. The MOS has R_{on}<*100Ω* and C_H=*10nF*.

a) Quote the maximum static error at the output.

b) Quote the maximum dynamic error given a sampling frequency of *500ksps* and a holt time of T_{hold}=*1.5μs*.


2k2
47k
A1
A2
V_{IN}
V_{OUT}
V_G
C_H

11.42

The µC has to be able to control the current that flows in the resistive sensor from *0.5mA÷32mA* through a *6bit* DAC.

a) Project the programmable current source able to inject the requested current in the sensor connected to ground with an expected value from *10÷100Ω*.

b) Amplify the signal in order to feed it as input to the ADC in the µC, thus to be able to use all its dynamic *0÷5V*.

11.43

The peek stretcher is fed at the input with a serial digital stream at 512kbit/s and has to feed two outputs: the first one is an positive analog one at *8bit*, with the input peeks stretched to 1ms, the second one is digital and provides a high value in time with the stretch.

a) Project the Analog Converter, without using any commercial ADC.

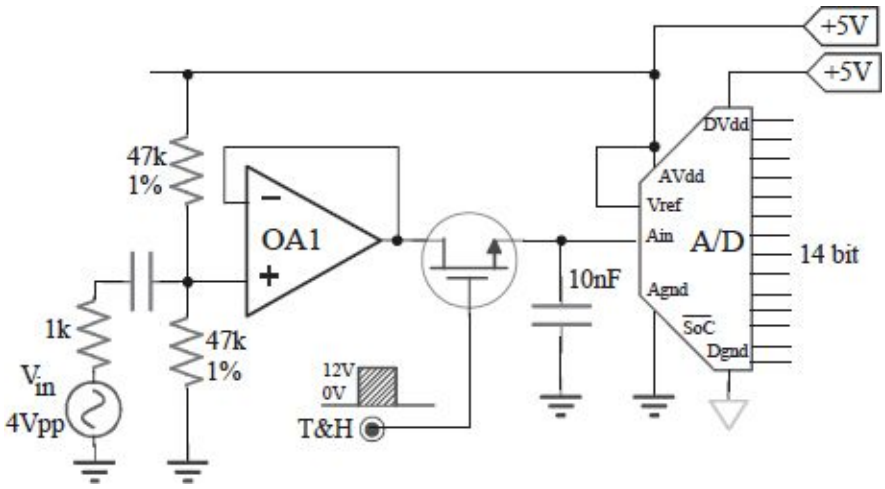
b) Project the Peak Detector, and in particular its catching accuracy (at least better than 1LSB).

11.44

The OpAmp in the circuit has $v_{in}/\sqrt{\Delta f}=10nV/\sqrt{Hz}$ and $i_{in}/\sqrt{\Delta f}=20pA/\sqrt{Hz^2}$, IB=-200nA, VOS=4mV, A0=100dB, GBWP=50MHz. After the S&H there's a *14bit* ADC with Vref=5V. The MOS has $R_{on}(V_{GS}=10V)=100\Omega$.

a) Draw the noise spectral density at the output of OA1.

b) Find out the rms error on the takeover caused by that noise.



11.45

It’s given a *12bit* pipeline ADC that works at *25MHz*, with *GainError=±1.5%*, *GainTempCo=±85ppm/°C*, *Offset error=±2.5%* and that uses *340mW*, ϑ_{ja}*=75°C/W.*

a) Compute the overall error in terms of codes, in the worst case scenario.

b) It’s given also *data latency=6.5convert cycles*. Draw the conversion timing (when *clock=high* we’re in tracking phase; the sampling is in time with the *falling edge*).

11.46

The circuit has to follow the input voltage (with a range *0÷5V*) through a simple digital network that drives a DigPot (Digital Potentiometer) with *64taps* by *100Ω* each, choose on a comparison.

a) Project the whole circuit in order to be able to follow full scale signals at *50kHz* frequency.

b) Modify the circuit in order to have at the output the quantizated current $I_{out}/V_{input}=1mA/V$.

11.47

In the circuit from **Es.11.47** introduce the following IC S&H.

a) Remove OA1 from the **11.47** and resize the circuit with the given S&H coupled in DC with the input in order to have the best condition for a signal at *200Hz* with amplitude *±100mV*.

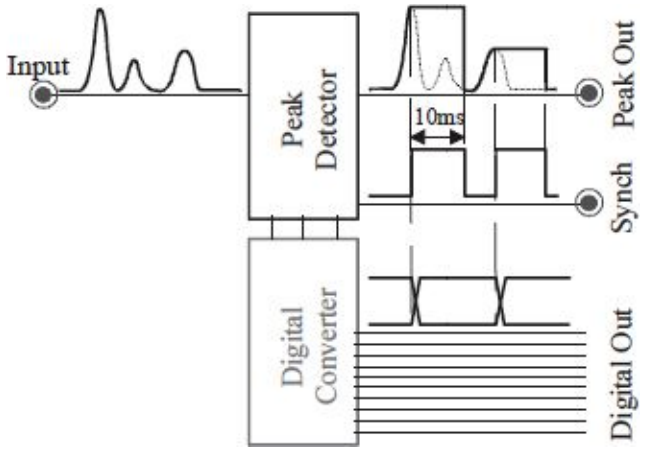
b) Estimate both $I_{out,max}$ and *SR* of the OTA and the OpAmp in the IC in order to guarantee a sample time less than *100ns* ($C_H=100pF$).

11.48

The peek stretcher has to feed two outputs: the first one is analog with the peek stretched to *10ms*, e second one is digital with a *high* level in time with the stretch.

a) Project the Peak Detector, with a major attention to its catching accuracy, that has to be better than *10mV* given an input between *0÷+5V*.

b) Project the Digital Converter provides the peek values, converted at *10bit*, at the beginning of the stretch (that means at the rising-edge of the *Synch* output).



11.49

The OpAmp has A_0=*100dB*, $I_b \leq$*100nA*, *GBWP=40MHz* and *20nV/√Hz* noise; the MOS has V_t=*1.5V*, $C_{GD}=C_{GS}$=*10pF* and $R_{DSon} \leq 50\Omega$; The input changes from *0÷2V*; the ADC has *12bit*.

a) Compute both *SNRideal* and *SNRreal* at the output of the ADC.

b) Compute the static errors stored after *2µs* in the Hold phase, given R_S=*10kΩ*.

c) Modify the circuit in order to use the whole *FSR* of the ADC. It has to be used a feedbacked commercial S&H.

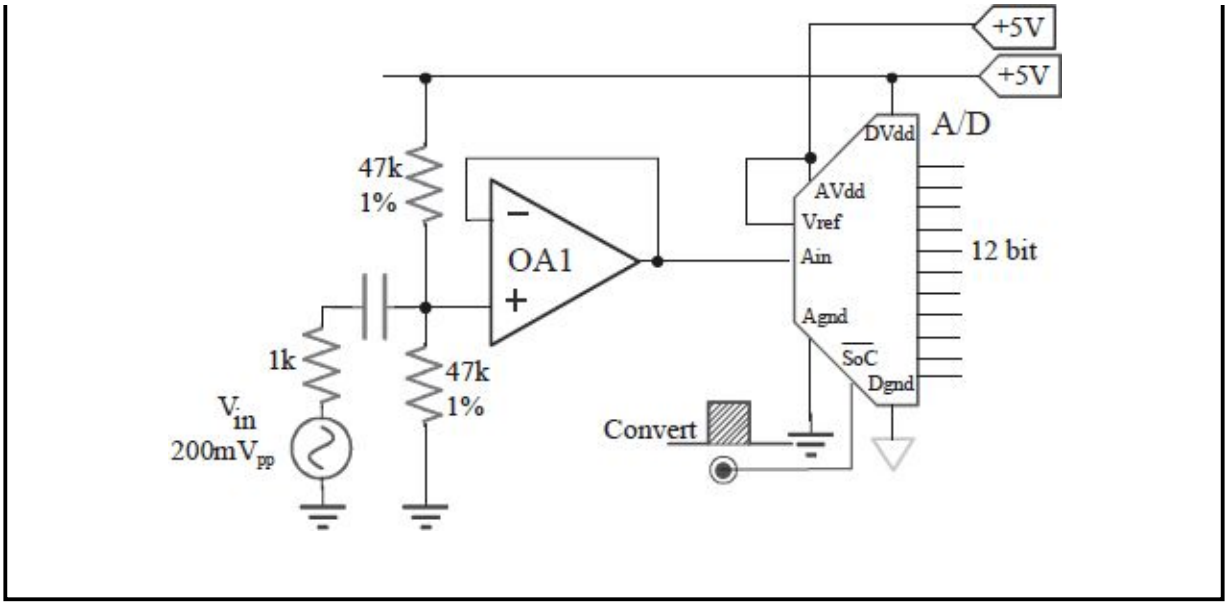
11.50

It's given an OpAmp with *A0=100dB*, *GBWP=3MHz*, *CMRR=80dB*, *SR=10V/µs*. The input signal has a bandwidth of *200Hz*. The ADC is a *12bit* SAR.

a) Compute the maximum sample time that allow the converter to work correctly without any *S&H*.

b) Compute ENOB and SNR given an input with ±*100mV* and the noises of the resistors.

c) Compute the error caused by the *CMRR*.


+5V
+5V
A/D
DVdd
AVdd
Vref
Ain
Agnd
SoC
Dgnd
12 bit
47k
1%
OA1
1k
47k
1%
V in
200mV pp
Convert

11.51

The DigPot has *64taps* and forces the output to GND if it isn't enabled. Sensor signal changes from *0÷5V*.

a) Draw the flow of the outputs when *Control=High* and the sensor gives a sinusoid with maximum amplitude and frequency equal to *1Hz*.

b) Discuss what would happen if we change all the resistor of the circuit with *1kΩ* resistors.

12

Microcontrollers

“She was a fast machine
She kept her motor clean
She was the best damn woman I had ever seen [...that I, ever seen]
She had the sightless eyes
Telling me no lies
Knockin’ me out with those American thighs
Taking more than her share
Had me fighting for air
She told me to come but I was already there
‘Cause the walls start shaking
The earth was quaking
My mind was aching
And we were making it and you.”

“You Shook me all night long”, AC-DC.

12.1. INTERNAL ARCHITECTURE

Microcontrollers (µC) are similar to microprocessors (µP), but they have additional functions on-chip at the expense sometimes of a lesser power. It’s the on-chip peripherals to provide their high flexibility of use and breadth of applications.

In order to simplify the analysis, among the various commercial products we will considered those of medium-low cost, ie 8 or 16-bit architectures and pinouts from 8 to 100 pins. Performance kinds and availability of devices are different, depending on the complexity of the product, but the basics of concepts and structures are common. We will illustrate low cost µC (with prices less that 10US$) because we can’t describe all the comercial µC and we dont want to speack of ideals model. Some examples will be Microchip PIC16C7X family and ST6xx STMicroelectronics family. A current market standard

consists of the ARM and CORTEX μC, distributed by various microelectronics brands.

Unlike old μP systems in which we have a Von Neumman architecture with data and address bus shared between cores, memory and peripherals; in the μC is more common an Harvard architecture system. In it we have a specific bus (both data and addresses) for the program instructions (that connects core and program memory, ROM or EEPROM) and another one for data (memory and peripherals). In the μC are present both the program memory (can be written only by the programmer) and the data memory (volatile and rewritable). There are also auxiliary registers such as program counter, stack registers and records of the status of internal devices. It ‘s exactly the presence of internal devices that differs μC from μP. While μP are suitable to do extensive computing, but are devoid of RAM, ROM and on-chip peripherals, μC are fully autonomous and complete: they contain RAM, ROM, sometimes E^2PROM, input/output ports, timer and ADCs.

Fig. 12.1 shows the block diagram and pinout of the ST62T65 microcontroller from STMicroelectronics, the real core is surrounded by on-chip peripherals. Some pins have two or even three functions that can be set via firmware by the designer.

This μC charateristic are:

- PORTS (PA, PB, PC): Input/Output ports (I/O), even analog ones (Ain), thanks to an internal analog to digital converter (ADC).
- TIMER (TIMin- TIMout): used for counters.
- Communications Port (Sin, Sout, Sck): serial ports to reduce the number of pins.
- RAM: small capabilities, can be used to store a few data (still not enough to store the data of a complete sampled signal).
- OTP/EPROM: the program memory. It can be either OTP (one-time programmable) or programmable and then erasable (optically) by the designer. For example, prototypes can be built on reprogrammable μC (more expensive) and then use non-reprogrammable devices for production. The OTP can be programmed by the designer (spending time and money for programming equipment) or by asking the same parent company to make a “mask programming” in the design phase of the form of metal (the so-called ROM version). The latter option is the more convenient the higher the number of chips produced. E^2PROM can also be used: one advantage is the

easy reprogrammation on-board without removing the chip from the board, in fact the erase is electrically and not through UV exposure.

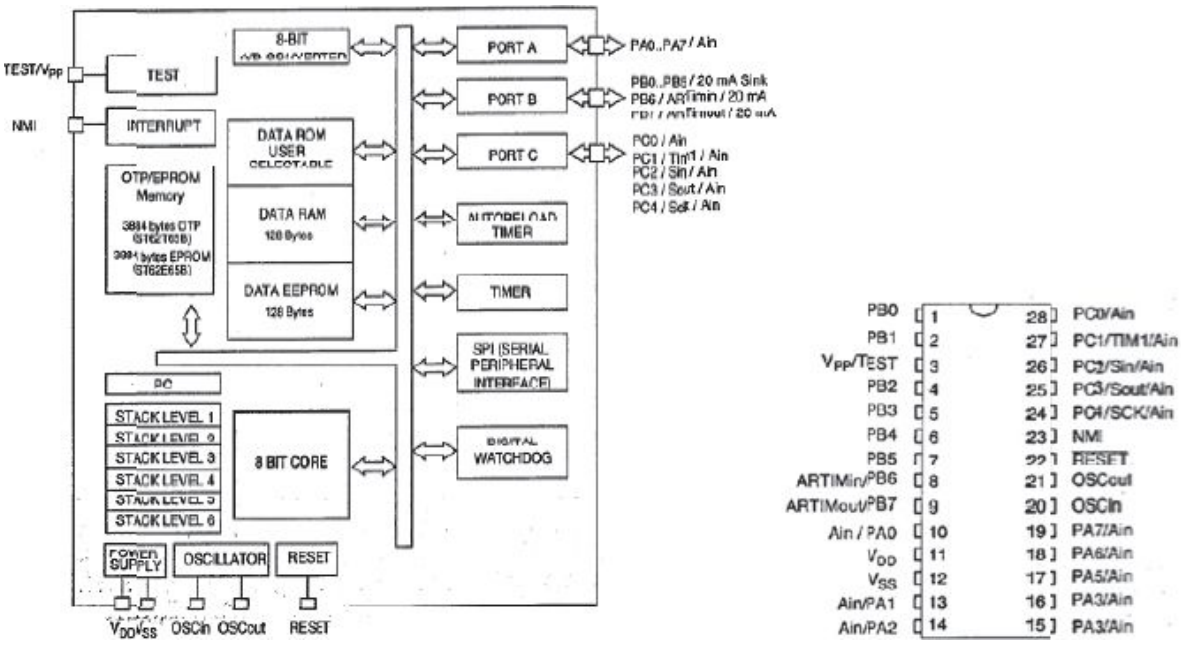


Fig. 12.1: Internal structure (left) and pins (right) of STMicroelectronics ST6265.

SGS-THOMSON MICROELECTRONICS

ST62T65B/E65B

8-BIT MCUs WITH A/D CONVERTER, AUTO-RELOAD TIMER, EEPROM AND SPI

- 3.0 to 6.0V Supply Operating Range
- 8 MHz Maximum Clock Frequency
- -40 to +85°C Operating Temperature Range
- Run, Wait and Stop Modes
- 5 Interrupt Vectors
- Look-up Table capability in Program Memory
- Data Storage in Program Memory: User selectable size
- Data RAM: 128 bytes
- Data EEPROM: 128 bytes
- User Programmable Options
- 21 I/O pins, fully programmable as:
  - Input with pull-up resistor
  - Input without pull-up resistor
  - Input with interrupt generation
  - Open-drain or push-pull output
  - Analog Input
- 8 I/O lines can sink up to 20mA to drive LEDs or TRIACs directly
- 8-bit Timer/Counter with 7-bit programmable prescaler
- 8-bit Auto-reload Timer with 7-bit programmable prescaler (AR Timer)
- Digital Watchdog
- 8-bit A/D Converter with 13 analog inputs
- 8-bit Synchronous Peripheral Interface (SPI)
- On-chip Clock oscillator can be driven by Quartz Crystal Ceramic resonator or RC network
- User configurable Power-on Reset
- One external Non-Maskable Interrupt
- ST626x-EMU2 Emulation and Development System (connects to an MS-DOS PC via an RS232 serial line)

PDIP28

PSO28

CDIP28W

(See end of Datasheet for Ordering Information)

DEVICE SUMMARY

DEVICE	OTP (Bytes)	EPROM (Bytes)	I/O Pins
ST62T65B	3884	-	21
ST62E65B	-	3884	21

Fig. 12.2: First page of STMicroelectronics ST6265 datasheet.

Figure 12.2 shows the first page of the datasheet of this component, in which you can see the different packages that can be provided. In particular, the EPROM version with internal program memory needs a package with transparent window to allow the cancellation by exposure to UV rays.

Figure 12.3 shows a summary table of the microcontrollers PIC16C7x family from Microchip. You can see immediately how these cheap components, the

EPROM is not very large (up to 4K) and how the packages are small in size (up to 40 pins). The package can be found is both of DIL packages (dual in-line with pins through-hole to be threaded into the holes on the circuit board and being soldered on the opposite side of the board respect to the component) and SMD (Surface-Mounted Device with legs welded on the printed circuit board on which is the component without holes). For µC with more pins (over 44), there are also PLCC packages (Plastic Lead Chip-Carrier) and QFP (Quad Flat-Pack, a square with legs along the sides).

To limit the size of the package and the cost of the µC, many lines of I / O are shared on the same pin. For this reason you can not use all ports simultaneously (in the case of ST6 called PA, PB, PC), analog inputs (Ain) and auxiliary control devices (such as for serial communication: Sin, Sout and Sck). The designer have to choose what to use, setting off the µC firmware in such a way that to took into account if the pin is configured as input or output, digital or analog.

Often to reduce the cost of the cheaper µC, as well as to shrink the package (which represents a considerable cost), a reduction in the on chip supported memory, as well as in the number of internal peripherals as shown in the table of Fig 12.3. The µC of Figure 12.4 is the smallest of the family PIC6C7x, while Figure 12.5 shows that with the largest number of internal peripherals.

The log at the ALU output (from now on called Working Register, W) is usually known as the accumulator. There are two buses: one controlled by the Program Counter, which is used to access the instructions (which are then suitably ‘decoded’ and executed by an appropriate logic) and the other, the data bus that allows communication between RAM and peripherals (I/O, ADC, timers, etc.). In every µC family, next to the PC (Program Counter) there is the stack, which saves the contents of the PC if a subroutine call happens. Having a finite number of levels, you can not handle too many calls at once, because the former would be lost.

In order to undertake the planning of the µC there is usually a special pin (TEST) to be asserted prior to power the chip. In this way the µC is configured to be programmed instead of for the normal operation.

PIC µC instructions are 14 bits long words (compared to the usual 8-bit), so you can make op-code (operative-code, i.e. instructions) of a single *word*, instead of a double call to the memory program.

TABLE 1-1: PIC16C7X FAMILY OF DEVICES

	Clock: Maximum Frequency of Operation (MHz)	Memory: EPROM Program Memory (x14 words)	Memory: Data Memory (bytes)	Peripherals: Timer Module(s)	Peripherals: Capture/Compare/PWM Module(s)	Peripherals: Serial Port(s) (SPI/I²C, USART)	Peripherals: Parallel Slave Port	Peripherals: A/D Converter (8-bit) Channels	Peripherals: Interrupt Sources	Features: I/O Pins	Features: Voltage Range (Volts)	Features: In-Circuit Serial Programming	Features: Brown-out Reset	Features: Packages
PIC16C710	20	512	36	TMR0	—	—	—	4	4	13	3.0-6.0	Yes	Yes	18-pin DIP, SOIC; 20-pin SSOP
PIC16C71	20	1K	36	TMR0	—	—	—	4	4	13	3.0-6.0	Yes	—	18-pin DIP, SOIC
PIC16C711	20	1K	68	TMR0	—	—	—	4	4	13	3.0-6.0	Yes	Yes	18-pin DIP, SOIC; 20-pin SSOP
PIC16C72	20	2K	128	TMR0, TMR1, TMR2	1	SPI/I²C	—	5	8	22	2.5-6.0	Yes	Yes	28-pin SDIP, SOIC, SSOP
PIC16C73	20	4K	192	TMR0, TMR1, TMR2	2	SPI/I²C, USART	—	5	11	22	3.0-6.0	Yes	—	28-pin SDIP, SOIC
PIC16C73A(1)	20	4K	192	TMR0, TMR1, TMR2	2	SPI/I²C, USART	—	5	11	22	2.5-6.0	Yes	Yes	28-pin SDIP, SOIC
PIC16C74	20	4K	192	TMR0, TMR1, TMR2	2	SPI/I²C, USART	Yes	8	12	33	3.0-6.0	Yes	—	40-pin DIP; 44-pin PLCC, MQFP
PIC16C74A(1)	20	4K	192	TMR0, TMR1, TMR2	2	SPI/I²C, USART	Yes	8	12	33	2.5-6.0	Yes	Yes	40-pin DIP; 44-pin PLCC, MQFP, TQFP

All PIC16/17 Family devices have Power-on Reset, selectable Watchdog Timer, selectable code protect and high I/O current capability.
All PIC16C7X Family devices use serial programming with clock pin RB6 and data pin RB7.

Note 1: Please contact your local sales office for availability of these devices.

Fig. 12.3: Classification of internal pheripereals and performance divided by family.

Device	Program Memory	Data Memory (RAM)
PIC16C710	512 x 14	36 x 8
PIC16C71	1K x 14	36 x 8
PIC16C711	1K x 14	68 x 8

Fig. 12.4: Block diagram of PIC16C7xx. To note both I/O ports, internal ADC and timer.

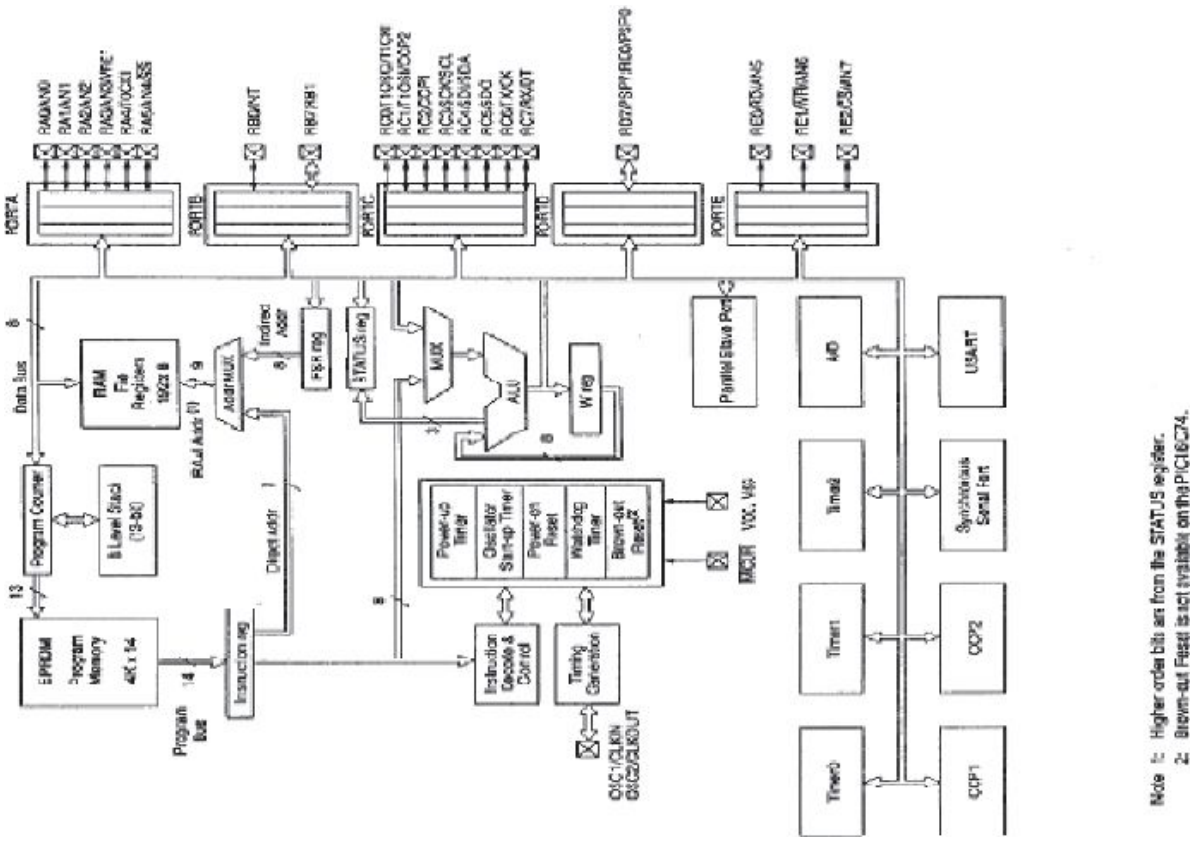


Fig. 12.5: In PIC16C74/74A you have four I/O ports, the ADC, three timers, two Capture and Compare units (CCP) and synchronous and asynchronous serial communication devices (USART).

12.2. µC DEVELOPMENT SYSTEMS

Development system

In order to familiarize with a particular µC, with the management of its internal devices and its own set of instructions, many manufacturers offer the designer a real development system, also called Demo Board, Starter Kit, Evaluation Board, Development System. These kits usually consist of a general purpose board with the µC, a numeric keypad, display, buttons, some digital inputs and outputs, if any analog input, and software that can develop the firmware and transfer it in the µC on the board. In this way, the designer can become familiar with the µC without having to design a prototype board, but can use the one already prepared and verified by the manufacturer. The project manager should devote himself exclusively to writing firmware more or less complex, able to handle all internal and additional external devices provided on the evaluation board.

Figure 12.6 shows the starter kit for STMicroelectronics' µC ST6xx family, complete with manual, adapter, parallel cable to connect to the computer and management and programming software. The cost of developing these systems can vary from one to several thousand euros, depending on the price of µC and the optional components on the board (display, OpAmp for conditioning signals from sensors, ADC, DAC or external buffer , ...).

Software

The designer has different software packages to easyer the programming of the µC and the writing of the code. It's properly this software that generates the real programming file (characterized by the real op-code in digital format) and translate the mnemonic codes of the listing, macroes and labels used in the program. The assembler is the software that allows you not to write directly in ASCII. Usually the program is written with mnemonic codes (add, goto, ...), as shown on the screen of Figure 12.7. You can often write directly in a routine high-level language (usually C). In the latter case, however, it should be noted that the translation from C to assembler language is not always optimized and, therefore, this option is feasible only for high-end µC. For those low-end average is preferable to become familiar with the assembler and implement the program directly with the menmonic codes.

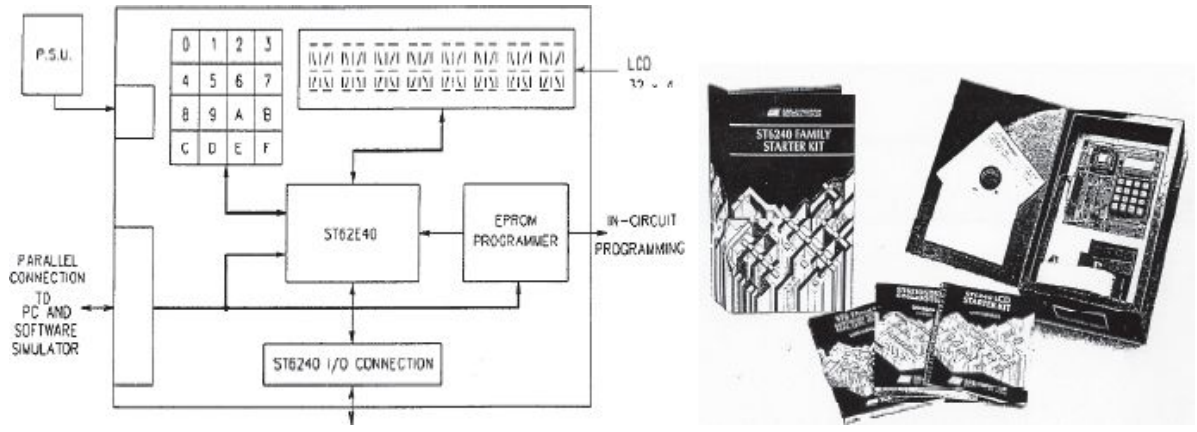


Fig. 12.6: Starter kit for ST6 family. In addition to the µC, note the presence of a numeric keypad, a eight characters display and the inputs/outputs available to the designer. The firmware is downloaded from your computer hrough parallel port.

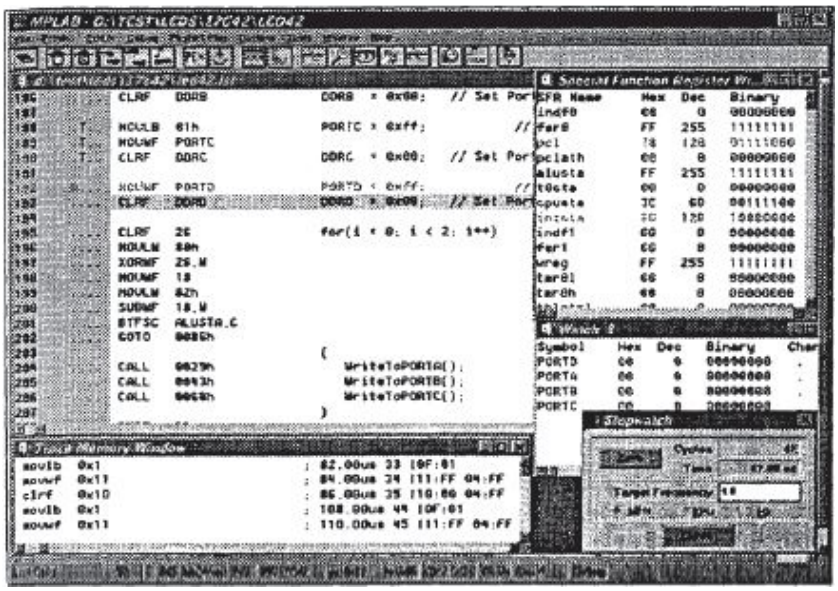


Fig. 12.7: Screenshot of an assembler/simulator program.

Finally, the software allows you to simulate a newly developed firmware (the hardware has not yet been realized). With this simulation, you can follow step by step execution of each instruction, control the flow of data between registers, monitor the bit of interest in the records “special function” and in the ports of input/output, to count the execution time and temporal durations of the cycles. This way you can debug the firmware, before to write it in the real μC and implement the hardware it will be inserted in.

On the web freeware and/or shareware versions of some assemblers, compilers, and translators for μC low-end can be found. On the other hand the most complex packages and the medium-high level μC, the cost can range from a few thausands to tens of thausands euros. The cost depends on the power of logic simulation, the ability to carry out simulations in real-time and, eventually, the ability to input excitement vectors (in the form of files) to digital and analog inputs, to simulate the real application.

Programmer

In order to put the firmware in the microcontroller or, in common speaking, to “burn the micro”, the programmer is used. It’s the same equipment that usually is able to program other digital devices such as EPROM, E^2PROM, Flash, PLD and FPGA programmable logic. The communication between computers, where the firmware is designed, and the programmer is usually done through RS232 serial port.

In case of industrial production of large volumes, it is convenient to switch to a gang-programmer, able to program in parallel, thus simultaneously, more chips, significantly reducing the total time. One of these products is shown in Figure 12.8.

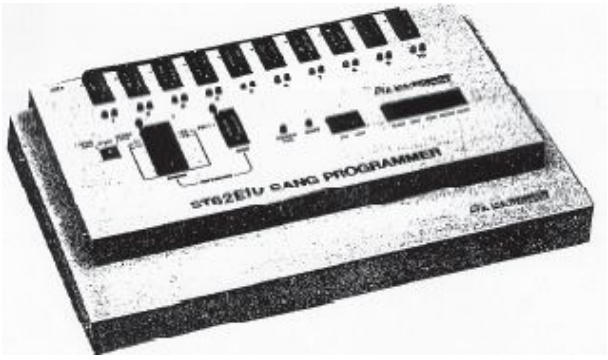


Fig. 12.8: Gang-programmer, for parallel programmino of more µC.

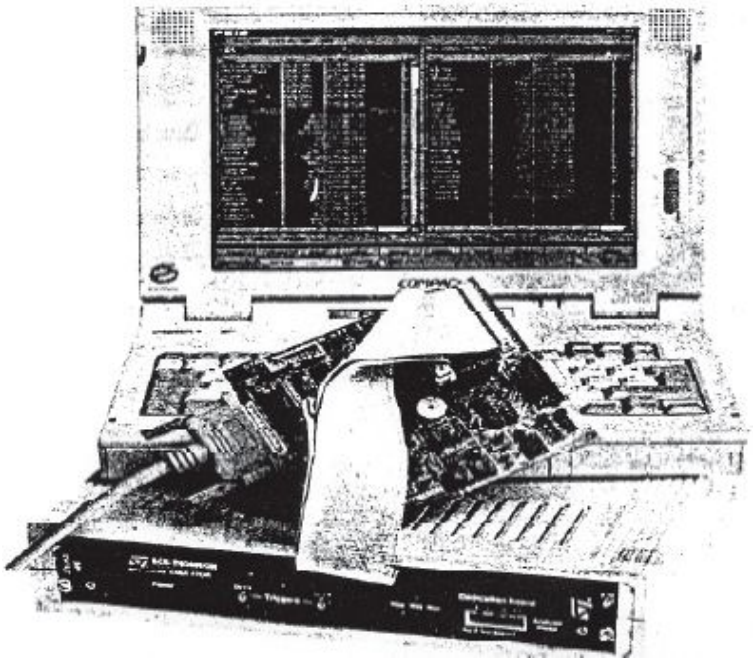


Fig. 12.9: µC emulator. To note the base that sostitute the real µC.

Emulator

Once programmed, the µC can be inserted in the prototype circuit and tested. Usually problems and malfunctions appers due to both the µC firmware and the board hardware. Debugging the entire system can become very difficult and expensive (in terms of time and cost), because it is often extremely difficult to

identify the cause of the malfunction. In all these cases, would be nice to “see” what is happening inside the µC, in order to monitor its registers, input/output ports, internal devices, to discover and solve the problem.

In this case the emulator comes to help, a hardware card times to times very complex that do the work of the microcontroller itself and can be inserted (through an appropriate adapter socket) instead of the µC during the prototyping board and verification of its functionality. A similar product is shown in Figure 12.9. The firmware is stored in the memory of the emulator, which will have at its pins the same behavior of the µC that it emulates. This practice can prevent the real µC to be programmed, but most important it makes possible to monitor the contents of all registers and devices internal to the emulated µC.

Some emulators are also equipped with analog and digital sensors able to capture and analize real time voltages or logic level on the board, in order to compare them with the expected valued during the stages of operations. This makes even easier to understand if the problem is in the firmware or in the hardware of the board. Their cost can range from a few thousands euros for the simpler emulators to dozens of thousands more for those with more accessories and for those able to emulate fast µC and high-end.

12.3. Program Memory

The program memory contains the firmware that the µC will execute. The main requirement that this memory must have is that it is not volatile. There are different types, briefly described below.

- ROM: it implies that it is “mask-programmed”, ie directly on the chip programmed by the manufacturer of µC, according to the list provided by the designer. This type of solution is suitable only for major productions (of the order of thousands of pieces).
- OTP: One-Time Programmable, it’s a memory programmable by the designer himself with just a PROM/µC programmer. This type of solution is convenient in the case of small productions, that is until a few hundred pieces, for more quantitative it would be too expensive in terms of time required for programming and for the operator costs. These memories are used after the complete setup of the program and then you are sure of its proper functioning. To cut costs, manufacturers often do not implement a true PROM in the µC, but they uses an EPROM without leaving the glass window on the package, thus making it impossible to delete and rewrite it.

- EPROM: erasable through UV, it requires a package including a window ultraviolet transparent for the cancellation, therefore, are more expensive than OTP but have the obvious advantage of being totally erased and reprogrammed. Their use is fundamental in the development and prototyping of the firmware or software setup.
- E^2PROM: are programmable and electrically erasable quickly (no need to use UV light as in the EPROM). Those are implemented in the more expensive μC, but for certain applications their use is advantageous in that it allows easy updating of firmware, even when the μC is mounted on the board (through the so-called In-System Programming).

To program μC we use the assembly language made up of instructions that manage the operation of the integrated. Since the assembler can be quite tricky and hard to develop, sometimes you prefer to use high-level languages (like C) that will be translated into assembly by a dedicated PC software (known as the compiler). In the following, ‘software’ will denote the high-level program while ‘firmware’ the low-level that manages the hardware inside and outside the μC.

The program memory is generally not very wide, from 1k to 4k or so, with words of 8 bits (ST) or 14bit (PIC). It’s usually inside the μC and doesn’t offer the possibility to be expanded outside. We will see that the 14-bit approach allows to write complex operations in a single memory location, even if it does not translate directly into a faster execution of instructions than the solution at 8 bits (which will often need two words memory).

The projects with μC are very simple in terms of hardware. In addition, the hardware is easily copied, it is enough to buy a circuit from a competitor and copy the PCB (major difficulties are encountered in the case of multi-layer PCB). It’s obvious that the originality of a project lies in the software. This inpluies that the part has to be protected is the code programmed into the firmware of the μC. The same way you can write it in the μC, you can also read the program memory of the μC. To avoid this, ie to prevent the reading of the stored program, the manufacturers provide the ability to burn a microfuse on-chip, preventing access to the final code from the outside. However, through reverse engineering it’s possible to chemically attack the package of μC and try to change this microfuse with a laser, or they could observe the Programmable ROM under the microscope directly.

As shown in Figure 12.10., not the entire program memory can be used by the designer. Often the manufacturers maps the memory in different areas; certain confidential (because of the management of peripheral or because of not

implemented in the chip), others containing the interrupt vectors explicitly and finally the remaining available to the user / designer. The first lines are used for the start of the μC: in the address 0000_H (Reset vector) lies the first instruction that runs when the μC is reseted, it will be a jump to the line of the first statement. There are other lines where interrupt vectors are stored inside the memory: each time an interrupt is called, the underway instruction is interrupted, the program counter is saved on the stack and a jump to the intervened interrupt memory cell is performed in order to to begin the execution of its routine.

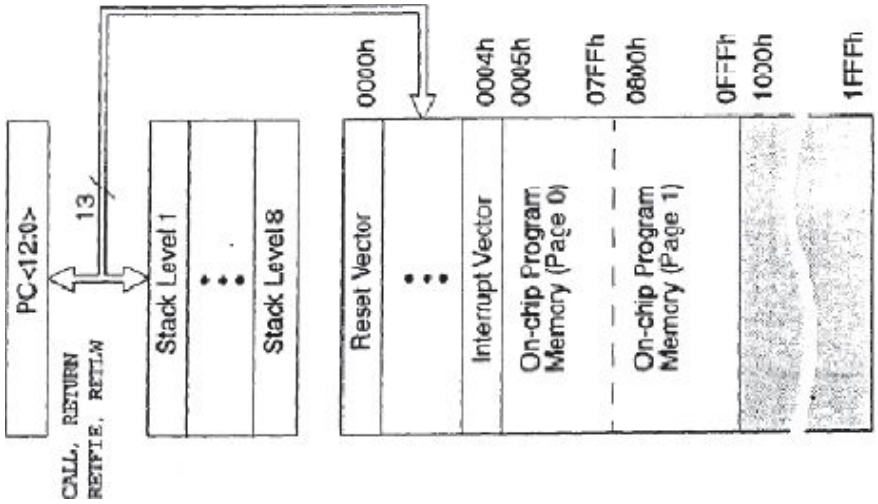


Fig. 12.10: Program Memory and Stack map of PIC16C73/74.

In the case of the PIC16C71 1k of program memory is available. To address this memory 10bit would be needed, that would rise to 13 in the case of 8k. To save money on a so large address bus, a memory mapped into pages is used. For example, Figure 12.10 shows the map to the PIC16C74A. The aim is to use fewer bits to access a page, so that the program counter can be shorter and jump instructions can occupy only one op-code.

In the case of the PIC memory is fully addressable with 13bit (the last cell is exactly equal to 1 $1FFF_H$ +4 +4 +4 = 13bit). Despite this, to access each page 11 bits are enough (that is equal to $07FF_H$ equal to 00’0111’1111’1111). The transition from page to page is done asserting the bits page (Page Select). If in future the manufacturer may (and will) integrate more memory on-chip, he will not change the program counter, but just add a bit to the page register, which will become two-bit.

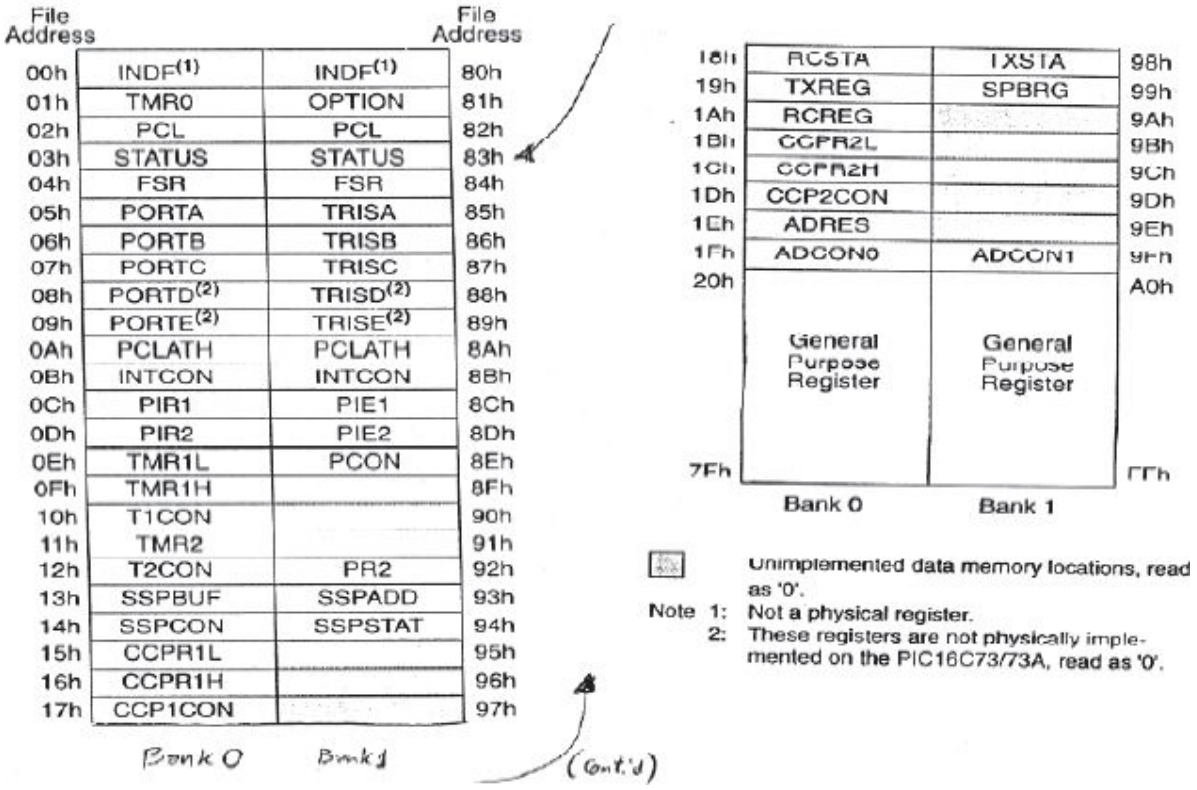


Fig. 12.11: Registry map of PIC16C73/74. To note both the division among two 128 byte data bank and the presence of dedicated "special function" registrers with theyr proper name and common one (general-purpose).

12.4. Data Memory

This memory is usually used to hold the temporary results of calculations performed by the µC, or numeric values acquired from the input ports, the ADC or internal timer. Usually in the low-end µC the data memory is small, ie a few tens or hundreds of bytes. Only in the high-end µC RAM you can get up to a few kbytes. If the memory is dynamic the µC will ensure a refresh, and this requires to work at a minimum clock frequency.

On the other hand, by implementing a static RAM, the operation of the µC will be also allowed at low clock rates (a few hundred kHz) or even continuous (one-step operation, also known as continuous or DC). The disadvantage of static RAM is, however, the increased size of the memory cell, which then leads to a lower memory to area occupied on silicon ratio (silicon estate).

In case you wish to place the µC in a "sleep" state in order to consume less power, it would be impossible to keep alive the circuitry necessary to refresh the

dynamic RAM. The ideal would therefore have, in addition to the dynamic RAM, an EEPROM memory for data less frequently to update, but not to lose, so it can save and keep the data even after the μC power off.

In addition to save space, the internal architecture of the data memory is organized at banks or register file, in order to realize the address with fewer bits and have compact op-codes. Figure 12.11 shows the typical data memory of the PIC. To switch from one bank to another simply set a bit to 0 or 1 in a special register. For example, in PIC16C7X this bit is the fifth content from the STATUS register and is called RP0: setting it to 0 or to 1, the μC will enter bank 0 or 1.

Memory addresses go from 00_H to FF_H but just 96 bytes of memory for each data bank are available to the user. In the case you try to read a cell that is not (for example, directing the Bank 3), the program may cause errors, because it would read nonsense (as the cell in question is seen in high impedance) or writing to nothing an important data, losing it.

Proceeding as just seen, each memory location within a bank can be addressed with only 7 bits. These bits can be part of the same op-code and be treated like privileged registers. This type of architecture is called a register file and allows you to target cells in RAM by their “name” (7 bit code), as were additional accumulator registers.

Only the working register W (in the PIC) is physically implemented within the ALU, it contains the results of arithmetic operations (SUM, SUB), logical (AND, OR, NOT) and shift. The other operand however, can be any register of the RAM (f stands for file register) or a constant passed immediately to the instruction.

For example, the sum of W and a generic register f uses only one op-code: ADDWF f, d op-code = 2000 0111 ffff dfff where the 7-bit address f specifies the location of data in RAM, and d indicates if to save the result in the register W (d=0) or f (d=1). The instruction: ADDWF 45h, 0;
sums W to the contents of the register 45_H and the result is stored in W. With:

ADDLW K op-code = 11 111x kkkk kkkk
the eight-bit constant K is added to the working register W. The bit x means “do not care”, so can take either the value 1 or 0.

It is evident that with a single 14 bit instruction complex codes can be described. Also, transactions between more memory from different cells are simplified by using the register-file architecture. For example, if you want to sum the value operandoA to operandoB it’s not needed to to go through the

accumulator, with the instructions: LDA operandoA ADD operandoB STA operandoB
But just the operations:

MOVFW operandoA ADDWF operandoB,1

In RAM special registers are also implemented, whose presence is necessary to control devices inside the µC, interrupts, data read and written from I/O ports, the contents of the timer, etc.. as shown in Figure 12.12. By setting the value 1 or 0 in the corresponding bit, you change the internal paths of signals in the µC or the internal structure of each device on chip. For this reason, the meaning of each bit must be perfectly clear to the designer that will have to know how to use, read or set any ot then at an appropriate time within the firmware. The other locations are intended for general-purpose registers available to the designer. For example, the register of state (Fig. 12.13) contains the flag bits, such as carry, overflow, zero, all alleged after each operation involving the working register W of the ALU. Referring again to Figure 12.11., we see that the status register is present in both the register bank, otherwise it would be impossible to change bit 5 to change bank. The Option Register (Fig. 12.14), however, manage some functions of timers, interrupts and pull-up door PB configuration.

12.5. Assembler Instructions

At this point it is convenient to see the instructions available in both the µC considered so far. As we shall see, it's a very simple instruction set, but more than enough to control applications that are intended for those kind of µCs.

PIC16Cxx has only 35 instructions of 14 bits each: one part of op-code defines the type of education, while the remainder contains the operands. The op-code fields are marked with letters, to clarify its origin. Figure 12.15 summarizes all the Microchip assembler instructions. For example, the ADDWF instruction adds the contents of W with that in the register of the RAM with address F specified by the user through the 7 bits (in the bank where you are). The result is stored in W or at F based on the value assumed by bit d. Other processors, such as ST6, use an op-code of two words of 8 bits each: the first byte specify the type of operation, while the latter specify the data.

12.5.1 Data address

To address a certain location in RAM you can specify the address within the op-code (eg with the 7-bit fff fffff seen in the examples above): this is called direct addressing, and is shown in Fig 12:17 left. However, if you were to scan an area of RAM without knowing a priori the extension or the point of departure or arrival, it would be usefull to have a different way to address, called indirect.

In µC there are registers for this purpose. For example, the PIC has a special register called FSR (File Select Register), which can be written, read, increased or decreased at will, and its memory address is 04_H, regardless of the Bank. Figure 12.17., right, shows its use. To access the memory location pointed to by this register, simply refer to the fictitious register INDF. Calling it, the µC will use the contents of FSR to access the location in RAM.

In the PIC, however, there isn't the statement '*move the data in* FSR', so you should always use the W register. More instructions are then needed to carry out quite complex operations, each of which is performed in an instruction cycle, corresponding to 4 clock periods. Using the ST6 microcontroller, you will note the presence of the '*move the data in* FSR' instruction. However, since the op-code requires 2 bytes for its implementation, the time is the same also for this device, using the same clock frequency.

PIC16C73/73A/74/74A SPECIAL FUNCTION REGISTER SUMMARY

Address	Name	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0	Value on: POR, BOR	Value on all other resets (2)
Bank 0											
00h(4)	INDF	Addressing this location uses contents of FSR to address data memory (not a physical register)								0000 0000	0000 0000
01h	TMR0	Timer0 module's register								xxxx xxxx	uuuu uuuu
02h(4)	PCL	Program Counter's (PC) Least Significant Byte								0000 0000	0000 0000
03h(4)	STATUS	IRP(7)	RP1(7)	RP0	TO	PD	Z	DC	C	0001 1xxx	000q quuu
04h(4)	FSR	Indirect data memory address pointer								xxxx xxxx	uuuu uuuu
05h	PORTA	—	—	PORTA Data Latch when written: PORTA pins when read						--0x 0000	--0u 0000
06h	PORTB	PORTB Data Latch when written: PORTB pins when read								xxxx xxxx	uuuu uuuu
07h	PORTC	PORTC Data Latch when written: PORTC pins when read								xxxx xxxx	uuuu uuuu
08h(5)	PORTD	PORTD Data Latch when written: PORTD pins when read								xxxx xxxx	uuuu uuuu
09h(5)	PORTE	—	—	—	—	—	RE2	RE1	RE0	---- -xxx	---- -uuu
0Ah(1,4)	PCLATH	—	—	—	Write Buffer for the upper 5 bits of the Program Counter					---0 0000	---0 0000
0Bh(4)	INTCON	GIE	PEIE	TOIE	INTE	RBIE	TOIF	INTF	RBIF	0000 000x	0000 000u
0Ch	PIR1	PSPIF(3)	ADIF	RCIF	TXIF	SSPIF	CCP1IF	TMR2IF	TMR1IF	0000 0000	0000 0000
0Dh	PIR2	—	—	—	—	—	—	—	CCP2IF	---- ---0	---- ---0
0Eh	TMR1L	Holding register for the Least Significant Byte of the 16-bit TMR1 register								xxxx xxxx	uuuu uuuu
0Fh	TMR1H	Holding register for the Most Significant Byte of the 16-bit TMR1 register								xxxx xxxx	uuuu uuuu
10h	T1CON	—	—	T1CKPS1	T1CKPS0	T1OSCEN	T1SYNC	TMR1CS	TMR1ON	--00 0000	--uu uuuu
11h	TMR2	Timer2 module's register								0000 0000	0000 0000
12h	T2CON	—	TOUTPS3	TOUTPS2	TOUTPS1	TOUTPS0	TMR2ON	T2CKPS1	T2CKPS0	-000 0000	-000 0000
13h	SSPBUF	Synchronous Serial Port Receive Buffer/Transmit Register								xxxx xxxx	uuuu uuuu
14h	SSPCON	WCOL	SSPOV	SSPEN	CKP	SSPM3	SSPM2	SSPM1	SSPM0	0000 0000	0000 0000
15h	CCPR1L	Capture/Compare/PWM Register1 (LSB)								xxxx xxxx	uuuu uuuu
16h	CCPR1H	Capture/Compare/PWM Register1 (MSB)								xxxx xxxx	uuuu uuuu
17h	CCP1CON	—	—	CCP1X	CCP1Y	CCP1M3	CCP1M2	CCP1M1	CCP1M0	--00 0000	--00 0000
18h	RCSTA	SPEN	RX9	SREN	CREN	—	FERR	OERR	RX9D	0000 -00x	0000 -00x
19h	TXREG	USART Transmit Data Register								0000 0000	0000 0000
1Ah	RCREG	USART Receive Data Register								0000 0000	0000 0000
1Bh	CCPR2L	Capture/Compare/PWM Register2 (LSB)								xxxx xxxx	uuuu uuuu
1Ch	CCPR2H	Capture/Compare/PWM Register2 (MSB)								xxxx xxxx	uuuu uuuu
1Dh	CCP2CON	—	—	CCP2X	CCP2Y	CCP2M3	CCP2M2	CCP2M1	CCP2M0	--00 0000	--00 0000
1Eh	ADRES	A/D Result Register								xxxx xxxx	uuuu uuuu
1Fh	ADCON0	ADCS1	ADCS0	CHS2	CHS1	CHS0	GO/DONE	—	ADON	0000 00-0	0000 00-0

Legend: x = unknown, u = unchanged, q = value depends on condition, - = unimplemented read as '0'.
Shaded locations are unimplemented, read as '0'.

Note 1: The upper byte of the program counter is not directly accessible. PCLATH is a holding register for the PC<12:8> whose contents are transferred to the upper byte of the program counter.

2: Other (non power-up) resets include external reset through $\overline{MCLR}$ and Watchdog Timer Reset.

3: Bits PSPIE and PSPIF are reserved on the PIC16C73/73A, always maintain these bits clear.

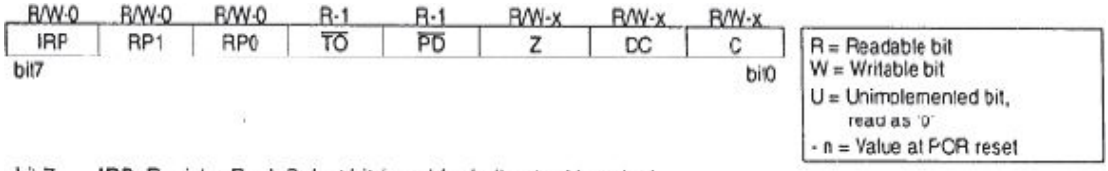
4: These registers can be addressed from either bank.

5: PORTD and PORTE are not physically implemented on the PIC16C73/73A, read as '0'.

6: Brown-out Reset is not implemented on the PIC16C73 or the PIC16C74, read as '0'.

7: The IRP and RP1 bits are reserved on the PIC16C7X, always maintain these bits clear.

Fig. 12.12: Short explanation of come "special funcion" registers for PIC16C73/74.



bit 7: **IRP**: Register Bank Select bit (used for indirect addressing)
1 = Bank 2, 3 (100h - 1FFh)
0 = Bank 0, 1 (00h - FFh)
The IRP bit is reserved on the PIC16C7X, always maintain this bit clear.

bit 6-5: **RP1:RP0**: Register Bank Select bits (used for direct addressing)
11 = Bank 3 (180h - 1FFh)
10 = Bank 2 (100h - 17Fh)
01 = Bank 1 (80h - FFh)
00 = Bank 0 (00h - 7Fh)
Each bank is 128 bytes. The RP1 bit is reserved on the PIC16C7X, always maintain this bit clear.

bit 4: $\overline{TO}$: Time-out bit
1 = After power-up, CLRWDT instruction, or SLEEP instruction
0 = A WDT time-out occurred

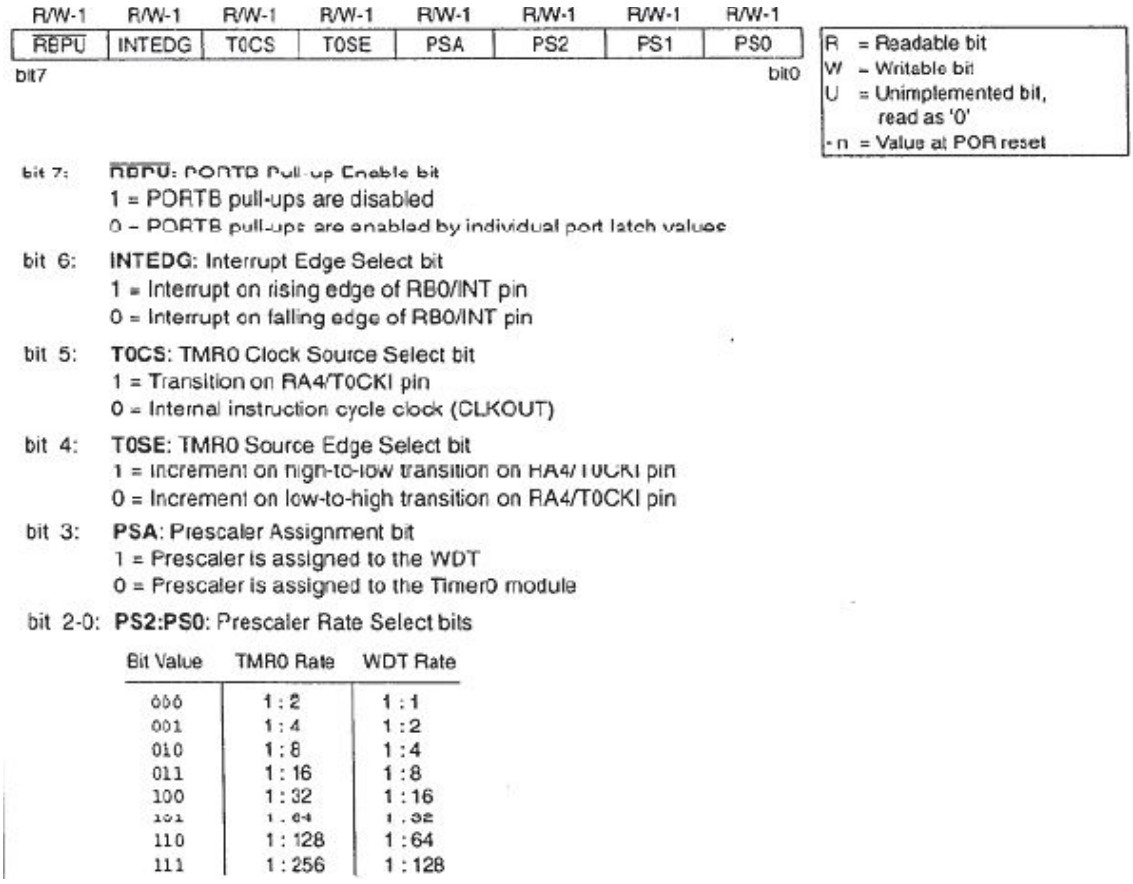
bit 3: $\overline{PD}$: Power-down bit
1 = After power-up or by the CLRWDT instruction
0 = By execution of the SLEEP instruction

bit 2: **Z**: Zero bit
1 = The result of an arithmetic or logic operation is zero
0 = The result of an arithmetic or logic operation is not zero

bit 1: **DC**: Digit carry/$\overline{borrow}$ bit (ADDWF, ADDLW, SUBLW, SUBWF instructions)(for $\overline{borrow}$ the polarity is reversed)
1 = A carry-out from the 4th low order bit of the result occurred
0 = No carry-out from the 4th low order bit of the result

bit 0: **C**: Carry/$\overline{borrow}$ bit (ADDWF, ADDLW, SUBLW, SUBWF instructions)
1 = A carry-out from the most significant bit of the result occurred
0 = No carry-out from the most significant bit of the result occurred
Note: For $\overline{borrow}$ the polarity is reversed. A subtraction is executed by adding the two's complement of the second operand. For rotate (RRF, RLF) instructions, this bit is loaded with either the high or low order bit of the source register.

Fig. 12.13: Contento f Status register (address 03_H and 83_H) for PIC16C73/74.



R/W-1	R/W-1	R/W-1	R/W-1	R/W-1	R/W-1	R/W-1	R/W-1
$\overline{\text{RBPU}}$	INTEDG	T0CS	T0SE	PSA	PS2	PS1	PS0
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
-n = Value at POR reset

bit 7: $\overline{\text{RBPU}}$: PORTB Pull-up Enable bit
1 = PORTB pull-ups are disabled
0 = PORTB pull-ups are enabled by individual port latch values

bit 6: **INTEDG**: Interrupt Edge Select bit
1 = Interrupt on rising edge of RB0/INT pin
0 = Interrupt on falling edge of RB0/INT pin

bit 5: **T0CS**: TMR0 Clock Source Select bit
1 = Transition on RA4/T0CKI pin
0 = Internal instruction cycle clock (CLKOUT)

bit 4: **T0SE**: TMR0 Source Edge Select bit
1 = Increment on high-to-low transition on RA4/T0CKI pin
0 = Increment on low-to-high transition on RA4/T0CKI pin

bit 3: **PSA**: Prescaler Assignment bit
1 = Prescaler is assigned to the WDT
0 = Prescaler is assigned to the Timer0 module

bit 2-0: **PS2:PS0**: Prescaler Rate Select bits

Bit Value	TMR0 Rate	WDT Rate
000	1 : 2	1 : 1
001	1 : 4	1 : 2
010	1 : 8	1 : 4
011	1 : 16	1 : 8
100	1 : 32	1 : 16
101	1 : 64	1 : 32
110	1 : 128	1 : 64
111	1 : 256	1 : 128

Fig. 12.14: Content od Option register (address 81_H) for PIC16C73/74.

INSTRUCTION SET

Mnemonic, Operands		Description	Cycles	14-Bit Opcode MSb			LSb	Status Affected	Notes
BYTE-ORIENTED FILE REGISTER OPERATIONS									
ADDWF	f, d	Add W and f	1	00	0111	dfff	ffff	C,DC,Z	1,2
ANDWF	f, d	AND W with f	1	00	0101	dfff	ffff	Z	1,2
CLRF	f	Clear f	1	00	0001	lfff	ffff	Z	2
CLRW	-	Clear W	1	00	0001	0000	0011	Z	
COMF	f, d	Complement f	1	00	1001	dfff	ffff	Z	1,2
DECF	f, d	Decrement f	1	00	0011	dfff	ffff	Z	1,2
DECFSZ	f, d	Decrement f, Skip if 0	1(2)	00	1011	dfff	ffff		1,2,3
INCF	f, d	Increment f	1	00	1010	dfff	ffff	Z	1,2
INCFSZ	f, d	Increment f, Skip if 0	1(2)	00	1111	dfff	ffff		1,2,3
IORWF	f, d	Inclusive OR W with f	1	00	0100	dfff	ffff	Z	1,2
MOVF	f, d	Move f	1	00	1000	dfff	ffff	Z	1,2
MOVWF	f	Move W to f	1	00	0000	lfff	ffff		
NOP	-	No Operation	1	00	0000	0xx0	0000		
RLF	f, d	Rotate Left f through Carry	1	00	1101	dfff	ffff	C	1,2
RRF	f, d	Rotate Right f through Carry	1	00	1100	dfff	ffff	C	1,2
SUBWF	f, d	Subtract W from f	1	00	0010	dfff	ffff	C,DC,Z	1,2
SWAPF	f, d	Swap nibbles in f	1	00	1110	dfff	ffff		1,2
XORWF	f, d	Exclusive OR W with f	1	00	0110	dfff	ffff	Z	1,2
BIT-ORIENTED FILE REGISTER OPERATIONS									
BCF	f, b	Bit Clear f	1	01	00bb	bfff	ffff		1,2
BSF	f, b	Bit Set f	1	01	01bb	bfff	ffff		1,2
BTFSC	f, b	Bit Test f, Skip if Clear	1 (2)	01	10bb	bfff	ffff		3
BTFSS	f, b	Bit Test f, Skip if Set	1 (2)	01	11bb	bfff	ffff		3
LITERAL AND CONTROL OPERATIONS									
ADDLW	k	Add literal and W	1	11	111x	kkkk	kkkk	C,DC,Z	
ANDLW	k	AND literal with W	1	11	1001	kkkk	kkkk	Z	
CALL	k	Call subroutine	2	10	0kkk	kkkk	kkkk		
CLRWDT	-	Clear Watchdog Timer	1	00	0000	0110	0100	$\overline{TO}$,$\overline{PD}$	
GOTO	k	Go to address	2	10	1kkk	kkkk	kkkk		
IORLW	k	Inclusive OR literal with W	1	11	1000	kkkk	kkkk	Z	
MOVLW	k	Move literal to W	1	11	00xx	kkkk	kkkk		
RETFIE	-	Return from interrupt	2	00	0000	0000	1001		
RETLW	k	Return with literal in W	2	11	01xx	kkkk	kkkk		
RETURN	-	Return from Subroutine	2	00	0000	0000	1000		
SLEEP	-	Go into standby mode	1	00	0000	0110	0011	$\overline{TO}$,$\overline{PD}$	
SUBLW	k	Subtract W from literal	1	11	110x	kkkk	kkkk	C,DC,Z	
XORLW	k	Exclusive OR literal with W	1	11	1010	kkkk	kkkk	Z	

Note 1: When an I/O register is modified as a function of itself (e.g., MOVF PORTB, 1), the value used will be that value present on the pins themselves. For example, if the data latch is '1' for a pin configured as input and is driven low by an external device, the data will be written back with a '0'.

2: If this instruction is executed on the TMR0 register (and, where applicable, d = 1), the prescaler will be cleared if assigned to the Timer0 Module.

3: If Program Counter (PC) is modified or a conditional test is true, the instruction requires two cycles. The second cycle is executed as a NOP.

Fig. 12.15: Instruction Set for PIC16Cxx family.

Going back to PICs, as a simple example, consider a program that will clear the RAM data memory locations between 20_H and $2F_H$ using the indirect addressing mode.

```
;INDIRECT ADDRESSING
```

```
    Movlw 0x20      ; initialize pointer
    Movwf FSR       ; to RAM
    NEXT clrf INDF  ; clear INDF register
    incf FSR,F      ; inc pointer
    btfss FSR,4     ; all done?
    goto NEXT       ; no - clear next
    CONTINUE        ; yes – continue
```

The loop ends when FSR=30_H=0011 0000, thus when the 5th bit switch to 1.

ST62 & ST63 INSTRUCTION SET

The ST62,63 instructions can be divided functionaly into the following seven groups.

- LOAD AND STORE
- ARITHMETIC AND LOGIC
- CONDITIONAL BRANCH
- JUMP AND CALL
- BIT MANIPULATION
- CONTROL
- IMPLIED

The following summary shows the instructions belonging to each group, the number of operands required for each instructions and the number of machine cycles. The flag behaviour is usually the same for both ST62 and ST63. The only difference is present for CP and SUB instructions as specified in the detailed description.

Note: For the following tables

Δ: Affected

*: Not Affected

Table 2. Load & Store Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
LD	1	4	Δ	*
LD r	2	4	Δ	*
LDI A	2	4	Δ	*
LDI	3	4	*	*

Table 3. Arithmetic & Logic Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
ADD	2	4	Δ	Δ
ADD (X,Y)	1	4	Δ	Δ
ADDI	2	4	Δ	Δ
AND	2	4	Δ	*
AND (X,Y)	1	4	Δ	*
ANDI	2	4	Δ	*
CLR A	2	4	Δ	Δ
CLR	3	4	*	*
COM	1	4	Δ	Δ
CP	2	4	Δ	Δ
CP (X,Y)	1	4	Δ	Δ
CPI	2	4	Δ	Δ
DEC	1	4	Δ	*
DEC A/r	2	4	Δ	*
INC	1	4	Δ	*
INC A/r	2	4	Δ	*
RLC	1	4	Δ	Δ
SLA	2	4	Δ	Δ
SUB	2	4	Δ	Δ
SUB (X,Y)	1	4	Δ	Δ
SUBI	2	4	Δ	Δ

Table 4. Conditional Branch Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
JRC	1	2	*	*
JRNC	1	2	*	*
JRR	3	5	*	Δ
JRS	3	5	*	Δ
JRZ	1	2	*	*
JRNZ	1	2	*	*

Table 5. Jump & Call Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
CALL	2	4	*	*
JP	2	4	*	*

Table 6. Bit Manipulation Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
RES	2	4	*	*
SET	2	4	*	*

Table 7. Control Instructions

Instruction	Bytes	Cycles	Flags	
			Z	C
NOP	1	2	*	*
RET	1	2	*	*
RETI	1	2	Δ	Δ
STOP	1	2	*	*
WAIT	1	2	*	*

Table 8. Addressing Modes/Instruction Table

Instruction	Inh	Dir	Sh Dir	Ind	Imm	PCR	Ext	BitDir	Bit Test	Flags	
										Z	C
ADD		X	X	X						Δ	Δ
AND		X	X	X						Δ	*
CALL							X			*	*
CLR A		X								Δ	Δ
CLR		X								*	*
COM	X									Δ	Δ
CP		X		X	X					Δ	Δ
DEC		X	X	X						Δ	*
INC		X	X	X						Δ	*
JP							X			*	*
JRC, JRNC						X				*	*
JRZ, JRNZ						X				*	*
JRR, JRS									X	*	Δ
LD, LDI					X					Δ	*
NOP						X				*	*
RES, SET								X		*	*
RET	X									*	*
RETI	X									Δ	Δ
RLC	X									Δ	Δ
SLA	X									Δ	Δ
STOP, WAIT	X									*	*
SUB		X		X	X					Δ	Δ

Notes:

INH. Inherent, DIR: Direct, Sh.DIR: Short Direct,

IND. Indirect, IMM: Immediate, PCR: Program Counter Reative

EXT. Exteeded, BIT DIR: Bit Direct, BIT TEST.: Bit Test

Δ. Affected

*. Not Affected

Fig. 12.16: Instructions of ST6, grouped by type.

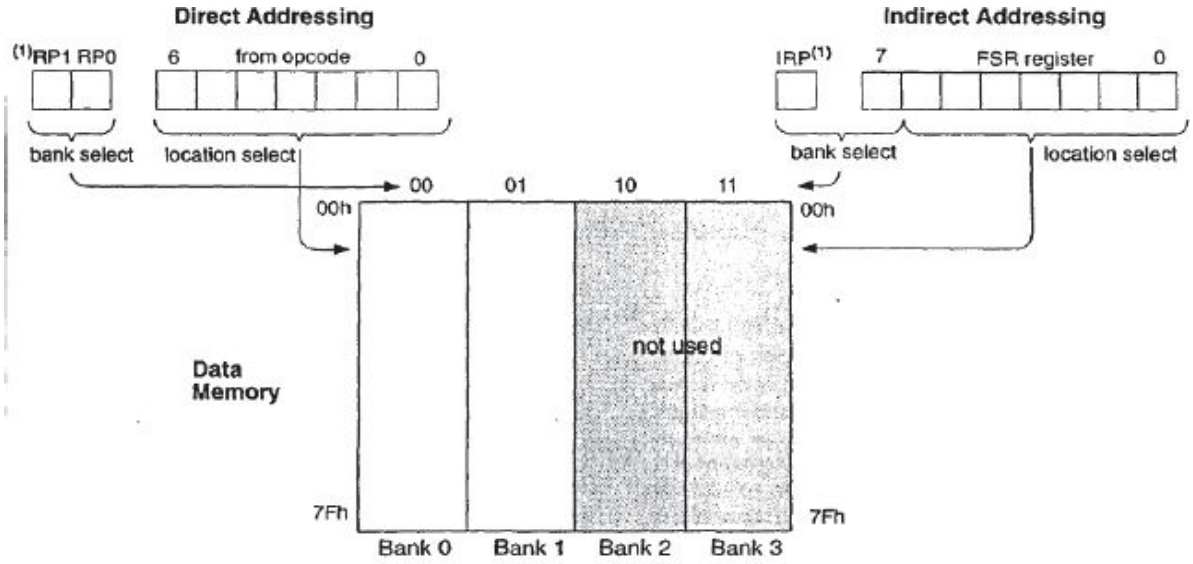


Fig. 12.17: Direct address (left) and Indirect (right) for a PIC.

On the other hand. Using the direct addressing we would have to use:

```
clrf 20h ; clear RAM at address 20h
clrf 21h ; clear RAM at address 21h
clrf 22h ; clear RAM at address 22h
clrf 23h ; clear RAM at address 23h
clrf 2Fh ; clear RAM at address 24h
```

In some μC there are also other ways to address the RAM, for example in ST6xx you have the following:

- INHERENT: the op-code contains all the information data (identical to the mode of the PIC, which specified the address ff ffff of the register file ffff ff);
- DIRECT: the address of the data in RAM follows following the op-code; the instruction is thus formed by two byte (eg LDA, 0A3h);
- SHORT DIRECT: it's possible to call 4 special registers X, Y, V and U using just one op-code byte (eg LDA, X);
- INDIRECT: 2 registers are used X or Y that contains the address of the ram cell to point. It's once again 1 byte long (eg LDA, (X)); • IMMEDIATE: Data is given just after the op-code and the RAM address in which we want to store. Lenghts goes from 2 to 3 bytes (eg LDI 34h, DFh).

ST6 have also 2 ways to jump: a RELATIVE PROGRAM COUNTER (eg JRC 3:if carry is 1 then PC=PC+3) and an EXTENDED one (eg JP $3FA_H$, thus PC=$3FA_H$).

STMicroelectronics microcontrollers have therefore address types much more extensive than those of Microchip, but these often require the use of multiple bytes. This fact shows that the comparison between different μC should not be restricted to the number of instructions provided, but also the length of the op-code, the number of memory accesses programs and the duration of the instruction cycle.

12.5.2 Instruction cycle

μC operations are timed by a clock. It is a microprogrammed machine and, ultimately, sequence. It is not true that for each clock period an entire operation is completed. In fact, the instruction cycle is executed in multiple clock periods. For example, in the PIC, in each period of the clock is carried out only one phase of an operation. As shown in Figure 12.18., in the first four clock periods, the μC goes to get (fetch) the next instruction addressed by the program counter, over the next four, decodes the instruction, which includes what type of command is and which operands are needed, and sending them directly to ALU or the correct device. Thus an instruction cycle (indicated by Tcy) consists of four clock periods and an instruction requires two instruction cycles, as shown in Figure 12.19. Just to exploit the parallelism of the current fetched education, to execute with the previous statement, it is convenient to use the pipeline architecture type. Thus in Tcy1 the statement loaded in the previous cycle is executed, but the μC proceeds also to read the new instruction (Fetch 2). That's the advantage of the dual bus structure: the core communicates with ROM and RAM through two distinct bus, in order to simultaneously access the program memory (for fetch) and the data (to obtain the operands).

In this way the throughput of the μC is doubled, in fact, despite an instruction usually takes two cycles to complete (8 clock periods), the number of instructions executed per second is equal to the inverse of Tcy (equal to 4 times clock). In the case of PIC, using the maximum clock frequency of 20MHz, it has a throughput of one instruction every 200ns, despite the execution time of 400ns.

In the flow of Figure 12.19 there's a thing to note: the subroutine called CALL SUB_1 is read in second instruction cycle (Tcy2). Therefore, the acquisition Fetch 4 is useless: the instruction 4 will not run (you have a Flush

instrad of Execute 4) and in Tcy4 you make the jump by reading the instruction at SUB_1. In this case, the actual instructions are four, but the cycles spent are five (not counting the Tcy0 fetch), because one went empty.

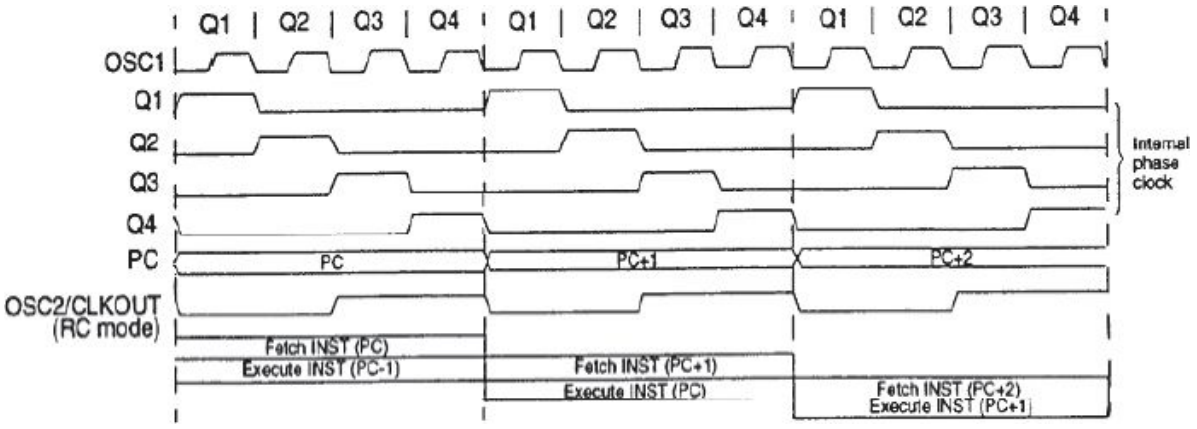


Fig. 12.18: Tcy instruction cycle, consistent in four clock periods. In the first cycle fetch is executed, thus the reading of the instruction to be executed, in the second cycle the proper execution.

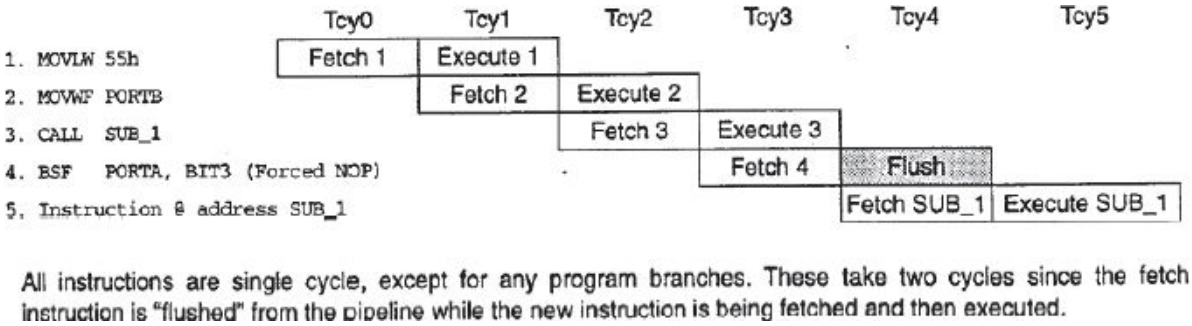


Fig. 12.19: Execution instruction flow. Note the pipeline at two levels, fetch at first (four clock periods) and execute at second (the other four clock periods).

12.6. I/O PORTS

I/O ports are the simplest example of the internal interface, which allows to come into contact with the outside world. The single pin I/O can be an input (I) or output (O) of the µC depending on how you set the port itself and therefore can be configured and used in different ways, summarized as follows:

- Digital Input: can be a normal CMOS input or a Schmidt trigger (which removes the uncertainty in the switching and increases the

Noise Margin). It's possible to have an internal pull-up (to save the external resistance), and leaving the pin floating.

- Input with Interrupt: as well as a simple data input, it can also be enabled to generate an interrupt in the µC.
- Analog Input: allows you to acquire analog signals and send them to the analog to digital converter (ADC, see Chapter 8) inside the µC, for further processing.
- Digital Output: Usually at CMOS levels, sometimes open drain (with only the output pull-down n-MOS, to enable wired connections, wired-and); it can have an internal pull-up resistor and can be three-state (ie, able to go high impedance and stay floating).

Figure 12.20 summarizes some possible settings for pin I/O in ST62xx microcontroller. Although architectural solutions chosen by the various µC manufactures may be slightly different one from another, this appears to be to more than adequate to describe what is on the market.

12.6.1 Configtion Registers

In order to configure each I/O, there is a rather complex structure for each pin to be set through firmware, as shown in Figure 12.21 for ST62xx. Each I/O port (A, B, C) has three 8 bits registers, one bit for each port line. These registers are the Data Register (contains the value read from or to be written to the corresponding port), the Direction Register (specify if the line is an input or output) and the Option Register (if the line is input, specify whether interrupt is enabled or not, if the line is output, specifying whether it is open-drain or normal).

The I/O digital port is simply summarized as shown in Figure 12.22. If the input selection is set to 1, the MOS are off, regardless of the value of output, the output is high impedance, so the pin is configured as an input. Conversely, if the input selection is kept low, the ports are transparent to the output value (denied) applied to the port and acts as an output pin. You may notice the presence of the two protection diodes.

In Microchip µCs (Fig. 12.23), the choice of how to use the I/O pin is done through the Special Function Register called TRIS: if the pin is 1 forces the pin three-state (so it can act as input), while to 0 pours the contents of the latch to the output (so the pin acts as output). The selection value is stored in the register by the WR TRIS signal from the µC core. Each 8-bit register TRIS bit will each connected to its corresponding port pin (TRISA, TRISB and so on). The bit

read from the port will be written into its corresponding PORT register bits. So read the port means to acquire the current state of the pin, while writing on the port means to write on the latch of the same port (Data Latch).

Note that at the μC startup, TRIS registers are preloaded with 1, configuring all pins as inputs, in order to avoid a short circuit outside of the μC. As an example, consider the initialization of the port A of the PIC:

```
EXAMPLE          INITIALIZING PORTA
CLRF PORTA       ; Initialize PORTA by clearing output data latches
BSF STATUS, RPO  ; Select Bank 1
MOVLW 0xCF       ; value used to initialize data direction 11001111
MOVWF TRISA      ; Set RA< 3 : 0 > as inputs RA< 5 : 4 > as outputs
                 ; TRISA< 7 : 6 > are always read as ‘0’
```

All operations performed on the I/O ports are of a Read-Modify-Write type: that is, in order to change a single bit output, without losing the values on other bits of the port, the core has to read the current situation of the port, modify the chosen bit and then pour the new output value.

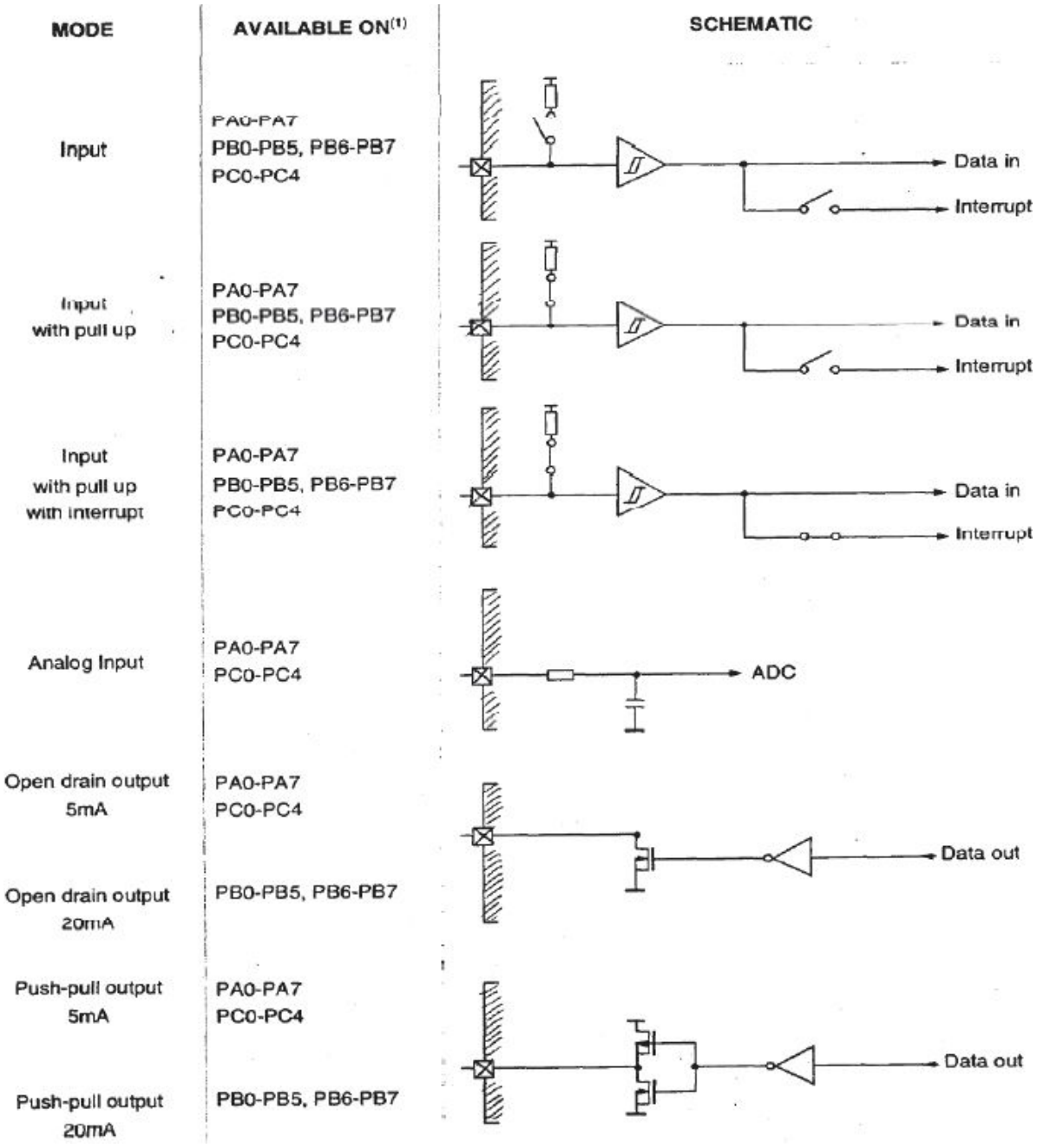


Fig. 12.20: Possible setup of ST62xx I/O pins.

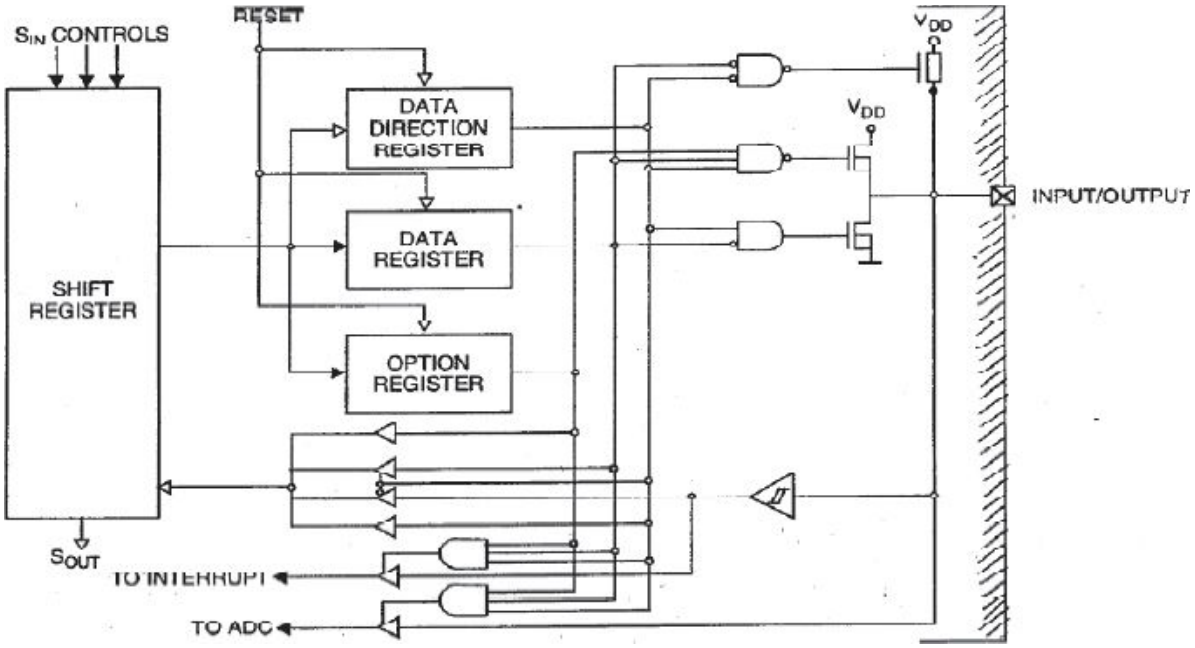


Fig. 12.21: Circuitry responsible for the management of each I/O pin.

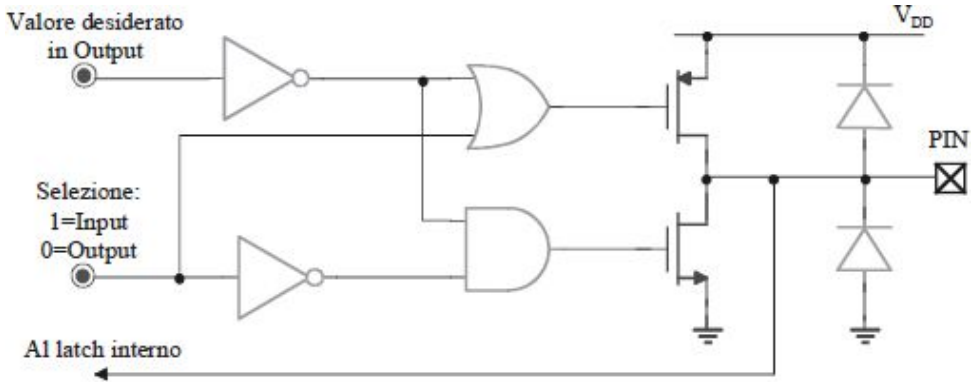


Fig. 12.22: Simplified management of outgoing or incoming data flow.

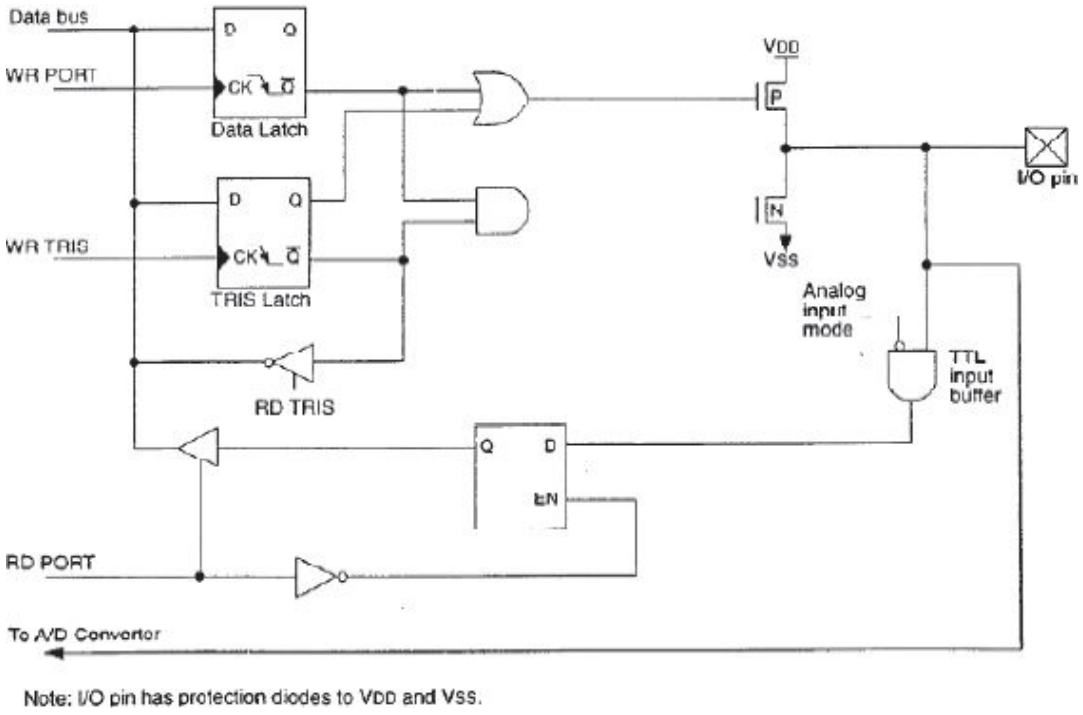


Fig. 12.23: Block diagram of a PIC16Cxx I/O pin.

12.6.2 Input/Output in the internal peripherals

Some pin I/O can be used by on-chip peripherals. For example, an input can be used as a trigger signal for the internal timers: this is the alternative role of bit 4 of port A in 16C7x PIC, which is called just RA4/TOCK1, as shown in Figure 12.24. In reference to configuration, the presence of RD TRIS buffer makes possible to know the current setting of the bits of the port. The table in Figure 12.25 summarizes how you can use various bits of Port A.

In more complex µC, there are ports that can be used as 8/16 bit bus to interface with microprocessors or memory external to the µC. In this case, in addition to the I/O data bus pins, there are pins available and designed to control writing (WR) and Read (RD) of external devices.

12.6.3 Interrupt Inputs

Finally, there are inputs that can be used to generate interrupts in the µC. In this case there will be a register to define the type of event that can trigger the interrupt, and in particular: • interrupt-on-change (takes place either on the rising or falling edge);

- interrupt-on-rising edge (takes place on the rising edge);
- interrupt-on-falling edge (takes place on the falling edge).

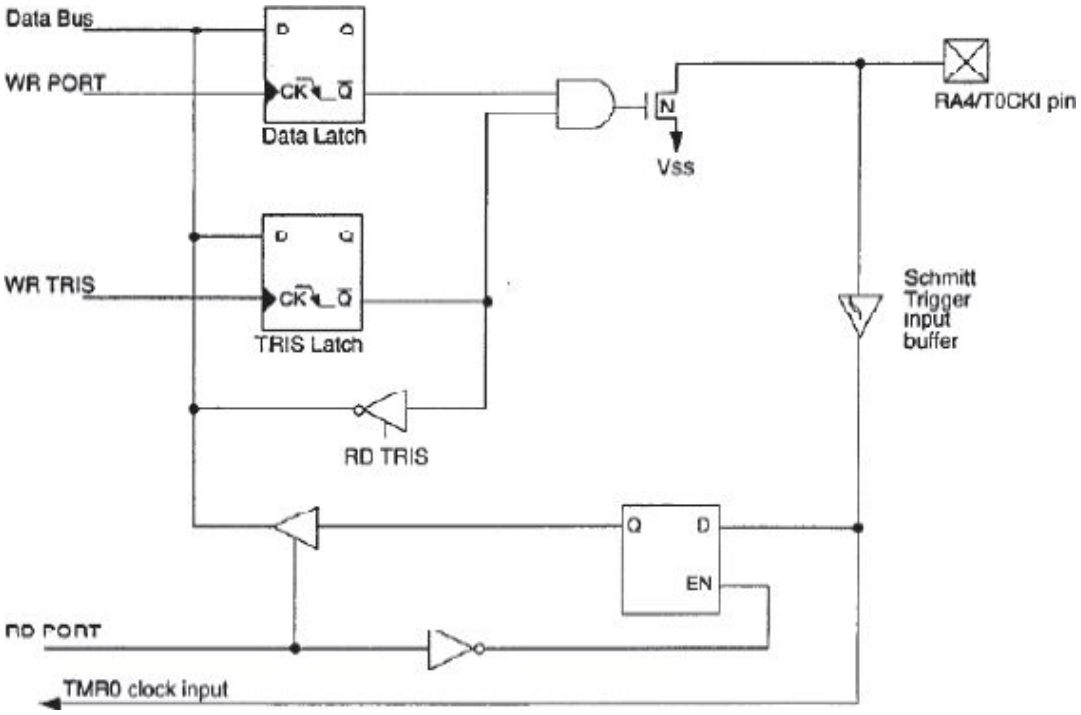


Fig. 12.24: Block diagram of a PIC16Cxx I/O pin that can be used also as input for the internal.

PORT FUNCTIONS

Name	Bit #	Buffer	Function
RA0/AN0	bit0	TTL	Input/output or analog input
RA1/AN1	bit1	TTL	Input/output or analog input
RA2/AN2	bit2	TTL	Input/output or analog input
RA3/AN3/VREF	bit3	TTL	Input/output or analog input/VREF
RA4/T0CKI	bit4	ST	Input/output or external clock input for Timer0 Output is open drain type
RA5/AN4/NOT SS	bit5	TTL	Input/output or slave select input for synchronous serial port or analog input

Legend: TTL = TTL input , ST = Schmitt Trigger input

Fig. 12.25: Functions of some I/O pins.

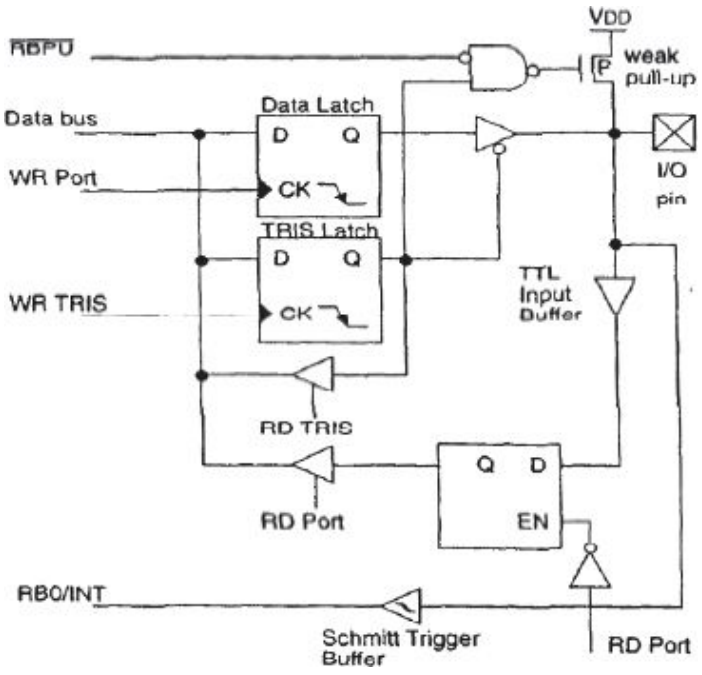


Fig. 12.26: I/O pin that can be used also as input able to generate an interrupt.

For example, 4-pin port B PIC16C7x (from RB0 to RB3, shown in Figure 12.26) offer the possibility of triggering an interrupt.

12.7. Interrupt Management

When an asynchronous event interrupt (Interrupt Request, IRQ) the main program is interrupted and runs a dedicated routine. Interrupt requests may come from multiple sources: • from pins of any I/O port if adequately abilitated;

- from the privileged Non-Maskerable Interrupt (NMI) pin, thus from outside;
- from internal devices (for example by the ADC, to communicate the end of conversion and the availability of the data, or by timers to signal the end of the counting).

To handle an interrupt the µC has to know the address in the program memory where the firmware management routine itself is stored, which may be different depending on the source that triggered the interrupt. This selection can be implemented in two different ways: an interrupt vector or polling.

Interrupt Vector

It is a memory area composed of several neighboring cells, usually 7, each of which contains the first address of the interrupt routine. Depending on the source that triggers the call, the core of the µC will be taken into the cell of the

vector address to jump to. In high-end μC there is a cell to any possible interrupt cause, thus for any internal device and dedicated pin.

To prevent the μC to be interrupted too often by the occurrence of different interrupts, usually the internal architecture makes it possible to “mask” the least significant, for example, originated from slower devices or of less importance. In this way, depending on what you want the μC to develop, the designer will decide whether or not to mask some interrupts of lower priority than a certain threshold set by firmware.

It could happen, maybe because of an error in the firmware or a real bug, that you set a priority too high, so that all interrupts are masked: in this way would not be possible to interrupt the normal operation of the μC. To avoid this, manufacturers nevertheless ensure the presence of a pin NMI, as said not-maskable, always able to trigger the call to the proper routine, thus leaving the μC from its deadlock or failure.

Polling

Low-end μC have only one interrupt cell to read from the start address of the only interrupt routine implemented. All other internal and dedicated pins are handled by this single routine, whatever source that triggered the interrupt. To identify the caller device, the designer has to develop a routine that goes for a test for the devices in order to locate the source of the interrupt. For example, it will check in succession which pin was asserted, or whether it was the end of conversion of the ADC, or the overflow of the timer, or by the arrival of data at a given serial port, and so on.

It is noted that this mode is definitely more expensive in terms of execution time of the previous one, which routed the core directly to the specific procedure for the source that led the interrupt.

Another observation, common to both modes of management is related to whether or not to handle nested interrupts, which are triggered when it is already running an administration routine of an interrupt triggered earlier. The ability to handle nested interrupts requires stack deep enough to save the contents of the program counter and other important registers, and a core complex enough to allow these “context switching”, every time a new interrupt is called. For these reasons only high-end μC allow the nesting of two (or so) interrupt levels, while for the low-end the designer has to write a firmware in order to to avoid responding an interrupt call if the μC is still running the previous interrupt handling routines. In particular, the first instruction of the

interrupt routine will disable all other interrupts and then re-enable them at the end of the procedure itself, precede the return.

In addition to disabling all interrupts in the first instruction of the routine, the contents of the registers that has not to be lost has to be stored, because they would be modified, even inadvertently, during the execution of the routine itself. Unlike the program counter which is always automatically saved on the stack, the other registers (W, status, etc..) are not saved in the low-end μC. Only at the end of the routine, once carried out the necessary operations, we must restore the contents of the registers before returning to the main program. Figure 12.27 shows an example of firmware for the rescue and subsequent restoration of the W and status registers in the PIC.

In addition to the interrupts from the on-chip peripherals, the μC allow you to configure some pins for external interrupt. To enable all or only some interrupts, you can act on some special registers contained in μC. Interrupt events from devices can be enable or excluded by setting the individual bits of these registers (such as TIMER, RB doors, A / D converter, synchronous or asynchronous serial devices, USART, etc.).. In addition, these records can be done by asking the polling of the current situation of a device. In the PIC there's the INTCON register (Fig. 12:28).

There is also a record enabling interrupts from other peripherals on-chip (Peripheral Interrupt Enable register, PIE), shown in Figure 12.29. Finally, there is the Peripheral Interrupt Register RIP (Fig. 12.30), which contains the "flag bit" of the states of individual on-chip peripherals that can cause an interrupt if enabled by the previous register.

It's important to note that the execution of the interrupt routine is not exactly coincident with the invocation of the interrupt itself. First, when the call interrupt, the μC performs clock cycles in order to end the ongoing operations, as shown in Figure 12.31. In addition, the INTF flag is read only during Q1 part of the clock. It's just after this moment that the program counter gets saved and jumps to the first instruction of the interrupt vector (in the example of Figure 12.31 is the address 0004_H). Ultimately this latency is approximately $3 \div 4 \cdot Tcy$, which may be a considerable time, especially in fast real-time applications. For example with a 20MHz clock and a $Tcy=4/f_{ck}=200ns$, this would mean the implementation of main body of the interrupt routine after at least 1μs.

```
MOVWF   W_TEMP              ;Copy W to TEMP register, could be bank one or zero
SWAPF   STATUS,W            ;Swap status to be saved into W
BCF     STATUS,RP0          ;Change to bank zero, regardless of current bank
MOVWF   STATUS_TEMP         ;Save status to bank zero STATUS_TEMP register
:
:(ISR)
:
SWAPF   STATUS_TEMP,W       ;Swap STATUS_TEMP register into W
                            ;(sets bank to original state)
MOVWF   STATUS              ;Move W into STATUS register
SWAPF   W_TEMP,F            ;Swap W_TEMP
SWAPF   W_TEMP,W            ;Swap W_TEMP into W
```

Fig. 12.27: I nstructions in the interrupt routine, to save and restore some registers.

INTCON REGISTER FOR PIC16C72/73/73A/74/74A (ADDRESS 0Bh, 8Bh)

R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-x
GIE	PEIE	T0IE	INTE	RBIE	T0IF	INTF	RBIF
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
- n = Value at POR reset

bit 7: **GIE:**(1) Global Interrupt Enable bit
1 = Enables all un-masked interrupts
0 = Disables all interrupts

bit 6: **PEIE**: Peripheral Interrupt Enable bit
1 = Enables all un-masked peripheral interrupts
0 = Disables all peripheral interrupts

bit 5: **T0IE**: TMR0 Overflow Interrupt Enable bit
1 = Enables the TMR0 interrupt
0 = Disables the TMR0 interrupt

bit 4: **INTE**: RB0/INT External Interrupt Enable bit
1 = Enables the RB0/INT external interrupt
0 = Disables the RB0/INT external interrupt

bit 3: **RBIE**: RB Port Change Interrupt Enable bit
1 = Enables the RB port change interrupt
0 = Disables the RB port change interrupt

bit 2: **T0IF**: TMR0 Overflow Interrupt Flag bit
1 = TMR0 register has overflowed (must be cleared in software)
0 = TMR0 register did not overflow

bit 1: **INTF**: RB0/INT External Interrupt Flag bit
1 = The RB0/INT external interrupt occurred (must be cleared in software)
0 = The RB0/INT external interrupt did not occur

bit 0: **RBIF**: RB Port Change Interrupt Flag bit
1 = At least one of the RB7:RB4 pins changed state (must be cleared in software)
0 = None of the RB7:RB4 pins have changed state

Note 1: For the PIC16C73 and PIC16C74, if an interrupt occurs while the GIE bit is being cleared, the GIE bit may be unintentionally re-enabled by the RETFIE instruction in the user's Interrupt Service Routine. Refer to Section 14.5 for a detailed description.

Fig. 12.28: Interrupt handling register: Each bit indicates whether the corresponding pin has unleashed (1) or not (0).

PIE1 REGISTER PIC16C73/73A/74/74A (ADDRESS 8Ch)

R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0
PSPIE(1)	ADIE	RCIE	TXIE	SSPIE	CCP1IE	TMR2IE	TMR1IE
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
- n = Value at POR reset

bit 7: **PSPIE(1):** Parallel Slave Port Read/Write Interrupt Enable bit
1 = Enables the PSP read/write interrupt
0 = Disables the PSP read/write interrupt

bit 6: **ADIE:** A/D Converter Interrupt Enable bit
1 = Enables the A/D interrupt
0 = Disables the A/D interrupt

bit 5: **RCIE:** USART Receive Interrupt Enable bit
1 = Enables the USART receive interrupt
0 = Disables the USART receive interrupt

bit 4: **TXIE:** USART Transmit Interrupt Enable bit
1 = Enables the USART transmit interrupt
0 = Disables the USART transmit interrupt

bit 3: **SSPIE:** Synchronous Serial Port Interrupt Enable bit
1 = Enables the SSP interrupt
0 = Disables the SSP interrupt

bit 2: **CCP1IE:** CCP1 Interrupt Enable bit
1 = Enables the CCP1 interrupt
0 = Disables the CCP1 interrupt

bit 1: **TMR2IE:** TMR2 to PR2 Match Interrupt Enable bit
1 = Enables the TMR2 to PR2 match interrupt
0 = Disables the TMR2 to PR2 match interrupt

bit 0: **TMR1IE:** TMR1 Overflow Interrupt Enable bit
1 = Enables the TMR1 overflow interrupt
0 = Disables the TMR1 overflow interrupt

Note 1: PIC16C73 and PIC16C73A devices do not have a Parallel Slave Port implemented, this bit location is reserved on these two devices, always maintain this bit clear.

Fig. 12.29: 2 Enable interrupts from device register: Each bit indicates whether the device is enabled on (1) or not (0) to trigger an interrupt.

PIR1 REGISTER PIC16C73/73A/74/74A (ADDRESS 0Ch)

R/W-0	R/W-0	R-0	R-0	R/W-0	R/W-0	R/W-0	R/W-0
PSPIF(1)	ADIF	RCIF	TXIF	SSPIF	CCP1IF	TMR2IF	TMR1IF
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
- n = Value at POR reset

bit 7: **PSPIF(1)**: Parallel Slave Port Read/Write Interrupt Flag bit
1 = A read or a write operation has taken place (must be cleared in software)
0 = No read or write has occurred

bit 6: **ADIF**: A/D Converter Interrupt Flag bit
1 = An A/D conversion completed
0 = The A/D conversion is not complete

bit 5: **RCIF**: USART Receive Interrupt Flag bit
1 = The USART receive buffer is full
0 = The USART receive buffer is empty

bit 4: **TXIF**: USART Transmit Interrupt Flag bit
1 = The USART transmit buffer is empty
0 = The USART transmit buffer is full

bit 3: **SSPIF**: Synchronous Serial Port Interrupt Flag bit
1 = The transmission/reception is complete
0 = Waiting to transmit/receive

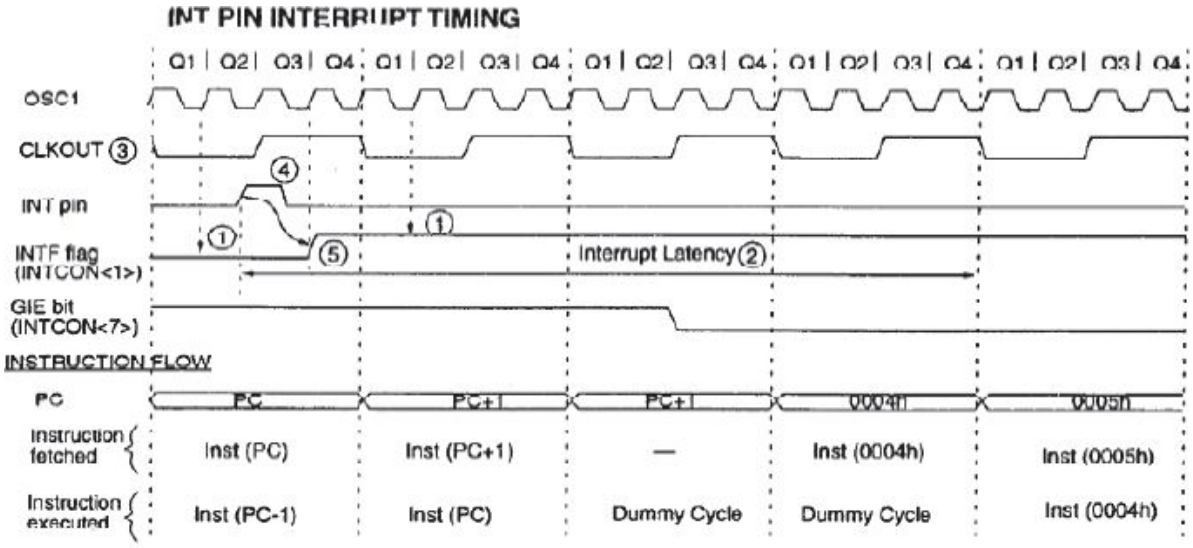
bit 2: **CCP1IF**: CCP1 Interrupt Flag bit
Capture Mode
1 = A TMR1 register capture occurred (must be cleared in software)
0 = No TMR1 register capture occurred
Compare Mode
1 = A TMR1 register compare match occurred (must be cleared in software)
0 = No TMR1 register compare match occurred
PWM Mode
Unused in this mode

bit 1: **TMR2IF**: TMR2 to PR2 Match Interrupt Flag bit
1 = TMR2 to PR2 match occurred (must be cleared in software)
0 = No TMR2 to PR2 match occurred

bit 0: **TMR1IF**: TMR1 Overflow Interrupt Flag bit
1 = TMR1 register overflowed (must be cleared in software)
0 = TMR1 register did not overflow

Note 1: PIC16C73 and PIC16C73A devices do not have a Parallel Slave Port implemented, this bit location is reserved on these two devices, always maintain this bit clear.

Fig. 12.30: Warning interrupt from device register: Each bit indicates whether the device has unleashed an interrupt on (1) or not (0).



Note 1: INTF flag is sampled here (every Q1).
2: Interrupt latency = 3-4 Tcy where Tcy = instruction cycle time.
Latency is the same whether Inst (PC) is a single cycle or a 2-cycle instruction.
3: CLKOUT is available only in RC oscillator mode.
4: For minimum width of INT pulse, refer to AC specs.
5: INTF is enabled to be set anytime during the Q4-Q1 cycles.

Fig. 12.31: Typical timing concerning an interrupt call from the dedicated INT pin.

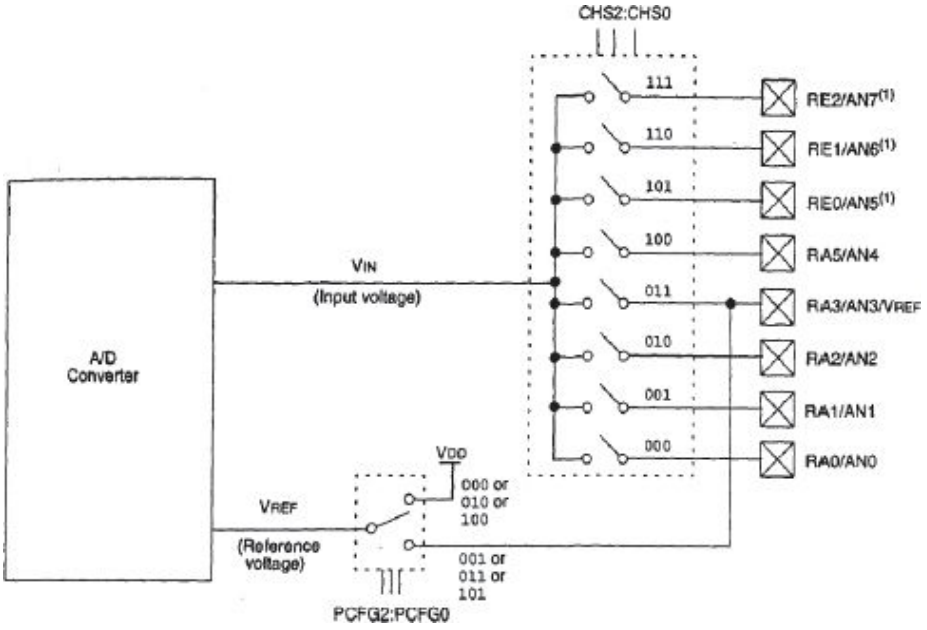


Fig. 12.32: Block diagram of the analog-digital conversion device of PIC16C7x.

12.8. ADC Converter

Many mid-range microcontrollers have an internal analog-digital converter, usually 8-bit, analog multiplexers and Sample & Hold, as shown in Figure 12.32. The most expensive µC can have up to 10 bit internal ADC and up to 12 input mux. The analog input is sent through internal lines CHS0:CHS2 to one of the switches (pass-transistor MOSFET) that selects the internal ADC. In some µC, as in the case of PIC, the reference voltage of the ADC can be set from outside, through one of the eight analog inputs available (the fourth from the bottom in Figure 12:32). The converter itself is usually a SAR, as shown in Fig 12:33.

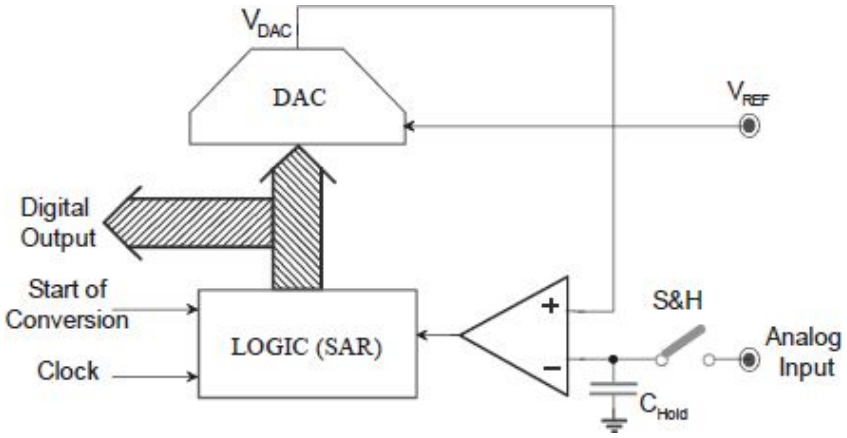


Fig. 12.33: Implementation of the ADC, based on successive approximation register (SAR).

The instant the input port is set (via CHS0, CHS1 and CHS2), the µC closes the Sample&Hold. Then, after waiting long enough so that the voltage across the capacitor is settled (Fig. 12:33), you can start the conversion (Start of Conversion, SoC) by setting to 1 bit GO/DONE in the ADC configuration registry, the ADCON0). At this point the µC will require a certain number of clock periods to complete the conversion (one for each bit of the converter). It should be noted, however, that the ADC work frequency is not the clock frequency of the µC but a submultiple. This is necessary to ensure the appropriate settling time of analog circuitry, to minimize conversion errors and possible problems (such as clock-feedthrough, cross-talk, etc..) from other digital blocks inside the µC. In general, therefore, the conversion is quite slow and takes a few microseconds in the mid-µC.

After that the result of the operation can be read directly in register ADRES. Of course the data is the correct one, only if all the bits at the output of the SAR are stable. To do this, the µC firmware can ask by polling the ADC GO/DONE bit in the configuration registry, in order to check if the µC has completed the

conversion (GO/DONE=0). That may also be the same ADC to trigger an interrupt when the data conversion ends.

If during the conversion no other instructions are to be made, you can send the μC in “Sleep Mode”, through the assertion of a specific bit in the register, or disable some if not most of the internal peripherals. In this way you can significantly reduce power consumption and so reduce the “noise” from digital devices on the conversion itself. In this way, if this option is enabled, it’s the same ADC that force the “Wake-up” of the entire μC after the conversion.

12.8.1 Errors and Conversion Timing

The diagram of Figure 12:34 shows a simplified model of the analog input of the μC through which it is possible to analyze the static and dynamic errors of the conversion and, therefore, to assess the accuracy reached by the ADC.

The I/O pin (currently set as analog input) is connected to an external signal generator, characterized by its series resistance R_s. Through the switch of the analog multiplexer (made with pass-transistor MOSFET, described in the chapter dedicated to them), the signal reaches the S&H condenser C_{HOLD}. C_{pin} is the input parasitic capacitance. Although the switch will be characterized by its series resistance value of R_{SS} different from zero. There is also a parasitic series resistance R_{IC} due to interconnections, bondings and the various MOSFET along the conductive path. You may also notice the two protection diodes for electrostatic discharge, always introducing a reverse current leakage. The latter, together with other undesired spurious currents is modeled by the current source $I_{leakage}$.

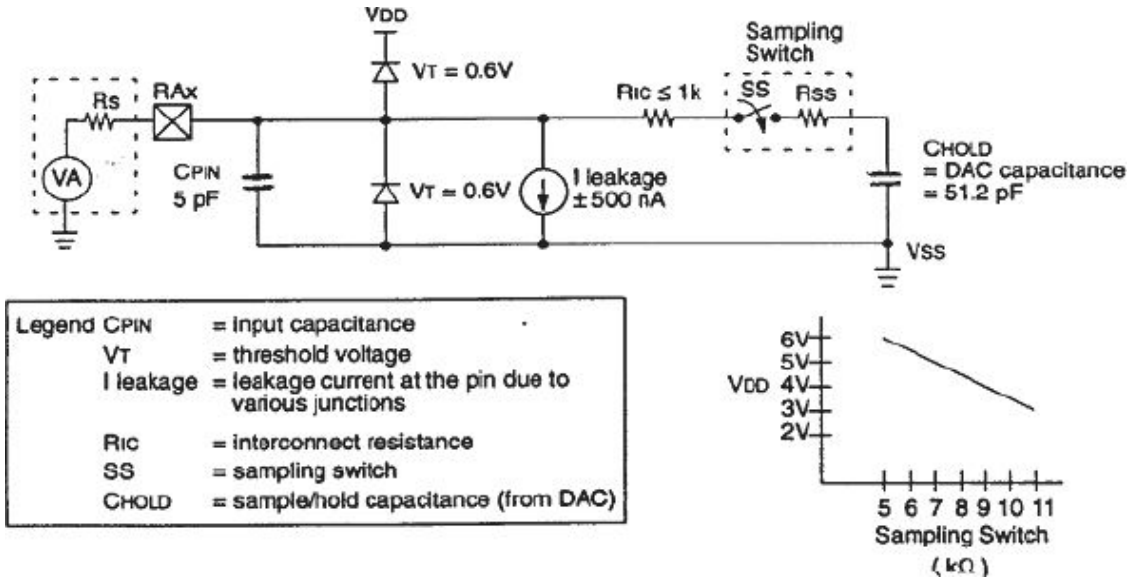


Fig. 12.34: Model of the analog input of the µC, which is necessary to estimate conversion errors.

Figure 12:34 shows the typical values measured by Microchip for their µC. You can see the series resistance of MOS with different power: the higher the tension the more ideal is the MOS function, and thus the lower is the R_{SS}. All of these parasites cause two kind of error: static and dynamic error.

Statis Error

Due to the leakage component, the voltage on pin RA_X is lower than V_A of a value equal to $I_{leakage} \cdot R_S$. Since the voltage converted by the ADC is stored in C_{HOLD}, it is important to minimize the entire fall between V_A and C_{HOLD}, to ensure that it is less than the error committed by the ADC itself, ie ½ LSB. With V_{REF}=5V and having 8bits, which is ½LSB=10mV, you must choose R_S<10mV/$I_{leakage}$=10kΩ.

Dynamic Error

When the switch is closed, the capacitor C_{HOLD} is charged with the time constant $\tau=(R_S+R_{IC}+R_{SS}) \cdot C_{HOLD}$. In order to evaluate the error between the exponential charge of C_{HOLD} and the ideal value to be converted (which would be reached at the asyntotic value), the case to be consider is the one with the maximum range, ie one in which the S&H capacitor is charging from a starting value 0V until reaching the value V_{REF}=5V.

The slope of the error is of the type $e^{-t/\tau}$; even in this case you must ensure that it is less than ½ LSB. Since the charge proceeds as a single time constant, it

is necessary to derive the minimum time T_a to ensure V_{ref} – ½LSB = (1- $e^{-Ta/\tau}$)·V_{ref} , equal to:

$$T_a = -C_{HOLD}\left(R_S + R_{IC} + R_{SS}\right)\cdot \ln\frac{½\,LSB}{V_{ref}} = 5.7\mu s$$

To be conservative, at this time you must add the settling time of the next OpAmp contained within the SAR (Fig. 12:33), which in the case of the PIC at 25°C is equal to 5µs. In addition, temperature dependence through appropriate thermal coefficientshas to be considered, again for Microchip, those are estimated at T.C.=0.05µs/°C; in the hypothesis to work at 50°C is obtained $T_{settling}$=6.25µs.

The full conversion time, from the moment you decide to get an input to the moment the numeric data is aviable in the output register of the ADC can be divided into 2 independent parts: acquisition time and conversion time.

Acquisition Time

The total acquisition time, ie the minimum time that has to be waited from the moment you have selected the analog input (by closing the switch) to the moment the conversion can start, is given by the sum of three contributions: $T_{acq} = T_a + T_{settling} + T\,.\,C\,.\cdot \Delta T = 12\mu s$

The µC does not discharge the capacity C_{HOLD} at the end of each conversion, just to speed up signal acquisition. To speed up further the acquisition, it is useful to reduce the R_S of the source. For example, being able to have R_S=0 (in reality it has to be made R_S<<R_{IC}+Rss≈8kΩ) T_{acq}=9µs could be attained. Typical values are tens of microseconds, long enough when you consider the fact that this often represent an amount of time equivalent to several tens or hundreds of instructions! In fact, in the case of PIC16Cxx at 20MHz, with an instruction cycle of 200ns, 12µs/200ns=60 instructions can be performed while waiting for the proper acquisition of 12µs at 8 bit.

Conversion Time

As soon as across the capacitor C_{HOLD} there is a stable value of the voltage to be converted, you can start the ADC conversion, setting bit GO/DONE=1 in the PIC16C7x. The conversion time is the time needed to reach to end the operation, at which point the ADC sets to 0 the GO/DONE bit. This time is directly related to the period of the clock that marks the steps of the state of the SAR ADC register (ie the duration of each step of approximation) in particular it will require at least 8 clock periods, one for each bit. More precisely, denoting

by T_{AD} the ADC clock time, ie the time required by the μC to evaluate each bit, PIC employs $9.5 \cdot T_{AD}$ for 8-bit conversion, to ensure the necessary steps to reset and initialization of the sequential network and analog component.

As mentioned above, ADC clock has a frequency lower than the μC clock: the greater its value, the faster the conversion, but at the same time, the greater will be conversion errors (due to noise, Clock feedthrough, charge injection, insufficient settling time, etc.).. With a prescaler, SAR clock of can be taken, properly divided, from the main μC clock (for example, in the PIC16Cxx you can choose via firmware between $T_{AD}=2 \cdot T_{OSC}$, $8 \cdot T_{OSC}$ or $32 \cdot T_{OSC}$), or an internal RC oscillator. Normally $f_{clock}/32$ is choosen, anyway, the manufacturer recommends in the databook $T_{ADminimum}>1.6\mu s$, to achieve the maximum precision allowed by the ADC. The selection is done by setting ADCS1 and ADCS0 bits in the corresponding special-function register ADCON0.

Ultimately, with a clock speed at 5MHz and choosing $T_{AD}=8 \cdot T_{OSC}=1.6\mu s$, conversion takes only $T_{conv}=9.5 \cdot T_{AD}=15.2\mu s$. Thus, the complete conversion (from the choice of the channel, to the availability of the result in the register ADRES) will take place in $T_{acq}+T_{conv}=12\mu s+15.2\ \mu s=27.2\mu s$.

12.8.2 Configuration registers

The block diagram of the converter (Fig. 12:33) shows how the pins of the multiplexer has to be configured as analog inputs to be used for this purpose and not as digital I/O. To do this you must configure the registry ADCON1 via firmware as shown in Fig 12:35: The letter D that appears in some cells means that the corresponding pin is controlled by the TRISA register and, therefore, is a digital I/O; instead the letter A indicates that the corresponding pin is configured as analog input.

ADCON1 REGISTER, PIC16C72/73/73A/74/74A (ADDRESS 9Fh)

U-0	U-0	U-0	U-0	U-0	R/W-0	R/W-0	R/W-0
—	—	—	—	—	PCFG2	PCFG1	PCFG0
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
- n = Value at POR reset

bit 7-3: **Unimplemented:** Read as '0'

bit 2-0: **PCFG2:PCFG0**: A/D Port Configuration Control bits

PCFG2:PCFG0	RA0	RA1	RA2	RA5	RA3	RE0(1)	RE1(1)	RE2(1)	VREF
000	A	A	A	A	A	A	A	A	VDD
001	A	A	A	A	VREF	A	A	A	RA3
010	A	A	A	A	A	D	D	D	VDD
011	A	A	A	A	VREF	D	D	D	RA3
100	A	A	D	D	A	D	D	D	VDD
101	A	A	D	D	VREF	D	D	D	RA3
11x	D	D	D	D	D	D	D	D	—

A = Analog input

D = Digital I/O

Fig. 12.35: ADCON1 Register to set the pins as digital I/O (D) or analog (A).

ADCON0 REGISTER, PIC16C72/73/73A/74/74A (ADDRESS 1Fh)

R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	R/W-0	U-0	R/W-0
ADCS1	ADCS0	CHS2	CHS1	CHS0	GO/DONE	—	ADON
bit7							bit0

R = Readable bit
W = Writable bit
U = Unimplemented bit, read as '0'
- n = Value at POR reset

bit 7-6: **ADCS1:ADCS0:** A/D Conversion Clock Select bits
00 = Fosc/2
01 = Fosc/8
10 = Fosc/32
11 = FRC (clock derived from an RC oscillation)

bit 5-3: **CHS2:CHS0:** Analog Channel Select bits
000 = channel 0, (RA0/AN0)
001 = channel 1, (RA1/AN1)
010 = channel 2, (RA2/AN2)
011 = channel 3, (RA3/AN3)
100 = channel 4, (RA5/AN4)
101 = channel 5, (RE0/AN5)(1)
110 = channel 6, (RE1/AN6)(1)
111 = channel 7, (RE2/AN7)(1)

bit 2: **GO/DONE:** A/D Conversion Status bit
If ADON = 1
1 = A/D conversion in progress (setting this bit starts the A/D conversion)
0 = A/D conversion not in progress (This bit is automatically cleared by hardware when the A/D conversion is complete)

bit 1: **Unimplemented:** Read as '0'

bit 0: **ADON:** A/D On bit
1 = A/D converter module is operating
0 = A/D converter module is shutoff and consumes no operating current

Fig. 12.36: ADCON0 Register to set the ADC clock frequency, the channel to be acquired, start and end of conversion and enabling the entire analog device on or off.

If an external V_{REF} is used, other than V_{DD}, you must set the RA3 pin as analog input and apply the desired voltage reference. Note, however, that the manufacturer guarantees a precision of 8 bits only when V_{REF} is the maximum possible, ie equal to V_{DD}. A lower reference values dramatically worsen the accuracy of the conversion, because of the noise contributions in the ADC increases relatively.

The other important register is the ADCON0 (Fig. 12:36), which allows you to set the frequency conversion (T_{AD} chosen as a multiple of the clock period of the μC, or using the internal RC oscillator), to choose the channel you want to read, to commence the conversion itself or turn off the entire ADC.

The firmware needed for an acquisition by a PIC is shown below:

DOING AN A/D

```
CONVERSION (PIC16C72/73/73A/74/74A)
BSF STATUS, RPO          ; Select Page 1
CLRF ADCON1              ; Configure A/D inputs
BSF PIE1,ADIE            ; Enable A/D interrupts
BCF STATUS, RPO          ; Select Page 0
MOVLW 0XC1               ; RC Clock, A/D is on , Channel 0 is selected
MOVWF ADCON0             ;
BCF PIR1, ADIF           ; Clear A/D interrupt flag bit
BSF INTCON, PEIE         ; Enable peripheral interrupts
BSF INTCON, GIE          ; Enable all interrupts
                         ; Ensure that the required sampling time for selected input
                         ; channel has elapsed. Then the conversion may started.
BSF ADCON, GO            ; Start A/D Conversion
:
                         ; The ADIF bit will be set and the GO/DONE bi
                         ; is cleared upon completion of A/D conversion.
```

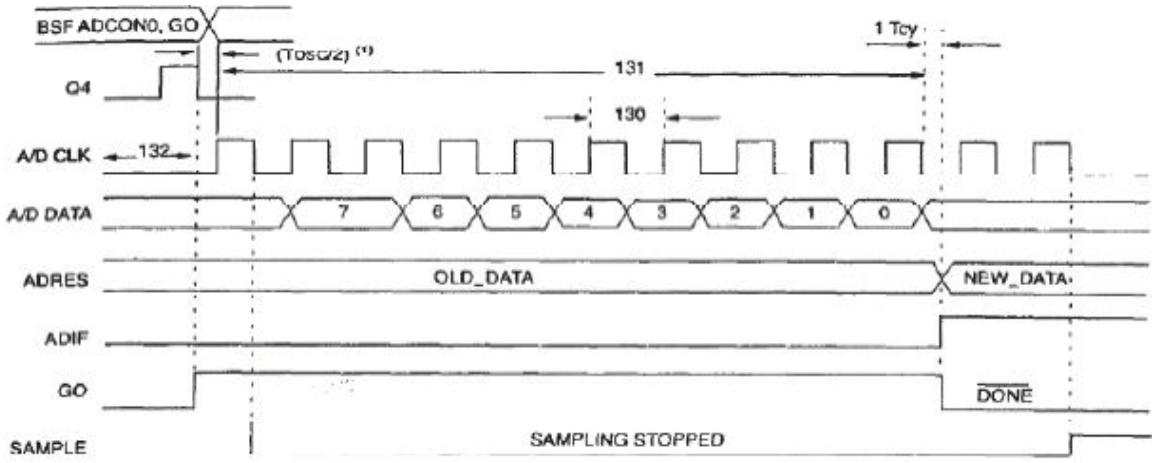
12.8.3 Performance

In the description of the performance of the ADC and its analog multiplexer, are specified both the dynamic parameters (timing and minimum delays) and the static ones (differential error, integral error, offset, full scale range of values , etc.). Figure 12:37 shows the schedule followed by the Microchip core during conversion, the values suggested by the manufacturer are listed in Fig 12:38.

Static performances are listed in Figure 12:39. The accuracy of ±1LSB is guaranteed only for a nominal 5V, indeed of 5.12V. This highlights even more the fact that, using a V_{REF} lower than V_{DD}, the accuracy of the conversion is intended to degrade significantly. Also, note that a converter that can be defined in 8bit, should have an accuracy better than ½LSB, and not just 1LSB. All these clues suggest that the accuracy of the converter in question is about that of a 7-bit ADC, although the resolution to be offered is the one of a 8-bit, thus 256

codes are available, but the last bit may be wrong (if a reading was 123, the correct value could be either 122, 123 or 124).

To improve (or at least not to decrease) the accuracy, the acquisition has not to take place during the digital I/O switching, because the heavy spikes could introduce errors. Some µC recommend to stop all other devices when you want to capture a precise value!



Note 1: If the A/D clock source is selected as RC, a time of Tcy is added before the A/D clock starts. This allows the SLEEP instruction to be executed.

Fig. 12.37: Timing during the analog to digital conversion.

A/D CONVERSION REQUIREMENTS

Parameter No.	Sym	Characteristic		Min	Typ†	Max	Units	Conditions
130	TAD	A/D clock period	PIC16C73/74	1.6	—		µs	Tosc based, VREF > 3.0V
			PIC16LC73/74	2.0	—	—	µs	Tosc based, VREF full range
			PIC16C73/74	2.0	4.0	6.0	µs	A/D RC Mode
			PIC16LC73/74	3.0	6.0	9.0	µs	A/D RC Mode
131	TCNV	Conversion time (not including S/H time) (Note 1)		—	9.5TAD	—	—	
132	TACQ	Acquisition time		Note 2	20	—	µs	

* These parameters are characterized but not tested.

† Data in "Typ" column is at 5V, 25°C unless otherwise stated. These parameters are for design guidance only and are not tested.

Note 1: ADRES register may be read on the following TCY cycle.

2: See Section 13.1 for min conditions.

Fig. 12.38: Durability requirements of the timing of Fig 12:37.

Parameter No.	Sym	Characteristic	Min	Typ†	Max	Units	Conditions
	NR	Resolution	—	—	8-bits	—	VREF = VDD = 5.12V, VSS ≤ AIN ≤ VREF
	NINT	Integral error	—	—	less than ±1 LSb	—	VREF = VDD = 5.12V, VSS ≤ AIN ≤ VREF
	NDIF	Differential error	—	—	less than ±1 LSb	—	VREF = VDD = 5.12V, VSS ≤ AIN ≤ VREF
	NFS	Full scale error	—	—	less than ±1 LSb	—	VREF = VDD = 5.12V, VSS ≤ AIN ≤ VREF
	NOFF	Offset error	—	—	less than ±1 LSb	—	VREF = VDD = 5.12V, VSS ≤ AIN ≤ VREF
	—	Monotonicity	—	guaranteed	—	—	VSS ≤ AIN ≤ VREF
	VREF	Reference voltage	3.0V	—	VDD + 0.3	V	
	VAIN	Analog input voltage	VSS - 0.3	—	VREF + 0.3	V	
	ZAIN	Recommended impedance of analog voltage source	—	—	10.0	kΩ	
	IAD	A/D conversion current (VDD)	—	180	—	μA	Average current consumption when A/D is on. (Note 1)
	IREF	VREF input current (Note 2)	—	—	1 10	mA μA	During sampling All other times

* These parameters are characterized but not tested.

† Data in "Typ" column is at 5V, 25°C unless otherwise stated. These parameters are for design guidance only and are not tested.

Note 1: When A/D is off, it will not consume any current other than minor leakage current. The power-down current spec includes any such leakage from the A/D module.

2: VREF current is from RA0 pin or VDD pin, whichever is selected as reference input.

Fig. 12.39: Static parameters and requirements of the ADC signal source.

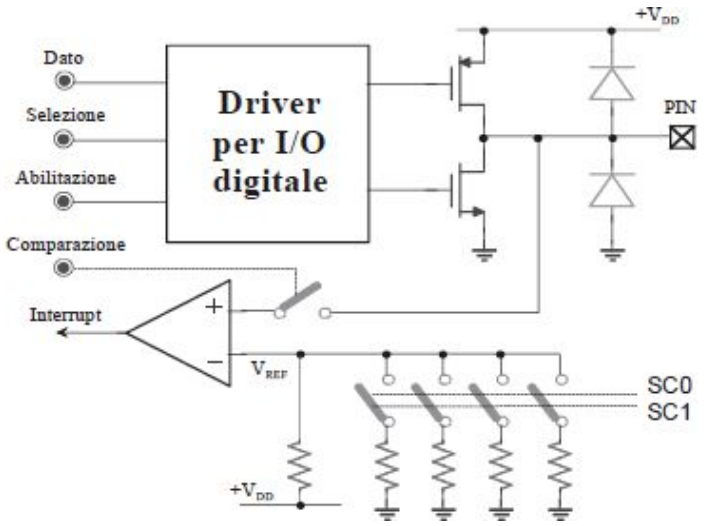


Fig. 12.40: Analog comparator in a μC: note the possibility to set the tripping threshold.

12.8.4 Comparator or Analog Watchdog

Some μC have some pin (one or two) that can be used to perform a comparison between an analog voltage and an external threshold inside the μC.

This threshold can be set via firmware, enabling a suitable resistor network partition, as shown in Figure 12.40. At a time when the threshold is exceeded, the comparator trips and you have the generation of an interrupt or a Watchdog.

This device is very useful when the µC must monitor a voltage level, from a sensor or, more usually, the battery that powers the µC itself. It is not needed in fact to waste an analog input mux and make an analog-to-digital conversion to read the voltage, the savings are significant both in terms of resources and efficiency of monitoring: you do not keep on performing ADC conversions to verify the crossing of a threshold. Sometimes there is also a dual comparator, the comparison between the two thresholds, both lower and upper adjustable via firmware. This makes it possible to allow the triggering of an interrupt when the monitored voltage gets out of the allowed range of values.

Analog Watchdog function is similar to the one of the analog comparator. The only difference is that when the voltage on the dedicated pin falls below a predetermined value an interrupt is not generated, but a full µC reset. The device acts as a real “watch dog” that forces a reboot of the µC when a voltage outside the allowed range is read.

12.9. DAC Converter

The implementation of a DAC within the µC is rare, because it leads to a wasteful consumption of silicon area and, above all, a high power consumption. In fact, while to provide a digital level (eg low) a MOS transistor is always on (the n-channel one) and the other off (the p-channel one), to provide a voltage between V_{DD}÷0, both transistors have to run, resulting in a static current consumption ways from negligible.

For these reasons, in the medium-low µC gamma, instead of a real internal DAC is preferred to use the PWM method (described in next section) to generate an analog voltage inside the µC. Only a few high-end µC have in them a DAC. For example, 7805x and 07x families from NEC have two analog outputs. Figure 12:41 shows the block diagram of NEC µPD78P054Y, in where you can see the large number of internal devices and, in particular, the DAC next to the ADC.

12.10. Exemple of use of I/O

A 16-key keyboard (such as 10 numeric and 6 auxiliary) has to be managed by a μC. We describe below some solutions.

***Exemple 1*: Digital Multiplexing**

Obviously, it is inefficient to use 16 digital inputs to monitor each key individually, it's much cheaper to operate a hardware multiplexing, scanning the keyboard for rows and columns, as shown in Fig 12:42, via an appropriate firmware.

Buttons has to be connected in a square matrix of 4 rows by 4 columns. The rows are driven in succession by 4 pins from PORT A, which will be configured as digital output. The columns are properly put in pull-down by four resistors, than connected to 4 pins of PORT B setted as digital inputs. A total of eight pins are used instead of sixteen.

Firmware will initially set all 8 pins as inputs and then configure as output PORT A pins one a time, bringing it to 1. In this way, reading the four bits of PORT B will recognize the key pressed. If a scan of the four lines revealed that more than one key has been pressed, you will have to discard the acquisition to avoid a possible misunderstanding caused by the reading of more keys simultaneously.

To prevent the µC to constantly scan the keyboard, even when no key is pressed, you can do as the pression of a button to trigger an interrupt. In this way the µC will know when to scan the ports to detect which key was pressed. In the absence of pressure, the µC can perform other operations in the main program. Although the time required by each

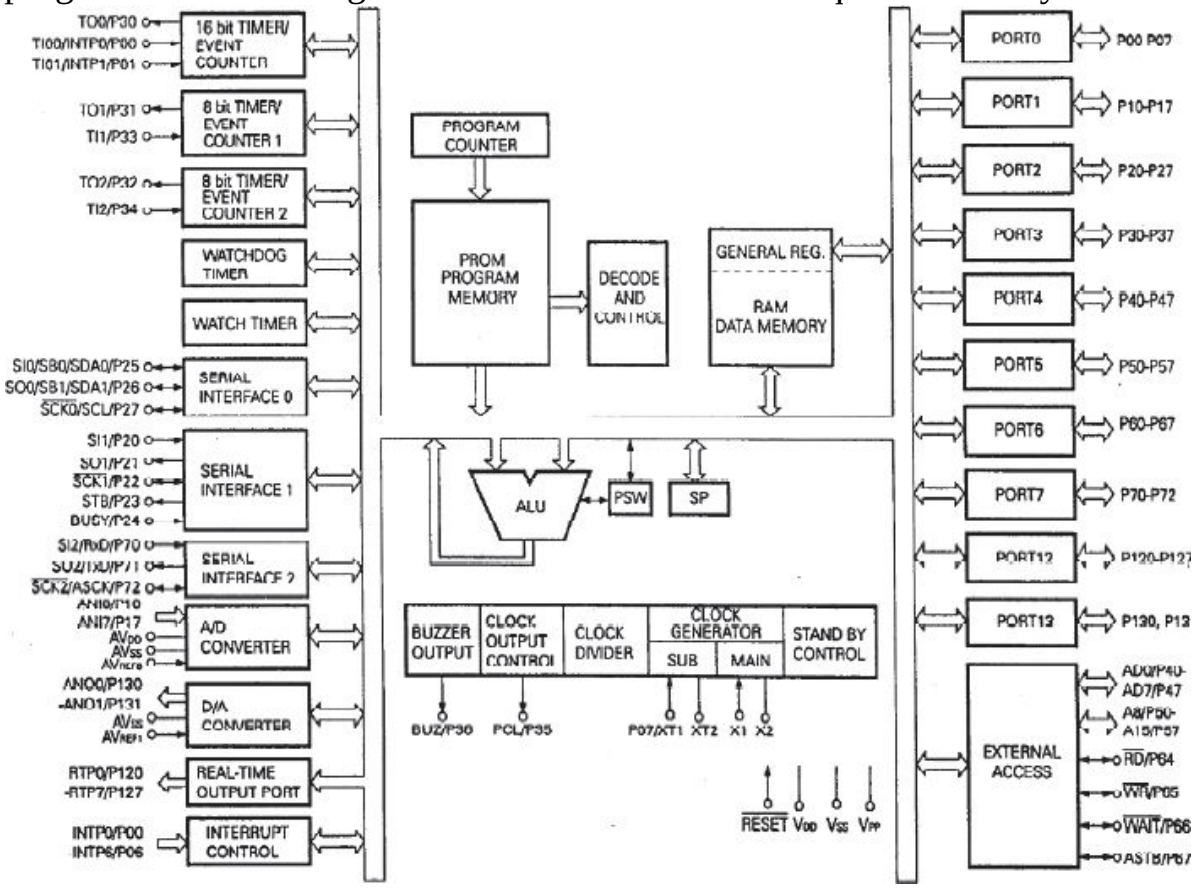


Fig. 12.41: Block diagram of the NEC µPD78P054Y: note the presence of both ADC and DAC.

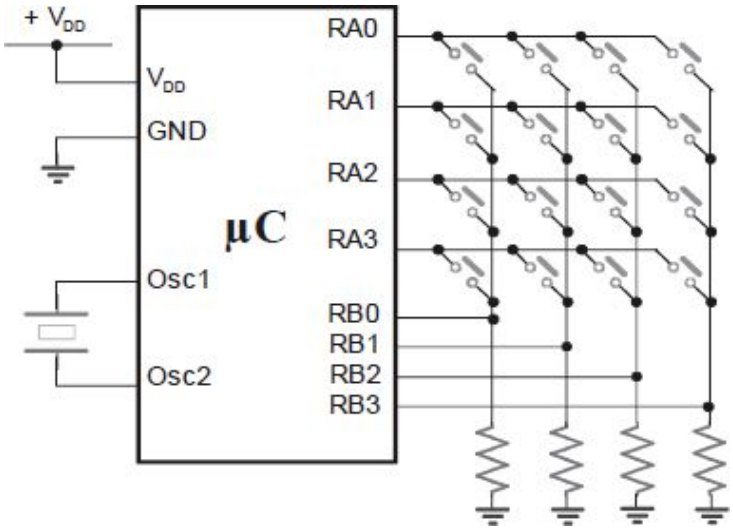


Fig. 12.42: Managing of a 16-key keypad by row and column multiplexing.

statement was 200ns for example, it would be very expensive to monitor the keyboard more often than once every 10,000 instructions (ie, a poll every 2ms) if this were to be used, for example once per second, or worse once an hour. Even more the press of a button by a human operator takes dozens (even hundreds) of millisecond.

Then in the main program, among other instructions you have to add the following commands to enable proper management interrupt:

Configurazione delle porte nel modo appena descritto;

Inizializzazione dei bit della PORT A a valore alto (tutti 1);

Abilitazione degli ingressi della PORT B a genere ‘interrupt on rising edge’;

In the management of the interrupt subroutine, without interrupt vectors, instructions will have to verify that the request comes from your keyboard:

Disabilitazione di tutti gli interrupt (per evitare la nidificazione);

Salvataggio di tutti i registri;

Si configurano come ingressi (alta impedenza) tutte le righe;

Si configura come uscita la prima riga e la si imposta a 1;

Si legge la PORT B, verificando quale bit è a 1;

Si configura come ingresso la prima riga;

Si configura come uscita la seconda riga e la si imposta a 1;

Si legge la PORT B, verificando quale bit è a 1;

Si configura come ingresso la seconda riga;

Si configura come uscita la terza riga e la si imposta a 1;

Si legge la PORT B, verificando quale bit è a 1;
Si configura come ingresso la terza riga;
Si configura come uscita la quarta riga e la si imposta a 1;
Si legge la PORT B, verificando quale bit è a 1;
Se nessun bit letto è 1 significa che nessun tasto è stato premuto;
Se solo un bit è 1 allora è stato premuto il tasto corrispondente;
Se più di un bit è 1 allora si scarta la lettura.
Ripristino di tutti i registri precedentemente salvati;
Abilitazione di tutti gli interrupt;
Ritorno al programma principale.

The time interval between the instant when you press a button and the instant in which the interrupt triggers, can, for example, be equal to 2µs. We must add to this one the time between the interrupt and the start of the scan: with the assumption that this would amount to 3µs, the delay between the moment when a key is pressed and the actual survey can be worth about 5µs. So it is impossible that the µC does not detect the touch of a button due to the too fast release of same, since the operation can not be faster than a few tens of ms.

There are other issues to be analyzed. One of them is bound to the keys bouncing that can be done at a distance of tens of milliseconds: the risk is the triggering of interrupts in succession, resulting in the execution of the routine and, ultimately, the detection of multiple pressures on the same key . A solution to this problem is to filter low-pass input signal to port B by inserting four small capacitors, with a time constant of tens of milliseconds. A firmware solution that will drop the second message if it's the same element of the keyboard can be preferred. Only a full scan without detecting any keys pressed is considered a valid successive press. Finally, the sensitivity to only "rising-edge" solve the problems arising from any continuous pressure on the button. If, however, a 'repeat fast' function would be enabled, so that pressing a button over 1s mean 'reps to push', we have to do it via firmware.

***Example 2*: Analog acquisition**

Using the ADC, you can still use the multiplexing of the lines, but with a single return line, so you can save I/O pins. The scheme is shown in Figure 12:43. This kind of deployment enables only one pin to the 'interrupt on rising edge', it is transformed into analog input to measure the voltage corresponding to the node. Upon receipt of the interrupt the columns has to be scanned slow enough in order to provide the necessary stabilization of the ADC, in a few tens

of microseconds. Obviously this system is slower than the previous one, because a conversion is required (with a duration of a few tens of microseconds).

An appropriate configuration of resistors is prepared: their value, around 1kΩ, was chosen to prevent that the leakage current of the ADC input don't results in an excessive error in terms of LSB. The voltage on the AN2/RB0 pin will be tied to column that belongs to the key pressed, as shown in the table of Fig 12:43.

With the button pressed in the first column (remember that the other lines should be set in three-state, thus as inputs) the output current is the maximum, that is 5V/10k Ω=0.5mA, a value small enough not to cause a significant lowering of the port V_{OH}, thus there is no need to use a buffer.

Note however that an outside buffer would be difficult to use ensuring the abilitation of only one row at a time. In fact, with a quad-buffer the outputs of the four-door can get only the values 0 or 1 and it would not contain the three-state, the simultaneous pressing of multiple keys on different rows cause short circuits and can damage the circuit.

This configuration problems results from resistors tolerance that prevent the voltage to be precise. In the case of resistance at 10%, the worst-case fluctuation occurs by pressing the fourth column, thus the tension could take the extreme values:

$$V_{max,min} = 5 \cdot \frac{(10K \pm 10\%)}{(10K \pm 10\%) + (3.3K + 6.7K + 20K \mp 10\%)}$$

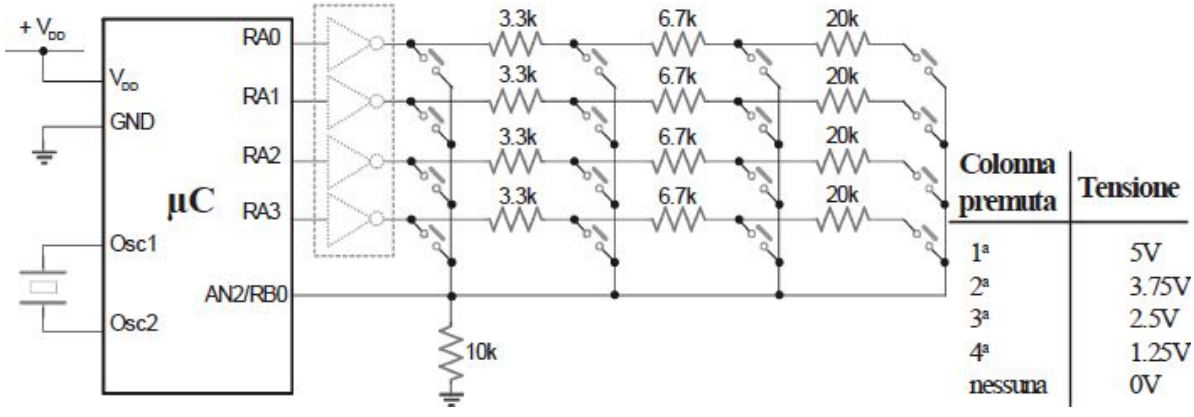


Fig. 12.43: Management of a keyboard by reading analog voltage. Note the inability to use the quad-buffer shown in hatch, as it would short-circuit.

The voltage can therefore change from V_{max}=1.45V to V_{min}=1.07V. Pressing the third column we would get: $V_{max,min} = 5 \cdot \frac{(10K \pm 10\%)}{(10K \pm 10\%) + (3.3K + 6.7K \mp 10\%)}$
thus: V_{max}=2.75V e V_{min}=2.25V. Last, for the second column we would get:

$$V_{max,min} = 5 \cdot \frac{(10K \pm 10\%)}{(10K \pm 10\%) + (3.3K \mp 10\%)}$$

thus V_{max}=3.94V e V_{min}=3.56V.

It's necessary to ensure that all the values in these ranges will correspond to the correct key. For example, referring again to the example in question, we can simply choose the thresholds of the ADC as follows: between 0÷1V thus 00_H÷33_H (no key), 1÷2V 34_H÷66_H (4th column), 2÷3V 64_H÷99_H (3rd column), 3÷4V $9A_H$÷CC_H (2nd column) and last 4÷5V CD_H÷FF_H (1st column). 8-bit ADC resolution is not needed then, but it is sufficient to monitor the first $\log_2 5 \approx 3$bit encoding.

Based on the foregoing, theoretically the number of columns could increase to the limit of 256 columns, but unfortunately, the range of values would get too small and would be subjected to numerous errors: close key columns may be interpreted as the same key, or even be misunderstood. In addition the voltage to be converted directly depends on the voltage line: it is 5V only theoretically, while more realistically it falls with increasing current drawn. For example, the current drawn when you press a key in the first column is 4 times greater than one in the last column.

Another implementation problem concern the management of the case in which several keys are pressed simultaneously. In the case of keys on different columns (not rows), prevails the key in the "lower" column. Instead, in the case of keys pressed on different rows "ghost key" issue may occur (by pressing three buttons down looks like a "fourth" kay in square position with the other three is pressed) in the multiplexing circuit of rows and columns of ***Examples 1*** and ***2*** but not in the case of the circuit under test, with a single analog input.

***Example 3*: Simgle analog line**

This method is more economical in terms of µC pins usage. It still uses a pin enabled to the "interrupt on falling-edge", which is then reset as analog input in order to read the voltage. The scheme is shown in Fig 12:44.

This method problems are the lack of immunity to noise and spikes (unless you put the filter capacitor C) and the problem caused by simultaneous pressure of keys.

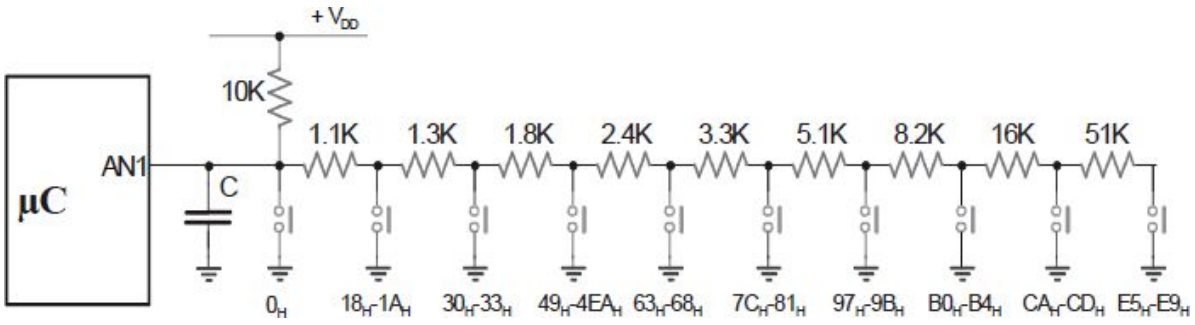


Fig. 12.44: Management of a keypad with a single analog input.

Resistors must be chosen in order to achieve uniform coding intervals, for example, just as shown in Figure 12:44. Again you can not recognize more than one key pressed simultaneously, as theone on the left wouldprevail on the other. The tolerance of the resistors has to be 2%, leading to a distance between the codification in the order of 20 codes. The maximum current consumption is 5V/10k Ω=0.5mA, the value can be decreased by increasing the value of all resistances, but doing so would increase greatly the current $I_{leakage}$ from the ADC.

The inclusion of a filtering capacitor significantly slows the acquisition, for example, with C=47nF, in the worst case, the data will stabilize with a time constant 10kΩ·47nF≈0.5ms. Thus, to reach the correct code within 3LSB we would have to wait for T=0.47ms·ln(256/3)=2ms. The wait would be extremely expensive, since in that time you can do 10,000 instructions (assuming f_{ck}=20MHz and an instruction cycle of 200ns). It's therefore preferable to give up the filter capacitor, threatening empty calls due to spikes.

A major problem of this circuit is the inability to implement at the input an "interrupt on falling edge", because the pressure of far keys is not enough to provoke it. The µC is thus forced to use a continuous polling.

12.11. Timer and Counter

Both timer and counter are located on the same device and perform the same counting of events function. However, while the timer counts the time, marked by the µC clock (or a submultiple), the counter counts pulses arriving at an

input pin, suitably qualified. The latter is used for the μC to count external events coming from sensors, photocells, microswitch, etc.

Usually in μC are one to three timer/counter peripheral that can be used in various ways. Each of them can generate its own interrupt at the end of the counter (overflow). Count frequency can be done by the internal clock or by an external one, provided through a dedicated I/O pin.

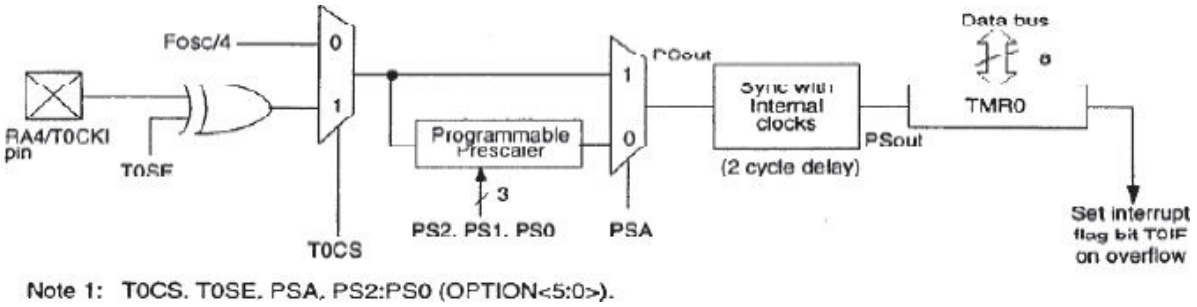


Fig. 12.45: Block diagram of TIMER 0 in PIC16Cxx microcontrollers.

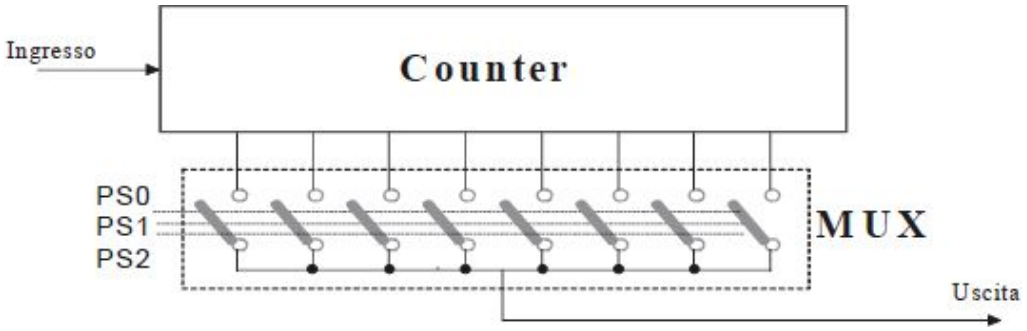


Fig. 12.46: Prescaler diagram to split input impulse.

However, even in the latter case there can be a synchronization system with the switching of the μC internal clock through a timing circuit, which stores the arrival of the event and makes it available in the cycle in which the core updates the internal timer. This capability is important if $f_{clock}<f_{timer}$ (eg if the μC works at low frequency, to reduce consumption), but unfortunately this may cause a one or two clock cycles delay.

A standard timer can be a 8bit one with an additional overflow bit, eventually with a programmable prescaler, that can be disabile in SLEEP mode (in that case it cannot be used for the wake up), and a syncronization unit. A common block diagram for TIMER0 in Microchip μC is the one shown in Fig.12:45. The register that controls TIMER0 in PIC16Cxx is the Option register. It’s possible

to select between timer mode or counter via firmware, the TOCS bit and its mux. In timer mode a submultiple of the μC clock frequency is used ($f_{osc}/4$), while in counter mode TOCKI pin is used to count external events, in rising-edge switch (if TOSE=0) or falling-edge (if TOSE=1), through the EXOR port. Using programmable prescaler (through PS2:PS0 bits), pulse counting can be further divided. In this way, you can have a timer that counts down the seconds, even if clock f_{osc} works in megahertz range, or a counter that counts input events in multiples of powers of two. The prescaler operates as a frequency divider and is obtained by a counter followed by a digital multiplexer output, as shown in Fig 12:46: with a 3-bit selection, contained in the Option register, 8 different divisions can be achieved, from a minimum ratio of 1:1 to a maximum of 1:128.

12.11.1 Timing

When you start the timer, there is always a time lag before you get the correct value of the count. For example, once you set the value in the timer with the instruction MOVWF TMR0 two instruction cycles have to be have to waited (eight clock cycles) before the timer starts to count and to be directly readable through the instructions MOVF TMR0, W. You can avoid this by increasing by two the initial value set in the timer. Figure 12:47 shows an example of timing.

As seen above, the timer can generate an interrupt. For example, it occurs when TIMER0 overflows, switching from FF_H to 00_H. So, if you want to do a count-down of x pulses, you can load the starting value of 256-x in irder to, come to FF_H, it will reset generating overflow and asserting the bit <2> (TOIF) of the same register. The interrupt triggered can be masked by putting to zero TOIE bit in the proper register (that is the special register INTCON, in the case of PIC).

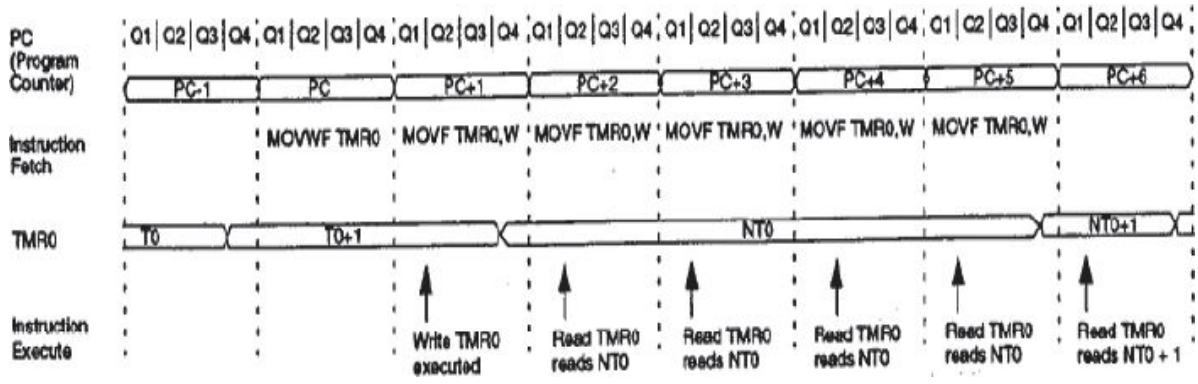


Fig. 12.47: TIMER0 timing with internal cloack and 1:2 prescaler.

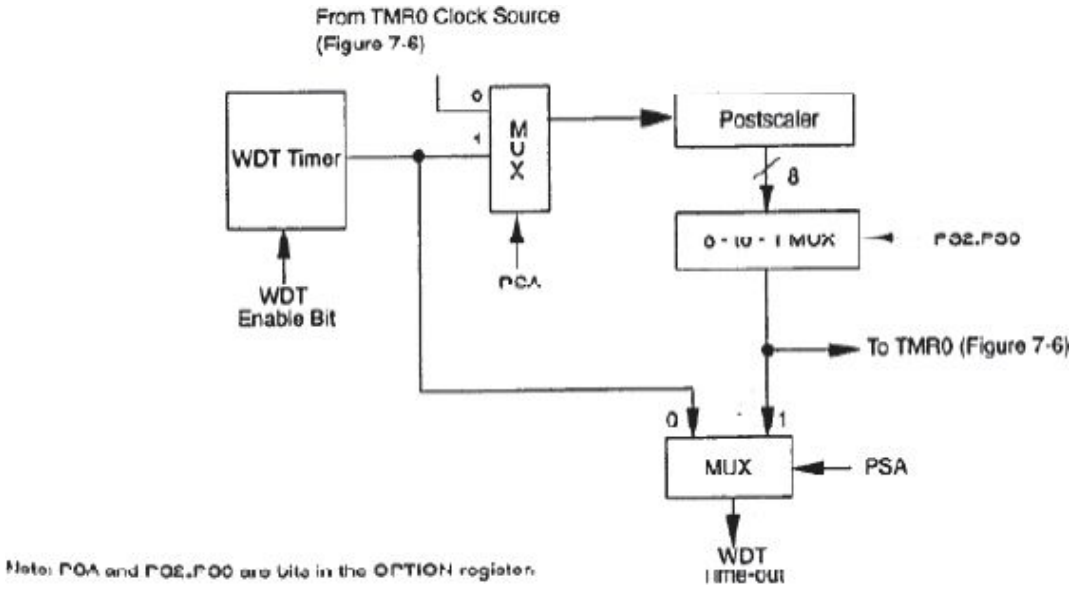


Fig. 12.48: Watchdog timer block diagram.

12.11.2 Watchdog

The μC, as every other state machine, has the risk to exit from the correct flow of processing and, ultimately, to get stalled because of firmware error, hardware, or disturbances in supply line (spike) or in I/O pins. To escape this state of impasse an internal circuit called Watchdog is provided, that after finding the abnormal operation of the μC, reset it avoiding the menance of proceeding in an uncontrolled way. The choice whether or not to enable it is done by asserting a bit in the planning stage of the μC.

This device consists essentially of a self-timer that continually counts down and when it reaches zero (ie if you detect the underflow) resets the μC. During normal operation, that is when everything is proceeding properly, firmware must prevent this, simply resetting the idle timer to FF_H, through a proper instruction. If at any point the μC went to stall or start processing a indefinite instructions, the μC will fail to reset the watchdog counter allowing it to go full-scale, thus triggering a reset of the entire machine. Typically the period of the watchdog is large enough, for example in PIC16Cxx is 18ms and can reach 2 seconds if you use an postscaler 1:128 (Fig. 12:48).

The watchdog can also be used to wake up the μC from a SLEEP state (wake-up planned) as an alternative method to interrupt. The ‘SLEEP’ is a low-

power state of the µC in which internal peripherals (such as ADC, various timers, all those that are not used) are disabled or disconnected from the core. Suppose you want to use the ADC inside the µC in order to capture slowly varying signals: you may want to lower the clock to 100kHz (each instruction takes 40µs) to perform slow and precise conversions. In this way a conversion typically require several hundred milliseconds; wanting to save power (and therefore battery life) the implementation of a so long wait loop in the firmware is unacceptable. Then, after setting the conversion, you can put the µC in SLEEP and make sure that it's not the ADC but the watchdog to trigger the interrupt for the wake up and prosecution of operations of the µC itself.

The presence of a watchdog is very useful in any programmed or sequential machine thus if there is no internal watchdog it's often a good idea to implement one outside of the µC. A simple external watchdog should be able to generate a reset without having to test all the bits of the machine and without the need for additional clock generators: in practice it has to be very simple and functional. There are some commercially available watchdog

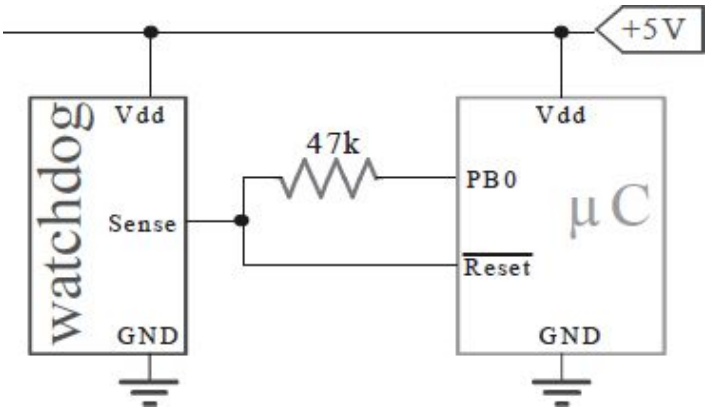


Fig. 12.49: Connection to an external watchdog.

with only 3 pins. Their operation is as follows: if the circuit does not receive the commutations on his only "sense" I/O line for more than 500ms, it forces itself on a low level that will act as a reset for the downstream electronics. A simple partitive external network allows you to create the appropriate levels and to avoid conflicts with the output level of the µC, as shown in Figure 12:49. Obviously after the reset you must ensure that the µC starts back from a known state.

12.11.3 Pulse-Width Modulation

The Pulse Width Modulation consist in a continuous change of the duty-cycle of a square wave with fixed frequency (Fig. 12.50). In the µC case, the

frequency is generated simply as a multiple of the clock F_{osc}. Taking the average of this signal, one gets a DC voltage (or low frequency) with amplitude range between 0 and V_{DD}. With this technique you can get an analog output from the µC (simply adding an external low-pass filter) controlled by a digital value (the duty-cycle) without the µC to possesses an internal DAC. It can be used, for example, to adjust the speed of a DC motor.

Power dissipation in the output MOS transistors is always very low: it will alternately one of the couple V_{DS}=0 and I_D=I_{MAX}, or V_{DS}=V_{MAX} and I_D=0. Of course, the durations of PWM control transitions can be considered negligible. If you are using the PWM as a real DAC, both stability and the absence of ripple in the output voltage depends on the filter used.

Wanting to reconstruct at the output a slow varying analog signal, you can generate the signal via firmware usinfg a timer inside the µC or you can generate through a device dedicated to the purpose. Figure 12:51 shows the block diagram of the second approach, employing the device Capture & Compare present in some µC capable to create a square wave PWM duty cycle controlled by the contents of a register. The TIMER is placed in counter mode and its output value is constantly compared with the one of the register: the output level will be low whenever the value will be higher than the register. The frequency of oscillation of the TIMER can be obtained directly from the clock through a prescaler. It's given a square wave with a period of $f_{PWM}=f_{ck(timer)}/2^n$.

The output will stay up for a time equal to: $$T_{on} = \frac{A}{f_{ck(timer)}} = \frac{T_{PWM}}{2^n} \cdot A$$

where A is the contento f the comparison register, with a digital value between $0 \div (2^n-1)$.

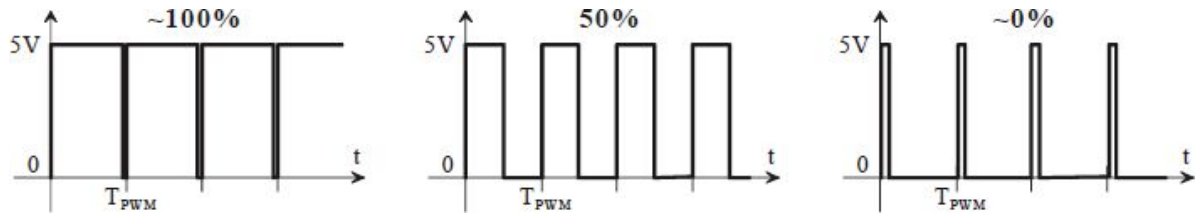


Fig. 12.50: The average value of a PWM wave is a function of its own duty-cycle.

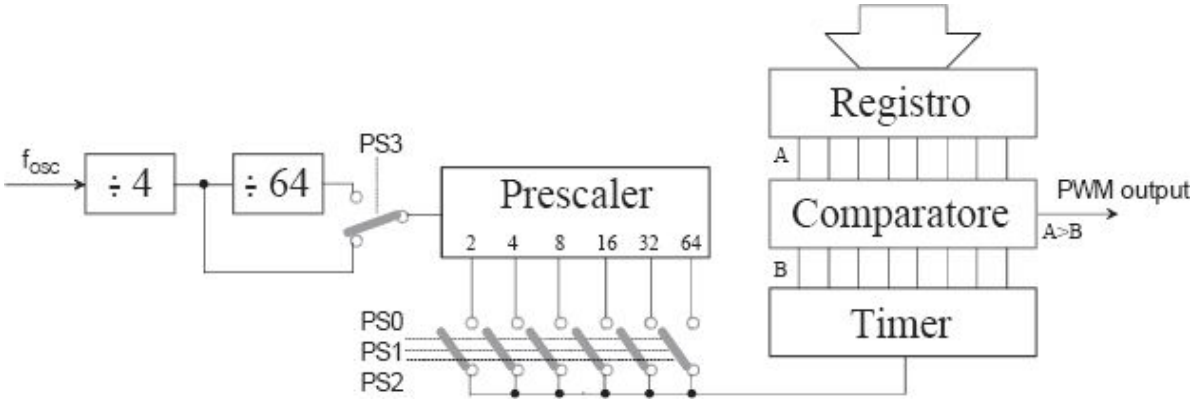


Fig. 12.51: PWM generation through a Capture&Compare unit.

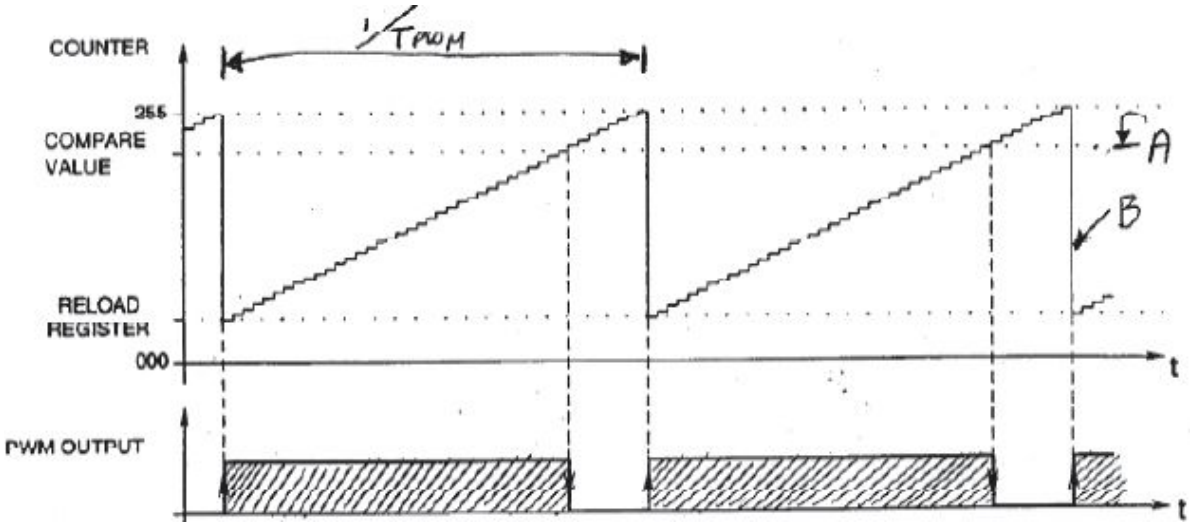


Fig. 12.52: PWM generation through a Capture&Compare unit.

12.12. Serial Comunication Devices

Serial ports are extremely useful because they allow the μC to talk and exchange data with other external devices, both on the same board (within the same board connected with serial EEPROMs, shift registers, display drivers, A/D converters, etc.). and on different boards (links to other boards). Using serial ports rather than parallel, let you save on both number of pins used and number of links.

There are basically three different protocols in the way to transfer data, the number of interconnections, and for the chance to own more than one master and carry arbitrage or not the bus. They are: Serial Peripheral Interface (SPI)

from Motorola, similar to Microwire from National Semiconductor; Inter-Integrated Circuit (I^2C) from Philips; and Serial Communication Interface (SCI), better known as the USART (Universal Synchronous/ Asynchronous Receiver/Transmitter).

12.12.1 Syncronous Serial Port SPI

You begin to consider the first and most simple: the SPI and his twin Microwire. Both have the advantage of being very simple, because it uses only three wires to transmit an 8-bit synchronous clock, but have the disadvantage of providing a single master and a unidirectional (from master to slave) clock line. The pins of interest are: Serial Data Output (SDO), Serial Data Input (SDI), Serial Clock (SCK). In addition, there may be further connected to Chip Enable (CE) in the case of multiple Slave.

Master is the μC that manages the clock and, ultimately, data transfer, the slave is the external device (up to another μC, which will always remain slaves). As shown in Fig 12:53, both are fitted with shift-register connected to form a closed loop, as if it were a single shift-register 16-bit shared between the two integrated. It's just the shift-register that allows the conversion of data from serial to parallel and vice versa: a parallel load 8-bit data out serially (led by the MSB) and are converted in parallel into the receiver. Basically it works like a ski-lift with the controls put in hand to the work ahead, but not to the downstream.

To initialize the synchronous serial port (SSP) in SPI mode you must specify whether the μC acts as a master (in this case SCK is an output) or as a slave (SCK becomes an input), polarity of the clock (rising or falling edge), clock frequency (for master) and finally mode of action of the slave (slave only). The data register is twofold: one to be transmitted in the shift register SSPSR and one in which received content is transferred SSPBUF (Fig. 12:53).

Once all 8 bits of interest has been received (ie, when the slave will be counted to 8 pulses), buffer flag is setted full and the interrupt flag (in the register PIR1<3>), data can then be taken by the μC core. Obviously it is preferable to use interrupts to avoid polling.

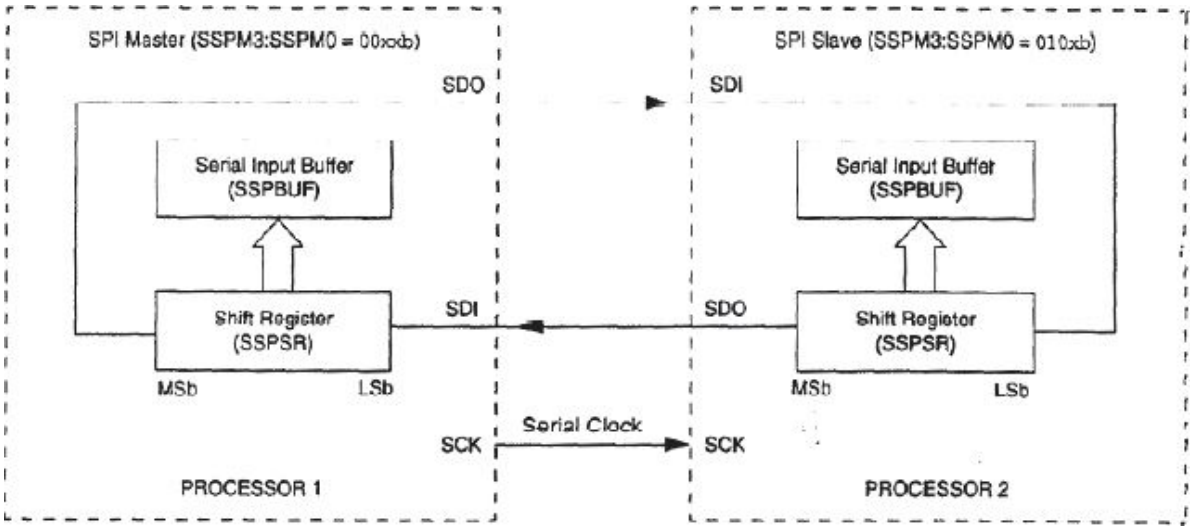


Fig. 12.53: SPI Connections between µC master (left) and slave device (right).

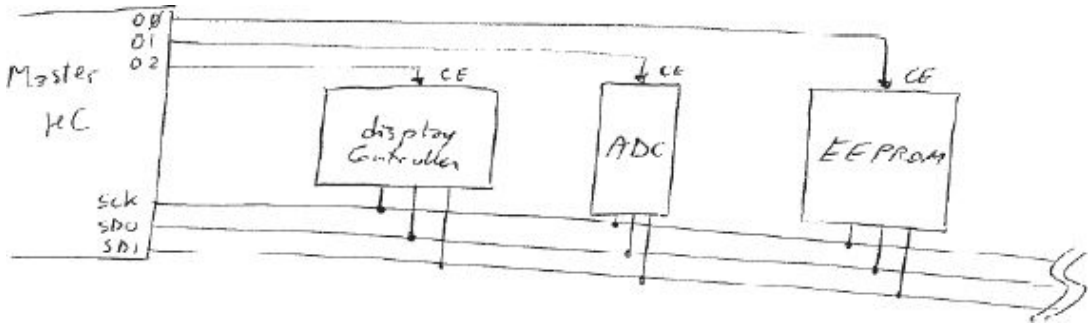


Fig. 12.54: SPI Connection to multiple slaves, through enable lines.

There are three different kind of data transmission; despite the master is unique, depending on who needs to read and/or write: • Write – master transmits useful data, while the slave answers useless data (dummy);

• Write and Read – Both master and slave transmits usefull data;

• Read – master sends dummy data, while the slave refers useful data.

It is therefore clear how this is a full-duplex protocol, because the data can be read and written simultaneously.

Master has always control over the operations: it decides the moment when to start transmission (enabling the clock), even when it's the slave that properly needs to (it can't actually "will" to). It's therefore necessary to write down a query slave protocol firmware, that lets you know when the data from the latter are 'good' and when 'dummy'. In addition, when a failure occurs, the slave does not have the opportunity to inform the master via the serial line, but must

wait until the latter questioned him, and then pass an appropriate emergency code. If slave is in SLEEP mode, the beginning of the data transfer will cause his ‘wake-up’.

In the case of multiple slaves, master will enable them one at a time, using additional chip enable (CE)lines, as shown in Fig 12:54.

12.12.2 Syncronous Serial Port I^2C

This protocol is more complex than the previous ones, because let you handle arbitration and synchronization of serial bus, using only two wires: Serial Clock (SCL) and Serial Data (SDA). Both lines are bidirectional because each device can act as a master, seizing the bus and sending the address of the device he wants to communicate, at this point, any other request of control of the bus will be rejected. It is therefore a serial communication system suitable for the connection of “smart” devices, able to handle the traffic on the bus and to recognize the call to its own address. Chip-enable lines are no more needed as for the SPI protocol, there are just two connections, whatever the number of devices, in contrast to what happens with the SPI.

Both lines have to be open-collector with an external pull-up resistor, so sizing is carried out according to the number of devices (both master and slave) connected. In this way, each pin relating the line can be an output (open drain) or input, depending on which device has control of the bus. Figure 12:55 shows the structure of serial bus I^2C.

Problems lie in the management of the protocols, it requires a complex hardware. A low-end µC without the interface for this protocol can communicate only through a firmware created ad hoc. Even those peripherals that will always act as slave have to be ‘smart’, to ensure addressing management by the master and the proper response to the caller.

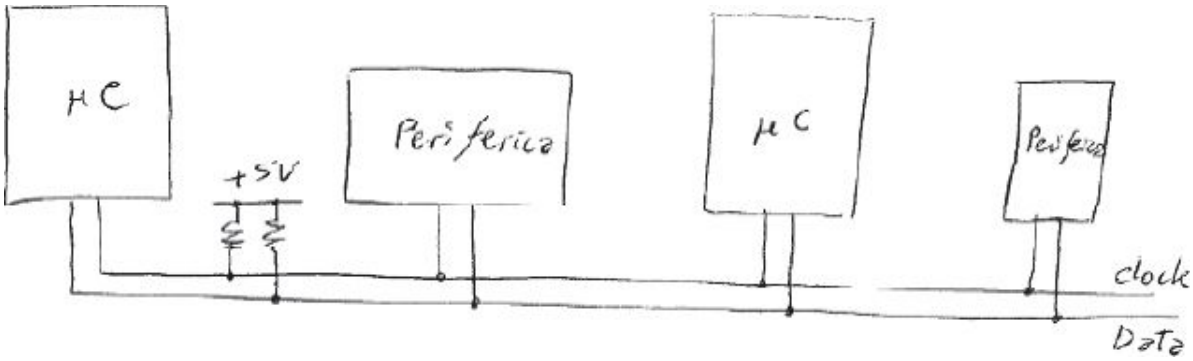


Fig. 12.55: I^2C serial communication: just two wires can handle multiple masters and multiple slaves.

I^2C standard provides a maximum speed of 100kbps (400kbps for the new fast-mode). Maximum lines load capacitance is 400pF and this limits the maximum number of devices connected. When there is no data transfer, both clock line (SCL) and data (SDA) are taken at high level by the external pull-up resistors. Request for use of the bus, thus the START condition, occurs when a device set the SDA down while SCL clock is high. Later, when the clock ‘raises up and then goes down’ data is stored (note: in this period the data should not change). The release of the bus, thus the STOP condition, occurs when SDA rises with high clock.

After taking control of the bus, the master transmits in the next 10-bit the address of device with which it wishes to communicate, then releases the bus and waits for the acknowledge from the latter. There are two formats to send the address of the device that the master wants to contact. The easiest is bit 7-bit followed by a R/W bit, which is used to determine if the master wants to read or write to the device. The second format instead uses a 10-bit address, first 5 bits (11110) are transmitted, then address 2-bit, R/W bit, and finally, after receiving the acknowledge, the remaining address bits.

All data has to be transmitted as byte (8 bits), there aren’t limitations on the number of bytes that can be transmitted for each transfer. After the master has sent each byte, the slave receiver must always generate an acknowledge bit /ACK. If not, the master aborts the transfer and generates the STOP condition.

There are two way to transfer data:

1) master send and slave receive; 2) master receive and slave send.

If the master is receiving data, it will generate the acknowledge bit at each received byte, except the last, for which no confirmation will be sent, so as to indicate the slave to interrupt the transmission. Consequently, the slave releases the SDA line, allowing the master to generate the STOP condition.

If the slave needs to delay the transmission of the next byte, it is sufficient to maintain the SCL line low, thus forcing the master state to wait. This technique can also be used for each bit.

When more than one master wants to transfer data at the same time an arbitrate on the SDA line has to be used during the period in which the line SCL remains high. The master which transmits a high when the other sends a low loses arbitration and immediately after finishing the reading of the bits composing the byte, it must switch to slave (because the new master may direct it as slave).

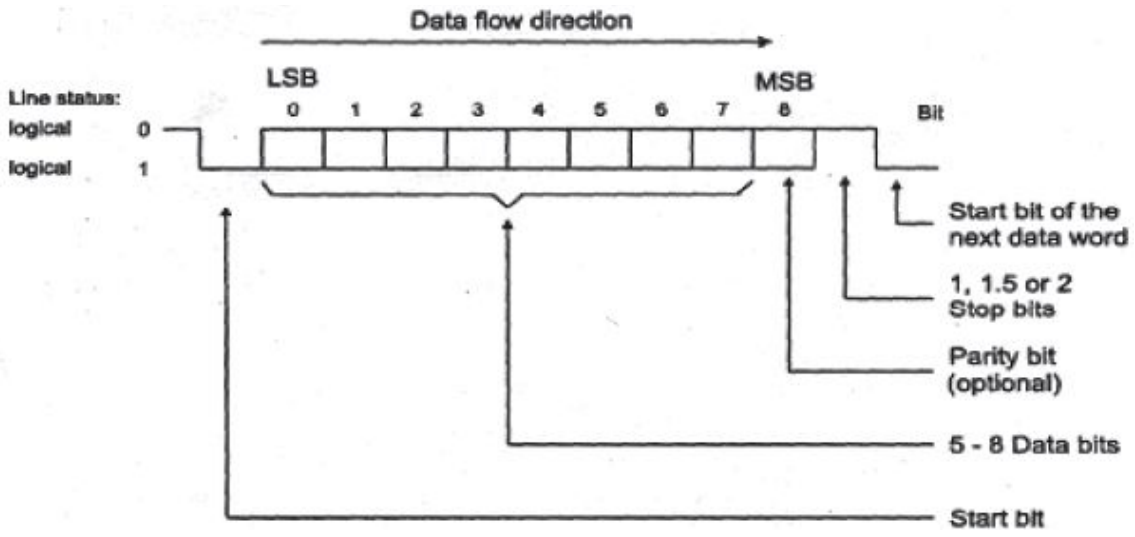


Fig. 12.56: Bytes structure of the USART asynchronous transmission.

12.12.3 USART port

SCI protocol (Serial Communication Interface) is also known as the USART (Universal Synchronous/Asynchronous Receiver/Transmitter). While the standard I^2C and SPI are used primarily for connections intrasboard, the SCI is intended for the connection interboard. SCI/USART protocol can be configured in full-duplex asynchronous, able to communicate with peripherals such as monitors and PCs, or half-duplex synchronous, to communicate with ADC, DAC, serial EEPROM, etc.. (both as master or slave).

Typical examples are RS232 serial protocols, or the industry standard 485 for both asynchronous and synchronous transmissions. From CMOS levels 0V to +5V, standard 232 requires the use of +12 V and -12V, through appropriate integrated level translators.

Transmission clock frequency (expressed in baud, thus bit/sec) is imposed by the firmware, along with other settings (polarity, presence of the parity bit, broadcast on 7 or 8 bits, etc..). In particular, to determine the baud-rate of the synchronization generator, a record with an appropriate value, so as to adequately divide μC clock f_{osc}, has to be set. In this regard, there are tables in the data-book of the various μC. USART asynchronous mode uses the standard NRZ (no return-to-zero) as transmission protocol, characterized by the structure shown in Fig 12:56: 1 start bit, 8 or 9 data bits, 1 stop bit. In some hardware where the parity bit (equal to 1 if the number of 1 to send is odd) is not implemented, it can be created via firmware and thus stored as the ninth bit.

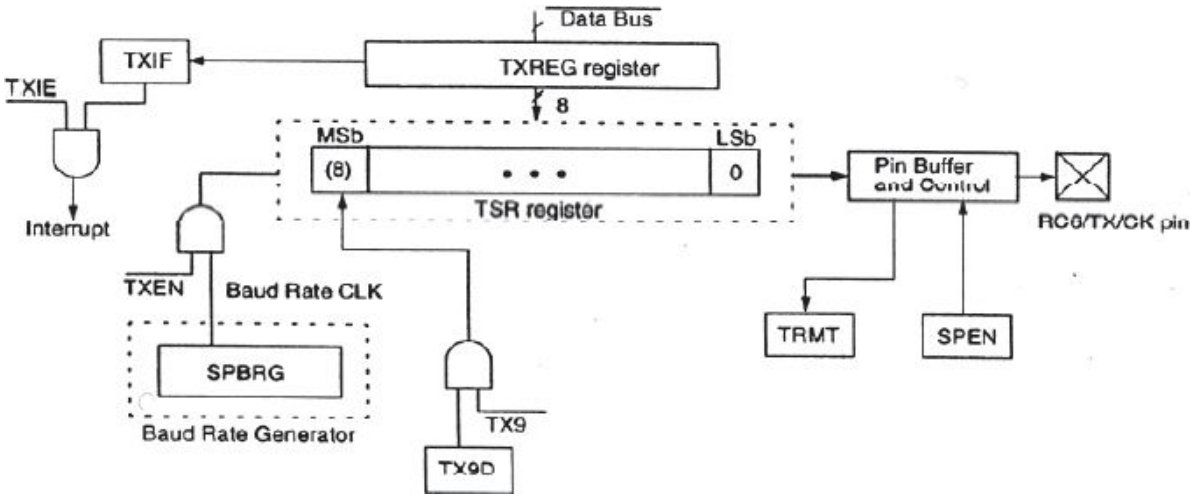


Fig. 12.57: Block diagram of the USART transmission device in PIC16xx.

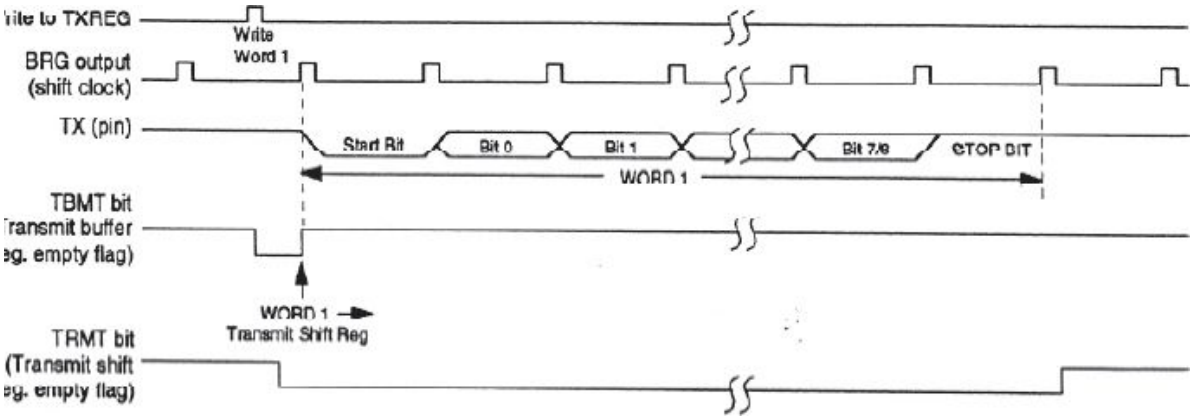


Fig. 12.58: Timing during a USART asynchronous transmission.

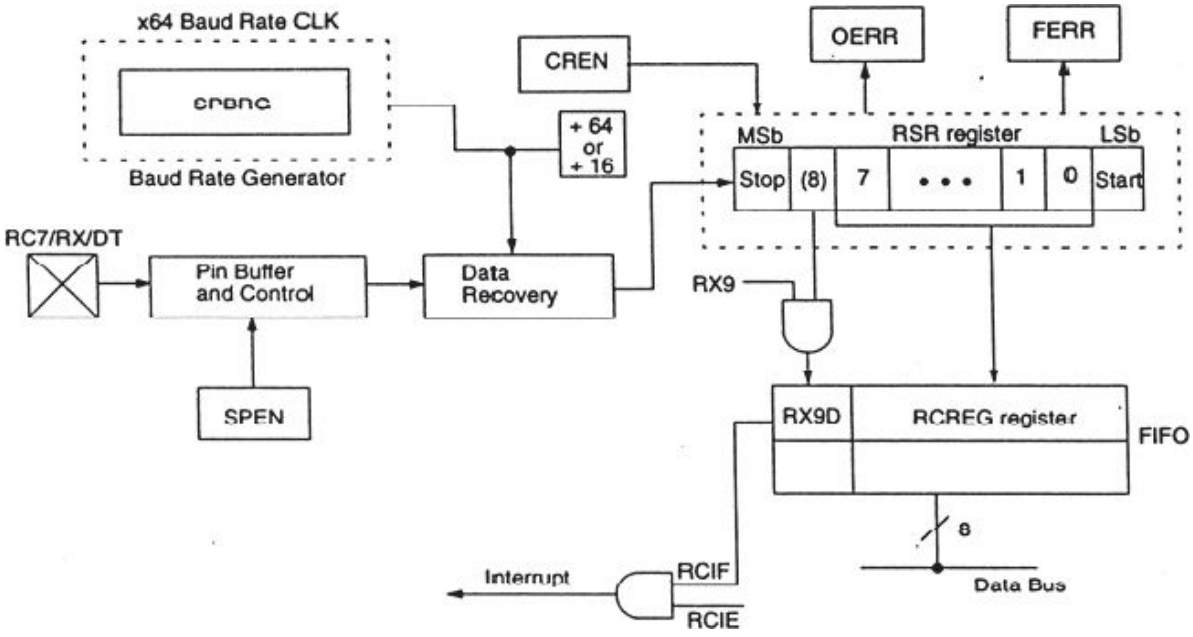


Fig. 12.59: Block diagram of the receiving USART device in PIC16xx.

Transmitter and receiver implemented in µC USART are independent, even if they use the same data format and the same baud rate. Figure 12:57 shows the block diagram of the Microchip USART Transmitter unit. The basic element is the shift-register TSR (Transmit Shift Register) tapping or inserting data in the register TXREG. Figure 12:58 shows a typical asynchronous transmission.

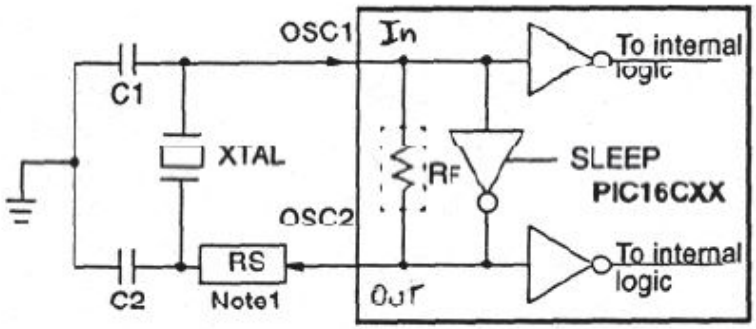
The receiver USART block picks up the asynchronous RX data, it reconstructs it in a data recovery block (a high speed shifter operating at a speed of 16 times baud rate) and submit it to the serial shifter (which operates at bit rate). This way, data present at the RX pin is sampled three times (compared to the baud rate), to determine if there is a low level or a high level in order to discard uncertain cases. The structure of the Microchip USART peripheral is shown in Fig 12:59.

12.13. Oscillator

Oscillator is the circuit that generates the clock required by the µC operation. Some µC can work with very low frequencies, up to the continuous (in the so-called one-step operation or DC), but there is usually a minimum f_{osc}, below which no one can get, and a maximum, which can not be exceeded sentence the proper functioning of µC to be guarantee. The lowest frequency is constrained by the presence or absence of dynamic memory within the µC, and thus the

need for a periodic refresh with a high enough rate. The maximum clock frequency is instead tied to the limitations in the switching speed of the internal logic.

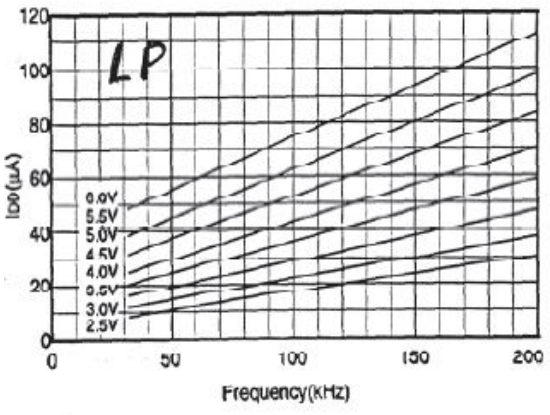
Circuits needed to realize the positive feedback in order to ensure the swing of a simple external circuit are usually already included inside the μC. Typically there are two pins OscIn and OscOut, as shown in Figure 12.60. These can connect an external quartz crystal in order to obtain the desired frequency of oscillation. It ‘s always recommended to include a parallel-cut, that is, two capacitors to determine the exact frequency of oscillation. If a stable oscillation has to be guarantee, precise and high-value (of the order of MHz) a quartz has to be used, otherwise if you want to work at lower frequencies is sufficient to use an RC network, obviously losing accuracy and stability.



See Table 14-1, Table 14-2, Table 14-3 and Table 14-4 for recommended values of C1 and C2.

Note 1: A series resistor may be required for AT strip cut crystals.

Fig. 12.60: Internal configuration of the PIC16xx oscillator when an external crystal is connected.



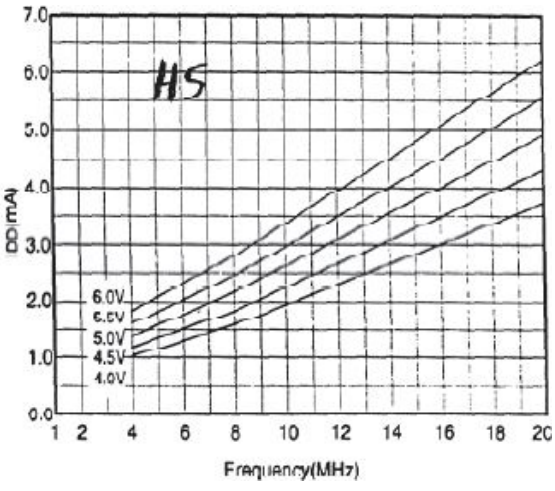


Fig. 12.61: Evolution of the current absorbed by the µC at various clock rate, in extreme cases of LowPower (left) and HighSpeed (right).

The advantage of a lower clock frequency is of course a lower power request (which goes up with increasing frequency in every CMOS network) and a lower generation of radiated interference noises. Microchip divides the range of clock frequencies into three bands: LP (low power application, up to 200kHz), XT (intermediate) and HS (high speed, up to 20MHz). Power request as a function of frequency and voltage (when the internal devices are not working) is shown in Figure 12.61.

Clock can also be supplied externally, because already available on the board or to ensure synchronization with other circuits. In this case we will only use the pin OSC1, while from OSC2 $f_{osc}/4$ will be taken, in order to power other external integrated, which require a clock synchronized with µC, as shown in Figure 12.62.

External oscillator can once again be made with a quartz in resonance both parallel and series, as shown in Figure 12.63. For applications where timing is not critical and it is not necessary to achieve high speed, you can use in place of quartz a simple and inexpensive external RC network. To do this the µC has to be seted appropriately, so that it modifies the internal hardware configuration of the pin OscIn, as shown in Figure 12.64. When tensions rise, the internal buffer switch and short circuit the capacitor. This regard, the manufacturer provides tables with the optimal values of R_{ext} e C_{ext} at the desired clock frequency.

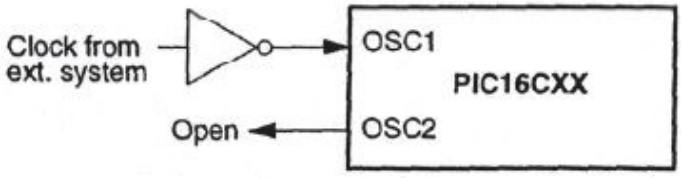


Fig. 12.62: Oscillator configuration when the clock is supplied from an external source.

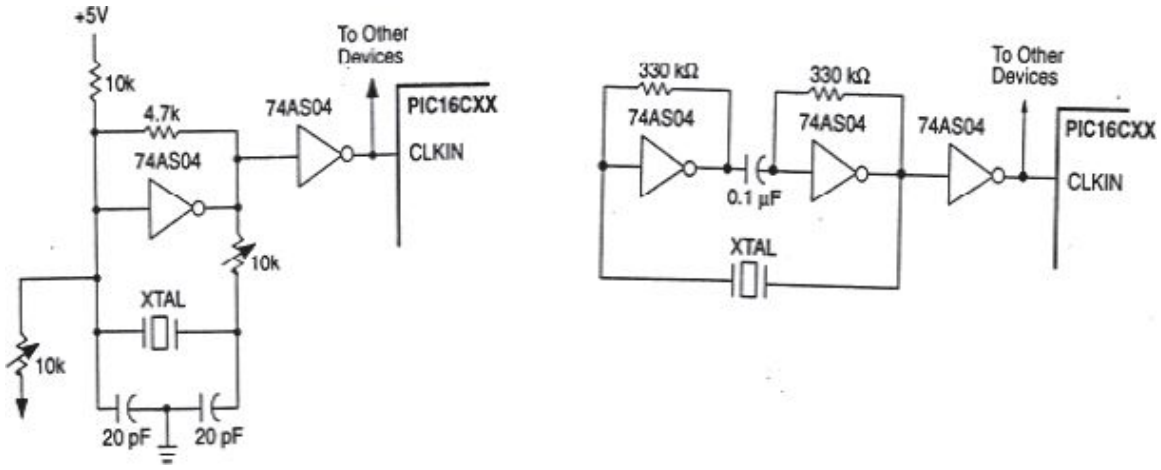


Fig. 12.63: Oscillating circuits external to µC, with crystals in parallel resonance (left) and series (right).

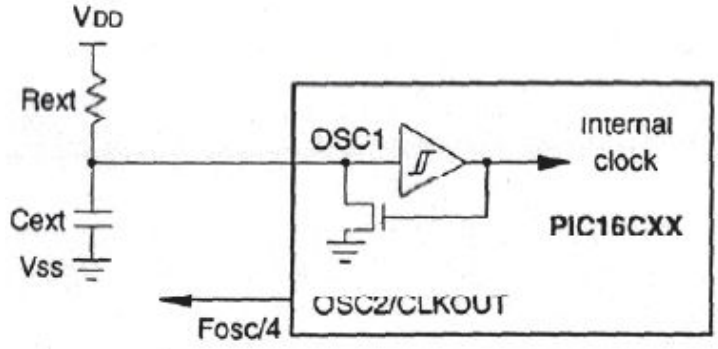


Fig. 12.64: Oscillator configuration when the clock is generated by an external RC network.

If an RC network is uses, the oscillation frequency is not stable (unlike what happened with the quartz) as this will vary with the supply voltage, temperature and depends on the type of capacitor used (changing capacitor, the parasitic capacitance change between the capacitor and printed circuit board, especially for low values of C_{ext}).

12.14. RESET

Reset is used for proper initialization of the sequential machine represented by the core of the µC. At first, when the µC is powered, it has to start from the first line of code and then run all the initializations. There are more reset sources, because the request can be supplied from outside or be generated internally in the µC itself. The main sources of reset are as follow: • Power-on Reset;

- Master-Clear Reset, during normal operation, using the external pin /MCLR;
- Master-Clear Reset, during SLEEP mode, again through pin /MCLR;
- Watchdog-timer reset, when watchdog timer gets an overflow (WDT);
- Brown-Out Reset, if the power supply has high fluctuations.

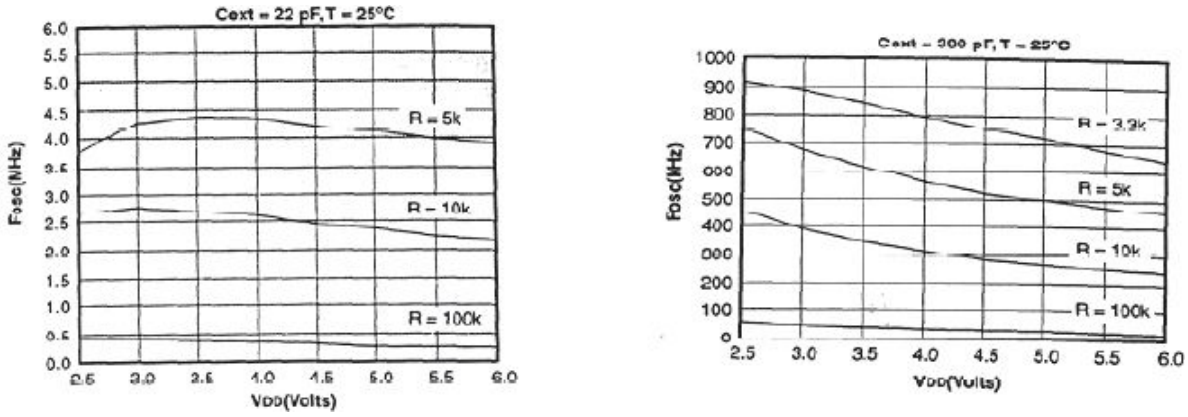


Fig. 12.65: Values suggested by Microchip for the components of Fig 12.64.

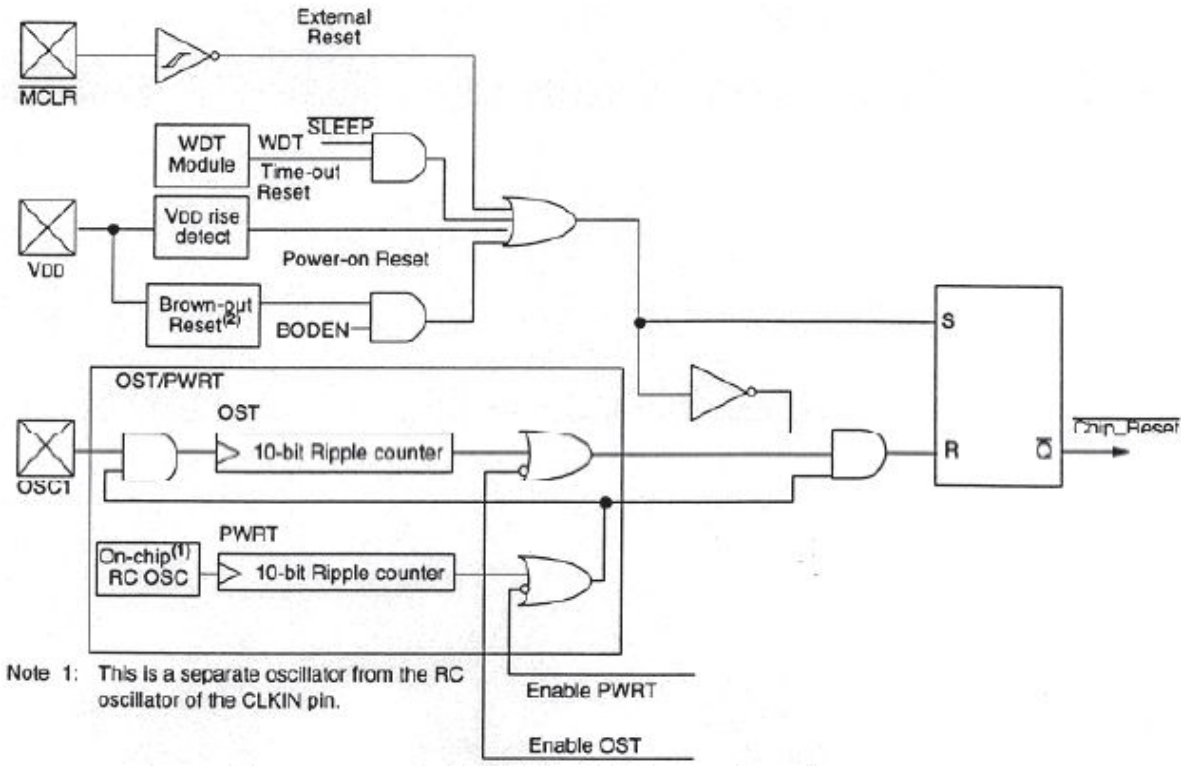


Fig. 12.66: Simplified block diagram of the internal reset units in the PIC16xx.

To analyze all modes of reset is useful to analyze the simplified block diagram shown in Figure 12.66., dedicated to all of these functions. The final flip/flop prevents the pulse to be too short and therefore not be properly interpreted by the core. Those individual unit that Set the flip/flop will therefore turn it off, working on the Reset. We will now consider in more detail various kinds of reset.

Master Clear

This reset can occur during normal operation of the firmware, or during SLEEP., the designer can attach a simple RC network to the pin dedicated to this purpose to ensure the automatic reset on power, as shown in Figure 12.67.

The external diode is used to help the internal protection diode to the MCLR pin: when power is switched off the charged capacitor is at a higher voltage than the μC power supply, eventually damaging the integrated in this way, with this extra diode the capacitor is shorted to ground quickly. Resistance R1 is designed to limit the current entrance to the port of μC in the presence of electrostatic discharge.

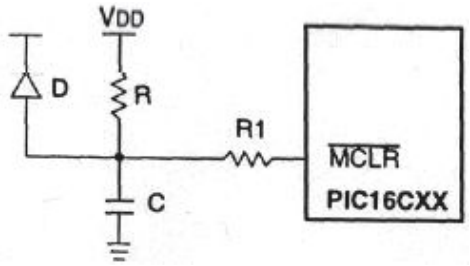


Fig. 12.67: Use of the Master Clear to reset the μC at the power on.

Power-on Reset

In addition to the external network just seen, μC have normally an internal circuit that generates a reset every time the voltage is applied (power-on). There is a particular inner block that records the rising edge of V_{DD} and reset the chip. In practice there is a reset every time the power crosses a threshold near $½V_{DD}$.

Power-on Timer and Oscillator Start-up timer

Once power supply is stable, we should wait until the clock begins to oscillate and thus for it to stabilize at the desired frequency. For this reason, some μC include within them the Power-on Reset Timer (PWRT), which has an

internal RC oscillator that counts a delay of 72ms, then allowing the µC to exit the reset state in which it was forced. Another possibility to exit from the reset can be delaying it further through the Oscillator Start-Up Timer (OST) block, which after counting 1024 clock pulses from the external oscillator tells the core to release and exit the reset state.

Brown-Out

The term “brown-out” means the momentary power fluctuation that can be followed by the real “black-out”. If the power suffer significant fluctuations (about 1V), the power-on reset would not intervene, despite the errors µC firmware may have suffered (due to failure to read or write some bits), hardware (piloting or false readings from devices external) or software (some bits of the registers or program memory may have changed). It is therefore necessary to have a brown-out detect circuit which operates when the supply voltage fluctuates by a certain amount of the nominal value. In PICs, for example, the reset occurs when V_{DD} drops below 3.8V for longer than or equal to 100µs and the reset will remain active for another 72ms from the last time the power is risen above 4.2V, as shown in Figure 12.68.

12.15. SLEEP-MODE

This mode of operation, also known as “Power-Down”, is invoked with the instruction SLEEP: the µC goes into a low-power state, where most of the internal devices are disabled or even disconnected from the power. Sometimes, even the clock oscillator is disabled (see Figure 12.60) and the whole core stops.

To further reduce power consumption, the designer must ensure that the outputs are placed high (V_{DD}) or low (GND) so as to avoid even the dissipation of pull-up or pull-down external resistors, or in any way on any external load. Furthermore, it’s better to avoid that the µC inputs to remain floating, as they may determine voltage levels intermediate between 0 and 5V and then set to cross-conduction the internal CMOS port. It ‘s always better to use external pull-up or pull-down.

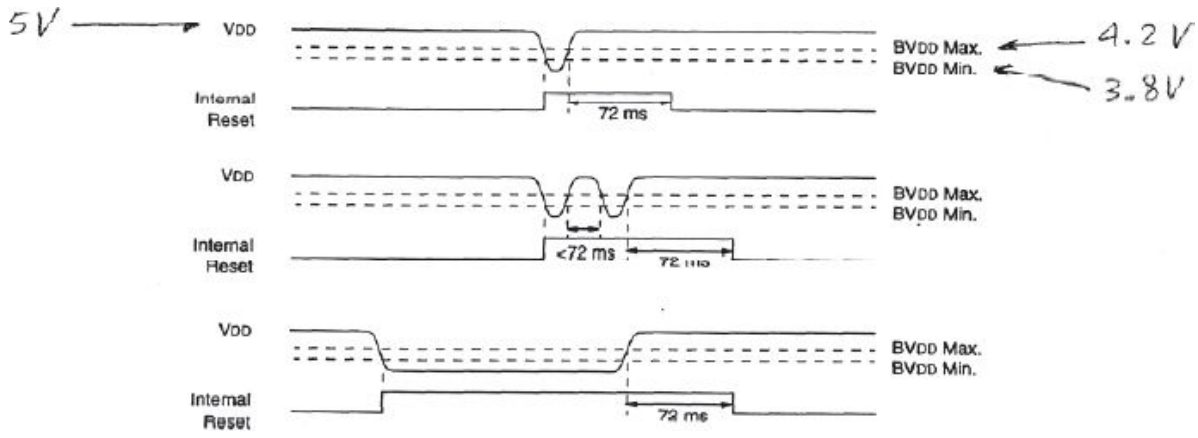
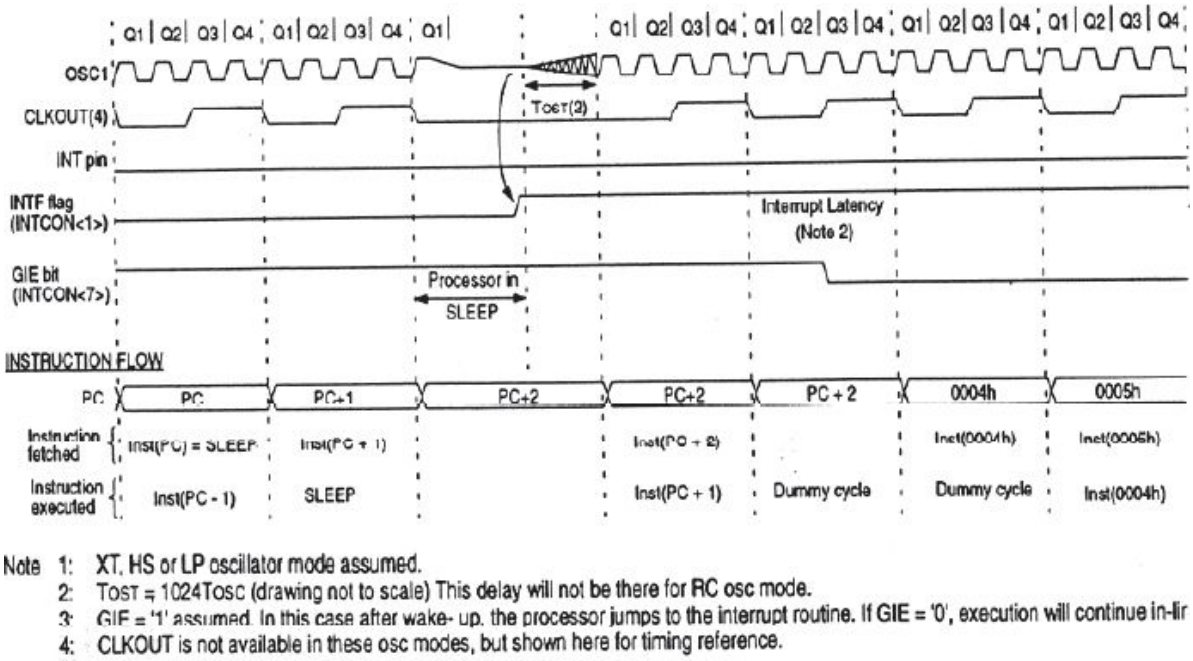


Fig. 12.68: Possible situations that trigger the brown-out reset.



Note 1: XT, HS or LP oscillator mode assumed.
2: TOST = 1024TOSC (drawing not to scale) This delay will not be there for RC osc mode.
3: GIE = '1' assumed. In this case after wake-up, the processor jumps to the interrupt routine. If GIE = '0', execution will continue in-lir
4: CLKOUT is not available in these osc modes, but shown here for timing reference.

Fig. 12.69: Wake-up from the SLEEP mode through interrupt.

The exit from this state can be done either through external reset, acting on the MCLR pin, or through interrupt (eg by watchdog). Once dropped out of the sleep-mode, as shown in Figure 12.69., the µC won't be ready immediatly but you should still wait a few time for the stabilization of the clock oscillator.

12.16. Short Review

Given the wide spread that the μC industry had in recent years in many sectors, for both industrial and consumer electronics, many manufacturers are entering the market with new products. In fact, now the μC are now commonly used in many electronic products and all the manufacturers, both the younger and those already established on analog products, have begun to produce them.

Most common are Motorola, National Semiconductor, Texas Instruments, STMicroelectronics, Microchip, ATMEL, Siemens, NEC, Intel.

12.16.1 STMicroelectronics

ST7 μC family is an example of mid-range that has improved the features of the low-end ST6 family, introducing a multiplier on-chip, creating a proper table of interrupt vectors and by increasing the stack. The internal architecture has been improved in order to speed up processing and exchanging data between different devices.

There’s an area in the memory map allocated to the Reset vectors, each of which contains the address to jump to the routine dedicated to process, there’s thus no longer the need for a firmware polling. The stack has been enhanced in order to handle routine or nested interrupts calls. It’s always to the designer not to exceProf. Fd the allowed range.

I/O ports have as the main feature to be very versatile, allowing bi-directional data transmission, often providing an internal pull-up, possibility to generate interrupts and to acquire analog signals.

An extended instruction set (for example, the presence of the multiplication operation) and an effective addressing modes, can streamline the code and faster execution of the program. Some μCs offer the availability to have an outgoing address and data bus for external memory to improve the efficiency of the project.

ST72101/ST72212/ST72213

8-BIT MCU WITH 4 TO 8K ROM/OTP/EPROM, 256 BYTES RAM, ADC, WDG, SPI AND 1 OR 2 TIMERS

PRELIMINARY DATA

- User Program Memory (ROM/OTP/EPROM): 4 to 8K bytes
- Data RAM: 256 bytes, including 64 bytes of stack
- Master Reset and Power-On Reset
- Run, Wait, Slow, Halt and RAM Retention modes
- 22 multifunctional bidirectional I/O lines:
  - 22 programmable interrupt inputs
  - 8 high sink outputs
  - 6 analog alternate inputs
  - 10 to 14 alternate functions
  - EMI filtering
- Programmable watchdog (WDG)
- One or two 16-bit Timers, each featuring:
  - 2 Input Captures
  - 2 Output Compares
  - External Clock input (on Timer A only)
  - PWM and Pulse Generator modes
- Synchronous Serial Peripheral Interface (SPI)
- 8-bit Analog-to-Digital converter (6 channels) (ST72212 and ST72213 only)
- 8-bit Data Manipulation
- 63 Basic Instructions
- 17 main Addressing Modes
- 8 x 8 Unsigned Multiply Instruction
- True Bit Manipulation
- Complete Development Support on PC/DOS-WINDOWS™ Real-Time Emulator
- Full Software Package on DOS/WINDOWS™ (C-Compiler, Cross-Assembler, Debugger)

PSDIP32

CSDIP32W

SO28

(See ordering information at the end of datasheet)

Device Summary

Features	ST72101G1	ST72101G2	ST72213G1	ST72212G2
Program Memory - bytes	4K	8K	4K	8K
RAM (stack) - bytes	256 (64)			
16-bit Timers	one	one	one	two
ADC	no	no	yes	yes
Other Peripherals	Watchdog, SPI			
Operating Supply	3 to 6 V			
CPU Frequency	8MHz max (16MHz oscillator)			
Temperature Range	- 40°C to + 85°C			
Package	SO28 - SDIP32			

Fig. 12.70: Front page of ST7 µC family.

12.16.2 Motorola

Motorola family of µC HC08 to HC11 is very large and powerful. It created a standard in the market, so that now its assembler is used and made compatible by different manufacturers. There are even conversion programs to adapt a code written for Motorola in order to translate it into that of another manufacturer, in order to facilitate the migration of users to new products.

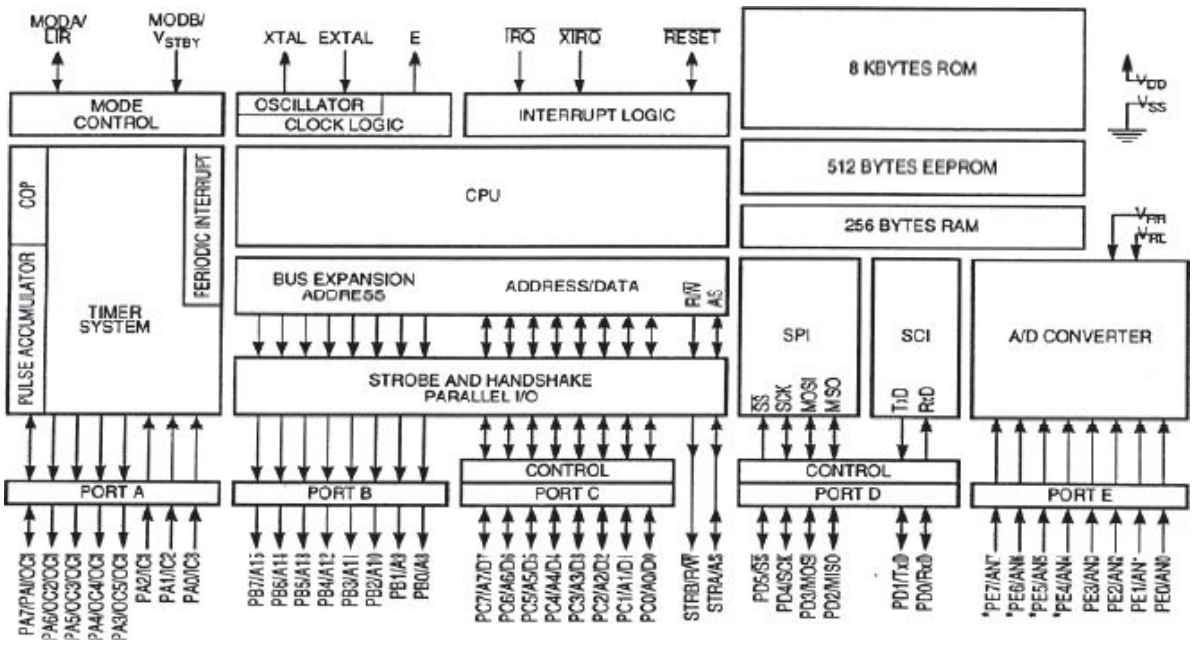
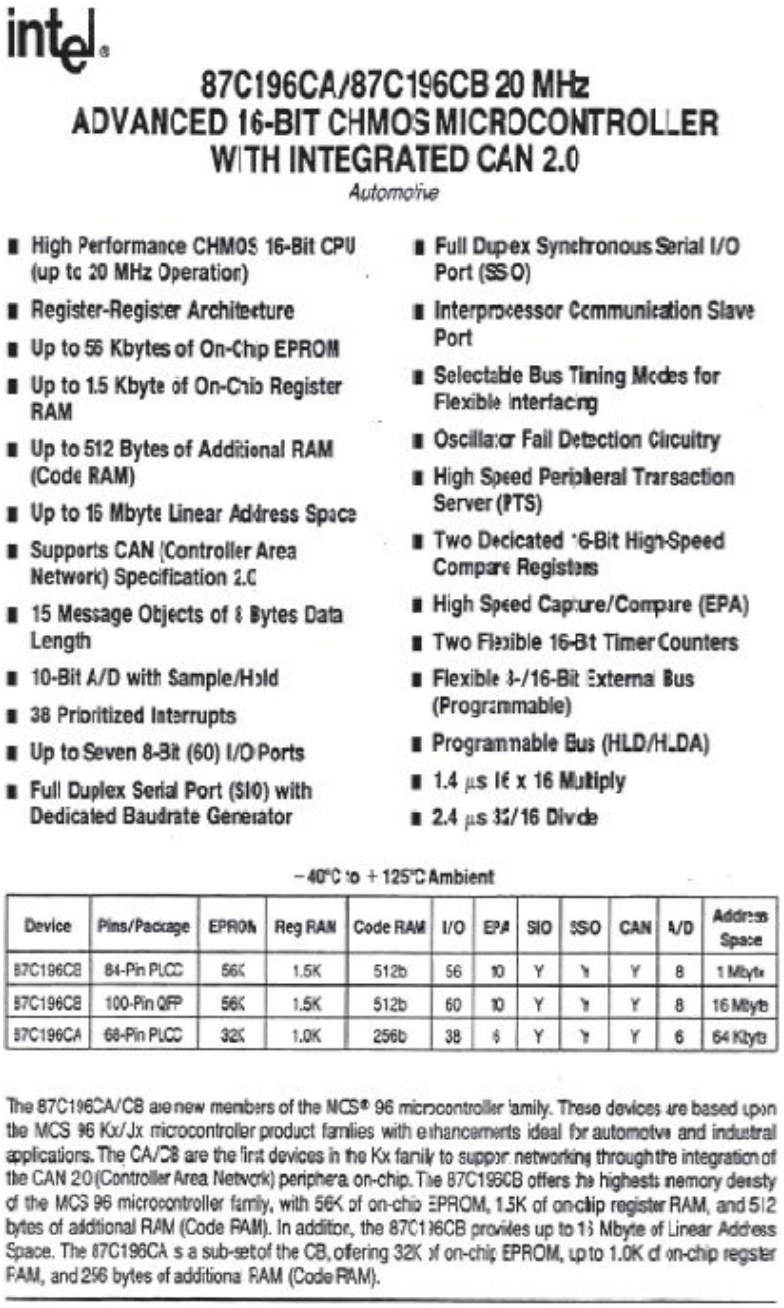


Fig. 12.71: Internal architecture of Motorola HC11 µC family.

12.16.3 Intel

Intel µC family, like those of other manufacturers, offers many differences other than evolved internal devices. As example CAN-bus communications, that is now a standard in industrial and automotive manufacturing.



intel

87C196CA/87C196CB 20 MHz
ADVANCED 16-BIT CHMOS MICROCONTROLLER
WITH INTEGRATED CAN 2.0

Automotive

- High Performance CHMOS 16-Bit CPU (up to 20 MHz Operation)
- Register-Register Architecture
- Up to 56 Kbytes of On-Chip EPROM
- Up to 1.5 Kbyte of On-Chip Register RAM
- Up to 512 Bytes of Additional RAM (Code RAM)
- Up to 16 Mbyte Linear Address Space
- Supports CAN (Controller Area Network) Specification 2.0
- 15 Message Objects of 8 Bytes Data Length
- 10-Bit A/D with Sample/Hold
- 38 Prioritized Interrupts
- Up to Seven 8-Bit (60) I/O Ports
- Full Duplex Serial Port (SIO) with Dedicated Baudrate Generator
- Full Duplex Synchronous Serial I/O Port (SSIO)
- Interprocessor Communication Slave Port
- Selectable Bus Timing Modes for Flexible Interfacing
- Oscillator Fail Detection Circuitry
- High Speed Peripheral Transaction Server (PTS)
- Two Dedicated 16-Bit High-Speed Compare Registers
- High Speed Capture/Compare (EPA)
- Two Flexible 16-Bit Timer/Counters
- Flexible 8-/16-Bit External Bus (Programmable)
- Programmable Bus (HLD/HLDA)
- 1.4 μs 16 x 16 Multiply
- 2.4 μs 32/16 Divide

−40°C to +125°C Ambient

Device	Pins/Package	EPROM	Reg RAM	Code RAM	I/O	EPA	SIO	SSIO	CAN	A/D	Address Space
87C196CB	84-Pin PLCC	56K	1.5K	512b	56	10	Y	Y	Y	8	16 Mbyte
87C196CB	100-Pin QFP	56K	1.5K	512b	60	10	Y	Y	Y	8	16 Mbyte
87C196CA	68-Pin PLCC	32K	1.0K	256b	38	6	Y	Y	Y	6	64 Kbyte

The 87C196CA/CB are new members of the MCS® 96 microcontroller family. These devices are based upon the MCS 96 Kx/Jx microcontroller product families with enhancements ideal for automotive and industrial applications. The CA/CB are the first devices in the Kx family to support networking through the integration of the CAN 2.0 (Controller Area Network) peripheral on-chip. The 87C196CB offers the highest memory density of the MCS 96 microcontroller family, with 56K of on-chip EPROM, 1.5K of on-chip register RAM, and 512 bytes of additional RAM (Code RAM). In addition, the 87C196CB provides up to 16 Mbyte of Linear Address Space. The 87C196CA is a sub-set of the CB, offering 32K of on-chip EPROM, up to 1.0K of on-chip register RAM, and 256 bytes of additional RAM (Code RAM).

Fig. 12.72: Frontpage of μC Intel 87C196 data-sheet.

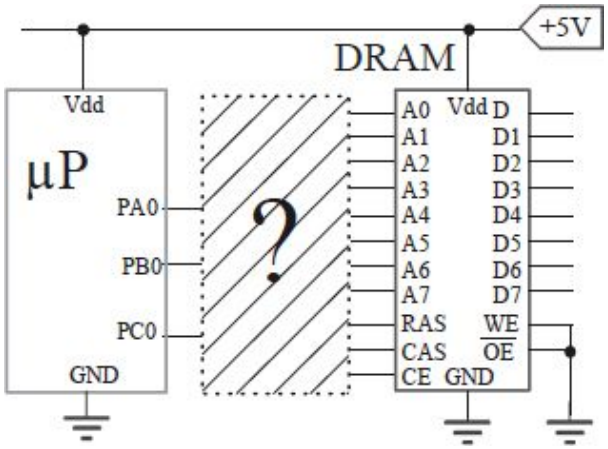
12.17. EERCISES

Exe. 1

The microcontroller uses only three lines to pilot a **serial** *64kbytes* dynamic RAM with multiplexed address.

a) Project the network with MSI commercial circuits.

b) Given a refresh time of *100ms* set the minimum scan frequency of the piloting network.

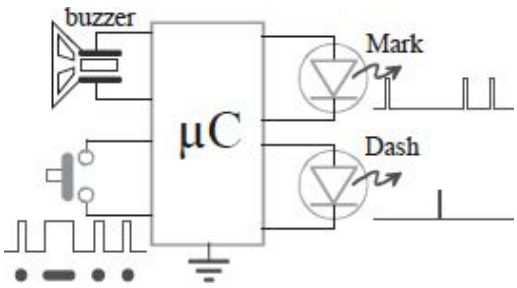


Exe. 2

A Morse code recognizer has to be achieved throu μC. At each correct press (*mark* longer than *500ms* or *dot* lower to *300ms*) the corresponding LED should light up for *300ms*, in case of wrong time the buzzer will sound (to be driven with rectangular pulses at about 1 kHz) for about *1s*.

a) Project the circuital scheme.

b) Write down firmware flow chart, and evidence how to manage the timing.

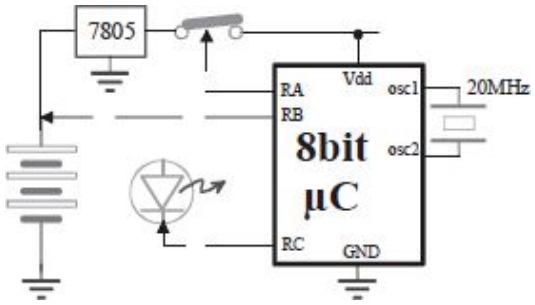


Exe. 3

Among other things µC has to monitor the *12V* power supply, shut down if the voltage drops below *6V* and switch back up when it raises over *7V*.

a) Choose the components and project the circuit to open and close the switch.

b) Drow the flow chart of the operations.

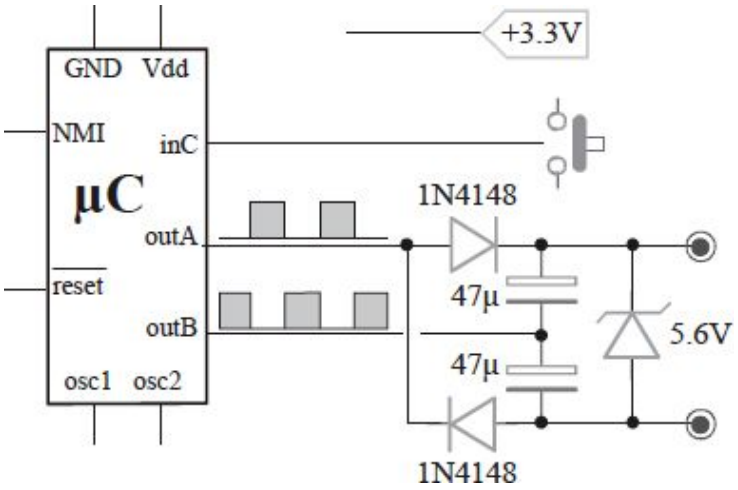


Exe. 4

Roject the charge pump with µC so that by alternate pressure of the button it will enable or disable the *1kHz* output shown in figure.

a) Draw the output waveforms and design all the electronics around the µC, used by the highlighted pins.

b) Write down the flow chart of operations that the µC has to do.

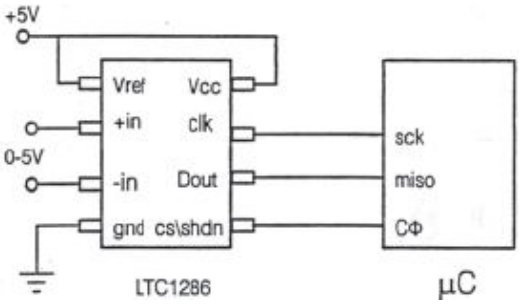


Exe. 5

Every *15min* the *12bit* ADC has to be read with a serial *8bit* µC clocked at *4MHz*.

a) Project the firmware with evidence for the *sleep* **wake-up** and counter register.

b) Project a circuit without µC but using only MSI, in order for any assertion of the *ACQUISITION* input to get the ADC reading and store it in a *32k x 8bit* SRAM.

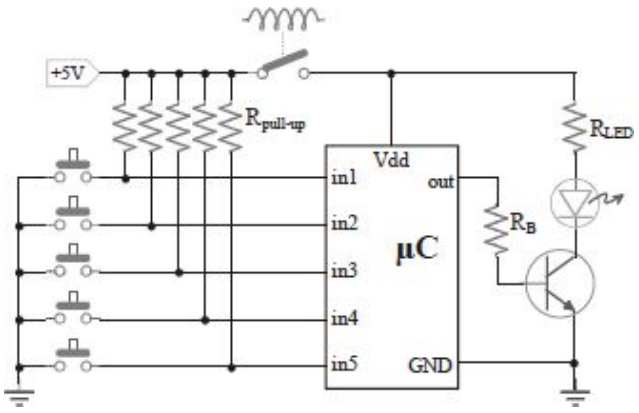

+5V
Vref
Vcc
+in
clk
0-5V
-in
Dout
gnd
cs\shdn
sck
miso
CΦ
LTC1286
µC

Exe. 6

In this infrared transmitter, you want at the touch of a button, the relay sends power to the µC (to conserve the battery). Then, the µC will make a number of pulses to the LED equal to the button pressed.

a) Project an simple external network to close the switch.

b) Sadly, we realize that in this way the µC burns: explain why and solve the problem.

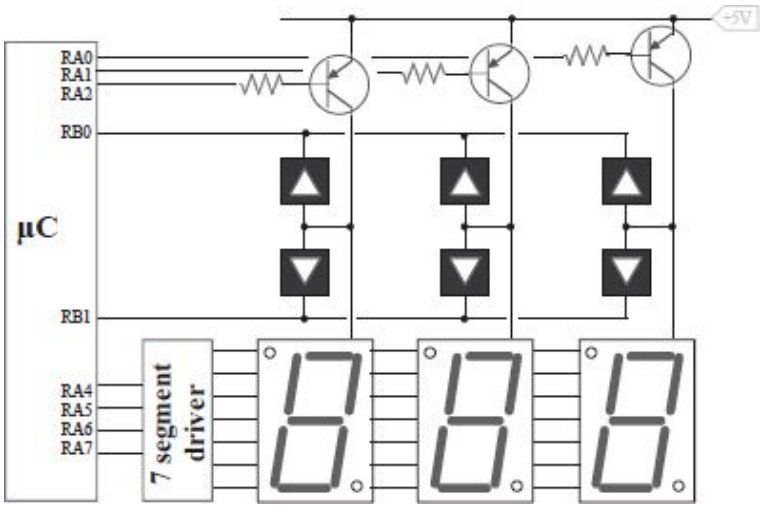


Exe. 7

Multiplex the three display (digits from “0” to “9”) and three pair of button *up* and *down*.

a) Whrite down the µC operation flow-chart (nitializations, cycles, reading, RAM writing, ...).

b) Show up how to handle the µC sleep and ke-up, when a key is pressed.



Exe. 8

A *4Mbit* DRAM (organized with byte) has to be read *8+8* I/O lines.

a) Project the full circuit and drow µC timing to pilot the memory.

b) Whrite down reading routine in order to use the address in registers *Ra1*, *Ra2* and *Ra3*, store the data read in *Rread* and withdraw the data to be written by *Rwrite*.

c) Whrite down refresh routine.